PISA 2015

Kristina Reiss, Christine Sälzer,
Anja Schiepe-Tiska, Eckhard Klieme,
Olaf Köller (Hrsg.)

Eine Studie zwischen Kontinuität und Innovation

WAXMANN

© Waxmann Verlag GmbH

Kristina Reiss, Christine Sälzer, Anja Schiepe-Tiska,
Eckhard Klieme, Olaf Köller (Hrsg.)

PISA 2015

Eine Studie zwischen Kontinuität und Innovation

Waxmann 2016
Münster · New York

© Waxmann Verlag GmbH

Redaktionsschluss: 03.11.2016

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.

Print-ISBN 978-3-8309-3555-0
E-Book-ISBN 978-3-8309-8555-6

AA @W+Q? 2,.
Rfgq umpi gq jgaclqcb slbcp_Apc_rgtc Amk kmlq ?rrpg`srgml+Qf_pc?jgic 2,.
Glrcpl_rgml_j Jgaclqc,

CC BY SA

Waxmann Verlag GmbH, 2019
Steinfurter Straße 555, 48159 Münster

www.waxmann.com
info@waxmann.com

Buchumschlag: Anne Breitenbach, Münster
Umschlagabbildung: Alexandra Goßner, Berlin
Satz: Stoddart Satz- und Layoutservice, Münster

© Waxmann Verlag GmbH

Inhalt

Vorwort **11**

1 PISA 2015 – die aktuelle Studie **13**
Christine Sälzer & Kristina Reiss

1.1 Ziele und Anlage des *Programme for International Student Assessment* 15
1.1.1 Funktionen eines internationalen Bildungsvergleichs 16
1.1.2 Das Grundbildungskonzept in PISA: *Literacy* 17
1.1.3 Grundmerkmale von PISA 18
1.2 Besonderheiten von PISA 2015 21
1.2.1 Teilnehmende Staaten 21
1.2.2 Testkonzeption und Testdesign 22
1.2.3 Computerbasiertes Testen 26
1.3 Durchführung von PISA 2015 in Deutschland 28
1.3.1 Population und Stichprobe: Wer nimmt an PISA teil? 29
1.3.2 Teilnahmeverpflichtung und Ausschlussgründe 34
1.3.3 Ablauf eines PISA-Testtags 35
1.3.4 Sicherung der internationalen Vergleichbarkeit 35
1.4 Interpretation der Befunde 39
1.5 Nationale und internationale Organisation der PISA-Studie 40
Literatur 42

2 Naturwissenschaftliche Kompetenz in PISA 2015 – Ergebnisse des internationalen Vergleichs mit einem modifizierten Testansatz **45**
Anja Schiepe-Tiska, Silke Rönnebeck, Katrin Schöps, Knut Neumann, Stefanie Schmidtner, Ilka Parchmann & Manfred Prenzel

2.1 Die Rahmenkonzeption zur Erhebung der naturwissenschaftlichen Kompetenz in PISA 2015 48
2.1.1 Naturwissenschaftliche Kontexte 49
2.1.2 Naturwissenschaftliche Teilkompetenzen 51
2.1.3 Naturwissenschaftliche Wissensbereiche 52
2.1.4 Motivationale Orientierungen und Einstellungen zu den Naturwissenschaften 55
2.2 Der Naturwissenschaftstest in PISA 2015 56
2.2.1 Entwicklung der Aufgaben und ihre Formate 56
2.2.2 Aufgabenbeispiele 59
2.2.3 Auswertung und Kompetenzstufen 68
2.2.4 Bemerkungen zur Vergleichbarkeit der Ergebnisse von PISA 2015 und PISA 2006 70
2.3 Ergebnisse des internationalen Vergleichs 70

© Waxmann Verlag GmbH

2.3.1 Die Gesamtskala naturwissenschaftliche Kompetenz im internationalen Vergleich ... 71
2.3.2 Naturwissenschaftliche Teilkompetenzen, Wissensbereiche und Wissenssysteme im internationalen Vergleich ... 79
2.3.3 Schulartspezifische Analysen der naturwissenschaftlichen Kompetenz in Deutschland ... 85
2.3.4 Die Veränderung der naturwissenschaftlichen Kompetenz zwischen PISA 2006 und PISA 2015 ... 89
2.4 Zusammenfassung und Diskussion ... 92
Literatur ... 95

3 Motivationale Orientierungen, Selbstbilder und Berufserwartungen in den Naturwissenschaften in PISA 2015 ... 99
Anja Schiepe-Tiska, Inga Simm & Stefanie Schmidtner

3.1 Theoretischer Hintergrund ... 101
3.1.1 Freude und Interesse an den Naturwissenschaften ... 102
3.1.2 Instrumentelle Motivation ... 103
3.1.3 Selbstwirksamkeitserwartungen ... 104
3.2 Naturwissenschaftsbezogene Berufserwartungen ... 105
3.3 Die Erfassung naturwissenschaftsbezogener motivationaler Orientierungen, Selbstwirksamkeitserwartungen und Berufserwartungen in PISA 2015 ... 106
3.3.1 Methodische Einschränkungen ... 106
3.3.2 Eingesetzte Skalen ... 107
3.4 Ergebnisse ... 109
3.4.1 Naturwissenschaftsbezogene motivationale Orientierungen und Selbstwirksamkeitserwartungen im internationalen Vergleich ... 110
3.4.2 Schulartspezifische Unterschiede in den naturwissenschaftsbezogenen motivationalen Orientierungen und der Selbstwirksamkeitserwartung in Deutschland ... 115
3.4.3 Veränderungen der Schülermerkmale zwischen PISA 2006 und PISA 2015 ... 116
3.4.4 Naturwissenschaftsbezogene Berufserwartungen im internationalen Vergleich ... 118
3.4.5 Naturwissenschaftsbezogene Berufserwartungen in Deutschland differenziert nach Schularten ... 122
3.4.6 Veränderungen der naturwissenschaftsbezogenen Berufserwartungen zwischen PISA 2006 und PISA 2015 ... 123
3.4.7 Der Zusammenhang zwischen naturwissenschaftsbezogenen Berufserwartungen und Schülermerkmalen ... 124
3.5 Zusammenfassung und Diskussion ... 126
Literatur ... 129

© Waxmann Verlag GmbH

4 Naturwissenschaftlicher Unterricht in Deutschland in PISA 2015 im internationalen Vergleich 133
Anja Schiepe-Tiska, Stefanie Schmidtner, Katharina Müller, Jörg-Henrik Heine, Knut Neumann & Oliver Lüdtke

4.1 Empirische Befunde zu lernwirksamen Merkmalen des naturwissenschaftlichen Unterrichts 135
4.2 Die Erfassung der Unterrichtsqualität in PISA 2015 138
4.2.1 Methodische Herausforderungen 138
4.2.2 Eingesetzte Skalen 140
4.3 Ergebnisse 142
4.3.1 Naturwissenschaftlicher Unterricht im internationalen Vergleich 142
4.3.2 Schulartspezifische Unterschiede in der Wahrnehmung des naturwissenschaftlichen Unterrichts in Deutschland 151
4.3.3 Die Veränderung der Wahrnehmung des forschend-entdeckenden Unterrichts zwischen PISA 2006 und PISA 2015 153
4.3.4 Muster des naturwissenschaftlichen Unterrichts in Deutschland 156
4.3.5 Zusammenhänge zwischen Unterrichtsmustern des forschend-entdeckenden Lernens und domänenübergreifenden Merkmalen der Unterrichtsqualität 162
4.3.6 Zusammenhänge zwischen Mustern des forschend-endeckenden Unterrichts und mehrdimensionalen Bildungszielen am Gymnasium 165
4.4 Zusammenfassung und Diskussion 166
Literatur 170

5 Schulische Rahmenbedingungen der Kompetenzentwicklung 177
Christine Sälzer, Manfred Prenzel, Anja Schiepe-Tiska & Marcus Hammann

5.1 Institutionelle und organisatorische Rahmenbedingungen 183
5.1.1 Leistungsvarianz und Gliederung des Sekundarschulwesens 183
5.1.2 Alter und Klassenstufe der PISA-Kohorte 189
5.1.3 Lernzeit für naturwissenschaftlichen Unterricht 197
5.2 Die Einzelschule: Gestaltung von Lernumwelten in Deutschland 201
5.2.1 Klassengröße 202
5.2.2 Schulentwicklung und Qualitätssicherung 203
5.2.3 Schulische Angebote, Ressourcen und Lerngelegenheiten für die Naturwissenschaften 207
5.2.4 Schülerverhalten und Schulklima 209
5.2.5 Zusammenfassung und Diskussion 213
Literatur 215

6 Mathematische Kompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven 219
Sabine Hammer, Kristina Reiss, Matthias C. Lehner, Jörg-Henrik Heine, Christine Sälzer & Aiso Heinze

6.1 Mathematische Kompetenz in PISA 2015 221
6.2 Mathematische Kompetenz im internationalen Vergleich 229
6.3 Vertiefende Analysen der mathematischen Kompetenz in Deutschland 236

© Waxmann Verlag GmbH

6.3.1 Unterschiede zwischen Schularten ... 236
6.3.2 Unterschiede zwischen Mädchen und Jungen ... 240
6.3.3 Veränderung der mathematischen Kompetenz seit PISA 2003 ... 240
6.4 Zusammenfassung und Diskussion ... 244
Literatur ... 246

7 Lesekompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven ... 249
Mirjam Weis, Fabian Zehner, Christine Sälzer, Anselm Strohmaier, Cordula Artelt & Maximilian Pfost

7.1 Lesekompetenz in PISA 2015 ... 252
7.2 Lesekompetenz im internationalen Vergleich ... 263
7.3 Vertiefende Analysen zur Lesekompetenz in Deutschland ... 270
7.3.1 Unterschiede zwischen Schularten ... 270
7.3.2 Unterschiede zwischen Mädchen und Jungen in der Lesekompetenz ... 274
7.3.3 Veränderung der Lesekompetenz seit PISA 2000 ... 276
7.4 Zusammenfassung und Ausblick ... 279
Literatur ... 281

8 Soziale Herkunft und Kompetenzerwerb ... 285
Katharina Müller & Timo Ehmke

8.1 Erfassung der sozialen Herkunft in PISA ... 288
8.2 Der internationale Vergleich: Zusammenhang zwischen sozialer Herkunft und naturwissenschaftlicher Kompetenz ... 290
8.2.1 Kopplung zwischen naturwissenschaftlicher Kompetenz und sozioökonomischem Status (HISEI) ... 291
8.2.2 Kopplung zwischen naturwissenschaftlicher Kompetenz und ökonomischem, kulturellem und sozialem Status (ESCS) ... 295
8.2.3 Veränderungen in den sozialen Disparitäten der naturwissenschaftlichen Kompetenz ... 298
8.3 Vertiefende nationale Analysen: die soziale Herkunft der Jugendlichen in Deutschland ... 300
8.3.1 Vorhersage der naturwissenschaftlichen Kompetenz durch Merkmale der sozialen Lage und des Bildungsweges ... 301
8.3.2 Beschreibung der sozialen Lage der Jugendlichen aus unterschiedlichen EGP-Klassen ... 303
8.3.3 Veränderung der sozialen Disparitäten der Lesekompetenz und der Bildungsbeteiligung in den EGP-Klassen zwischen PISA 2000 und PISA 2015 ... 306
8.3.4 Beschreibung der sozialen Lage der leistungsstarken und leistungsschwachen Jugendlichen ... 308
8.4 Zusammenfassung und Diskussion ... 310
Literatur ... 313

© Waxmann Verlag GmbH

9 Naturwissenschaftliche Kompetenz von Schülerinnen und Schülern mit Zuwanderungshintergrund 317
Dominique Rauch, Julia Mang, Hendrik Härtig & Nicole Haag

9.1 Einleitung 317
9.1.1 Methodische Vorbemerkungen 322
9.2 Jugendliche mit Zuwanderungshintergrund im europäischen Vergleich 324
9.3 Jugendliche mit Zuwanderungshintergrund in Deutschland 332
9.4 Zusammenfassung und Diskussion 339
Literatur 344

10 Die Lernumgebung in der Familie und die elterliche Unterstützung 349
Nina Jude, Silke Hertel, Susanne Kuger & Christine Sälzer

10.1 Theoretische Rahmung und Indikatoren in PISA 2015 350
10.2 Erfassung der häuslichen Lernumgebung sowie der elterlichen Unterstützung in PISA 2015 353
10.3 Die häusliche Lernumgebung von Fünfzehnjährigen aus Sicht der Eltern und das elterliche Engagement an der Schule ihres Kindes 355
10.4 Elterliche Einstellungen, Überzeugungen, Motive und Interessen 362
10.5 Zusammenhänge der elterlichen Unterstützung mit den Kompetenzen der Schülerinnen und Schüler in den Naturwissenschaften 366
10.6 Diskussion und Ausblick 369
Literatur 371

11 Fünfzehn Jahre PISA: Bilanz und Ausblick 375
Kristina Reiss & Christine Sälzer

11.1 Zwischen Kontinuität und Innovation 376
11.2 Disparitäten: Thema des PISA-Schocks und dauerhafte Herausforderung 377
11.3 Spitzenförderung 380
11.4 Fazit 381

12 Kompetenzmessung in PISA 2015 383
Jörg-Henrik Heine, Julia Mang, Lars Borchert, Jens Gomolka, Ulf Kröhne, Frank Goldhammer & Christine Sälzer

12.1 Einleitung und Überblick 383
12.2 Population und Stichprobe der PISA-Teilnehmer in Deutschland 385
12.2.1 Populationsdefinitionen, Stichprobenplan und Ziehung der Stichproben 385
12.2.2 Realisierte Schul-, Schüler- und Lehrerstichproben 395
12.2.3 Gewichtung als Adjustierung unterschiedlicher Ziehungswahrscheinlichkeiten 396
12.3 Tests zur Erfassung der Kompetenzen 398
12.3.1 Das PISA-Testmaterial im Multi-Matrix-Design 398
12.3.2 Psychometrische Modelle zur Testauswertung 402
12.3.3 Das *1- und 2-Parameter-Logistische-Modell*: eine vergleichende Diskussion von Merkmalen 404
12.4 Umstellung auf computerbasiertes Assessment (CBA) 407

© Waxmann Verlag GmbH

12.5 Reliabilität, Stichprobenfehler und Beurteilerübereinstimmung ... 411
12.5.1 Kodierung offener Antworten und Beurteilerübereinstimmung ... 411
12.5.2 Designeffekte, Stichprobenfehler und Standardfehler ... 413
12.6 Itemkalibrierung, Populationsmodell, Plausible Values und Kompetenzstufen ... 415
12.6.1 Bestimmung der Aufgabenschwierigkeit – Itemkalibrierung ... 416
12.6.2 Hintergrundmodell als latente Regression für Populationsmodelle und Ziehung von Plausible Values als Messwerte für Kompetenzen ... 417
12.6.3 Definition von Kompetenzstufen ... 419
12.7 Betrachtung von Trends zwischen PISA 2000 und PISA 2015 ... 420
12.8 Zusammenfassung und Ausblick ... 423
Literatur ... 424

13 Die Erhebung von Lern- und Lehrkontexten, Hintergrundmerkmalen und nichtleistungsbezogenen Lernergebnissen in PISA 2015 ... 431
Nina Jude, Eckhard Klieme, Susanne Kuger & Fabian Zehner

13.1 Konzeptionelle Grundlagen der Fragebögen für PISA ... 433
13.2 Entwicklung der Fragebögen für PISA 2015 ... 437
13.3 Skalierung der Fragebögen: Vorbereitung der Datenauswertung ... 440
13.4 Innovationen 2015: Computerbasierte Erhebung und ein internationaler Lehrerfragebogen ... 441
13.5 Ausblick ... 443
Literatur ... 444

14 Glossar zentraler Begriffe zu den PISA-Ergebnissen ... 447
Christine Sälzer, Kristina Reiss & Jörg-Henrik Heine

Anhang ... 459

Abbildungsverzeichnis ... 498

Tabellenverzeichnis ... 501

Die Autorinnen und Autoren dieses Berichtsbandes ... 504

© Waxmann Verlag GmbH

Vorwort

PISA ist nach mittlerweile sechs abgeschlossenen Erhebungsrunden ein fester Bestandteil des Bildungsmonitorings in den OECD-Staaten allgemein und speziell in Deutschland. Die Studie gibt insbesondere Auskunft darüber, welche Arbeit in Schule und Unterricht fachlich und pädagogisch geleistet wird und worin ihre Ergebnisse bestehen. Sie erlaubt ein internationales Benchmarking im Vergleich mit den OECD-Staaten und den Partnerstaaten, wobei die wiederholte Messung der Kompetenzen im Drei-Jahres-Rhythmus eine Analyse der Veränderungen der Leistungsfähigkeit von Bildungssystemen ermöglicht. PISA ist damit für die Praxis genauso wie für die Bildungspolitik von hoher Bedeutung. Auch Deutschland hat von diesem internationalen Vergleich in den zurückliegenden Jahren erheblich profitiert. Insbesondere PISA 2000 war Anlass für einen kritischen, gemeinsamen Blick von Bund und Ländern auf das deutsche Bildungssystem. In der Folge wurden erhebliche Anstrengungen unternommen, das Kompetenzniveau der Schülerinnen und Schüler in Deutschland zu verbessern. Diese Anstrengungen waren erfolgreich: Lagen die Ergebnisse der Tests im Jahr 2000 unter dem Durchschnitt der OECD-Staaten, so sind sie jetzt durchweg signifikant über diesem Durchschnitt und zeigen die sehr erfreuliche, positive Entwicklung der Kompetenzen der Fünfzehnjährigen in Deutschland. Auch beim Abbau sozialer Disparitäten sind Fortschritte gemacht worden, was sich gerade im internationalen Vergleich zeigt.

PISA 2015 markiert den Abschluss des zweiten PISA-Zyklus, in dem nun jede der drei Domänen Lesen, Mathematik und Naturwissenschaften zweimal den Schwerpunkt bildete. Trotz der scheinbar eingekehrten Routine ist die Studie aber auch der Beginn einer neuen Phase in der Testung von Fünfzehnjährigen. Verschiedene Neuerungen kamen dieses Mal zum Tragen, die nicht ganz ohne Einfluss auf die Ergebnisse gewesen sein dürften. So wurden beispielsweise Testaufgaben und Fragebögen nicht mehr in der traditionellen Version mit Papier und Bleistift vorgegeben, sondern die Bearbeitung fand vollständig am Computer statt. Dadurch konnten die Schülerinnen und Schüler flexibler durch die Testsituation geführt werden, und es wurden in den Naturwissenschaften neue Aufgabenformate möglich, die stärker an experimentelles Arbeiten und damit an einen zeitgemäßen Unterricht angelehnt sind. Auch bei der statistischen Modellierung der Kompetenzen und Fragebogenskalen gab es eine wesentliche Änderung, die bisherige klassische Rasch-Skalierung wurde durch eine differenziertere Skalierung ersetzt, die Spezifika der Aufgaben noch besser berücksichtigt. Die Änderungen im Design und in der Skalierung ermöglichen ein genaueres Bild der Bearbeitung durch die Schülerinnen und Schüler, verbessern die Messungen der Kompetenzen und erlauben den Einsatz neuer, interaktiver Testaufgaben. Sie erschweren allerdings die Interpretation der Veränderungen in den Werten zwischen früheren PISA-Erhebungen und dem Jahr 2015. Folgt man der Rationale „*If you want to measure change, don't change the measure*“, so soll-

© Waxmann Verlag GmbH

ten Analysen dazu angestellt werden, welche Effekte die Änderungen möglicherweise auf die Befunde in PISA 2015 hatten. In der Tat weisen unsere diesbezüglichen nationalen Analysen darauf hin, dass der neue Testmodus „Computer“ auf die Kompetenzwerte der Schülerinnen und Schüler in Deutschland einen Einfluss gehabt haben könnte und die Aufgaben letztendlich etwas schwieriger waren als auf Papier. PISA 2015 ist insbesondere keine einfache Fortsetzung der bisherigen Studien, sondern durch die innovativen Aspekte als ein Neubeginn zu sehen, der Veränderungen in der Lern- und Lebenswelt – wie etwa durch die Digitalisierung – gerecht wird und die Aussagekraft der PISA-Studien auf lange Sicht verbessert.

Vor dem Hintergrund dieser Herausforderungen werden die Befunde im hier vorliegenden nationalen Berichtsband interpretiert und in den Kontext des deutschen Bildungssystems eingeordnet. Wir argumentieren insbesondere im Hinblick auf die Trends in den erreichten Leistungen vorsichtiger als in den letzten Berichtsbänden. Sicher scheint allerdings zu sein, dass sich Deutschland in Bezug auf die Kompetenzen weiterhin auf guten oberen Plätzen im Vergleich der OECD-Staaten befindet und auch PISA 2015 ein positives Bild für das deutsche Bildungssystem zeichnet.

München, Kiel und Frankfurt im November 2016
Kristina Reiss, Christine Sälzer, Anja Schiepe-Tiska, Eckhard Klieme und Olaf Köller

© Waxmann Verlag GmbH

1 PISA 2015 – die aktuelle Studie

Christine Sälzer & Kristina Reiss

Nach PISA 2000, 2003 und 2006 ist mit der neuen Studie der zweite vollständige PISA-Zyklus (2009, 2012 und 2015) abgeschlossen und jeder der drei Kompetenzbereiche Lesen, Mathematik und Naturwissenschaften war nunmehr bereits zweimal inhaltlicher Schwerpunkt der PISA-Studie. Vorbereitung und Durchführung orientierten sich auch 2015 weitgehend an den bewährten Abläufen. Die wichtigste Änderung für die Testteilnehmerinnen und Testteilnehmer war die Umstellung auf computerbasiertes Testen. Zu ihren Erfahrungen im Umgang mit Computern befragt, geben sich Jugendliche in Deutschland hier zwar teilweise durchaus selbstbewusst, im Unterricht nutzen sie den Computer im internationalen Vergleich aber eher wenig. In dieser Erhebungsrunde wurde außerdem ein anderes Modell zur Skalierung der Daten verwendet, das neben der Aufgabenschwierigkeit und der Fähigkeit der Jugendlichen auch explizit die Trennschärfe von Testaufgaben berücksichtigt und damit das Potenzial, tatsächlich zwischen besserer und schlechterer Leistung zu unterscheiden.

Im Rahmen der PISA-Studie – initiiert durch die *Organisation for Economic Co-operation and Development* (OECD) – bearbeiten weltweit jeweils mehrere Hunderttausend fünfzehnjährige Schülerinnen und Schüler alle drei Jahre ein umfassendes Erhebungsprogramm aus Tests und Fragebögen. Gegenüber der ersten Erhebung im Jahr 2000, als 32 Staaten an der Studie teilnahmen, wuchs die Anzahl der beteiligten Staaten auch in der sechsten Erhebungsrunde weiter an: 72 Staaten, darunter 35 OECD-Staaten und 37 sogenannte OECD-Partnerstaaten, setzten PISA 2015 um. Die weiterhin steigende Beteiligung an PISA bekräftigt das nach wie vor hohe internationale Interesse an einem regelmäßigen Bildungsmonitoring, das aussagekräftig und zuverlässig Erträge des Bildungssystems dokumentiert, Herausforderungen aufzeigt und im Sinne eines Benchmarkings Hinweise auf Möglichkeiten für Verbesserungen gibt. Nach mittlerweile sechs abgeschlossenen Erhebungsrunden ist auch eine Betrachtung von Entwicklungen über die Zeit möglich, die sich in den einzelnen Teilnehmerstaaten teils deutlich voneinander unterscheiden. Sicherlich muss man diese vorsichtig interpretieren, denn auch Erhebungsmethoden und Aufgaben haben sich im Laufe der Zeit weiterentwickelt.

© Waxmann Verlag GmbH

Als internationale Bildungsvergleichsstudie ist PISA in Deutschland ein wichtiger Bestandteil der Gesamtstrategie zum Bildungsmonitoring, welche 2006 von der Ständigen Konferenz der Kultusminister der Länder (KMK) verabschiedet und 2016 aktualisiert wurde (KMK, 2006; 2016). PISA erfüllt dabei in erster Linie den Zweck, die Kompetenzen der Schülerinnen und Schüler am Ende der Sekundarstufe I zu beschreiben und international zu verankern. Von besonderem Interesse sind Vergleichsperspektiven, die Entwicklungen über die Zeit im Sinne von Trends erfassen. Solche Trendinformationen sind für die Steuerung von Bildungssystemen hilfreich, denn mit ihnen können Wirkungen von Maßnahmen abgeschätzt und problematische Entwicklungen frühzeitig erkannt werden. Erst der internationale Vergleich ermöglicht die Beurteilung, ob beispielsweise eine Klassenwiederholungsrate von etwa 20 Prozent bei den Fünfzehnjährigen in Deutschland (vgl. Sälzer, Prenzel & Klieme, 2013) ein auffallender Wert ist oder nicht.

Vor der Beteiligung an der *Third International Mathematics and Science Study* (TIMSS) im Jahr 1995 hatte Deutschland mehrere Jahrzehnte lang an so gut wie keiner international vergleichenden Schulleistungsstudie teilgenommen, weshalb die mittelmäßigen Ergebnisse aus TIMSS 1995 und wenige Jahre später aus PISA 2000 die heute als „PISA-Schock“ (Roeder, 2003) bekannte Reaktion auslösten. Die gravierenden Probleme, die durch die Befunde aus TIMSS 1995 und PISA 2000 sichtbar wurden, waren durch den langen Verzicht auf einen internationalen Vergleich (Prenzel, 2005) und das empirisch nicht hinterfragte Vertrauen in die Leistungsfähigkeit des deutschen Schulsystems (Gräsel, 2011) mit erheblicher Verzögerung entdeckt worden. Mittlerweile hat die Beteiligung Deutschlands an internationalen Bildungsvergleichen zwei Jahrzehnte Tradition, und die Trendinformationen aus diesen Studien liefern Anhaltspunkte für Entwicklungen und mögliche Wirkungen bereits ergriffener Maßnahmen. In Deutschland stehen hier insbesondere die seit 2003 sukzessive eingeführten Bildungsstandards für unterschiedliche Schulstufen und -abschlüsse im Mittelpunkt. Die Ergebnisse aus PISA 2015 für Deutschland sollen auch aus dieser Perspektive interpretiert und diskutiert werden.

Nachdem sich in den bisherigen fünf PISA-Erhebungsrunden Verbesserungen der durchschnittlichen Kompetenzen bei den fünfzehnjährigen Schülerinnen und Schülern in Deutschland gezeigt haben, wurden die Ergebnisse von PISA 2015 gespannt, aber durchaus unaufgeregt erwartet. Die Öffentlichkeit wird von der OECD wiederum in mehreren umfangreichen Berichtsbänden über die internationalen Ergebnisse informiert, die mit einer überschaubaren Anzahl einfacher Rankings ergänzt werden. Diese Berichtsbände dokumentieren den staatenübergreifenden Bildungsvergleich aus einer weltweiten Perspektive. Konkrete Fragen aus nationaler Sicht, etwa mit Blick auf bereits ergriffene Maßnahmen oder Besonderheiten des Schulsystems, können in den OECD-Berichten naturgemäß nicht aufgearbeitet werden. Es ist das Anliegen dieses nationalen Berichtsbandes, der im Zentrum für Internationale Bildungsvergleichsstudien (ZIB) verfasst wurde, einen kohärenten Eindruck der Ergebnisse aus PISA 2015 für Deutschland zu geben. Der Bericht präsentiert die Ergebnisse mit dem Blick auf Deutschland und soll durch die Erläuterung von Testkonzeptionen, Methoden und

© Waxmann Verlag GmbH

Ergebnissen helfen, die aktuellen Befunde und Trends über die Zeit zu interpretieren. Wie in früheren PISA-Erhebungsrunden wurden dazu umfassende, für Deutschland spezifische Datenanalysen durchgeführt und zusätzliche Daten aus nationalen Ergänzungen einbezogen.

Die sechste Erhebungsrunde brachte eine wesentliche Neuerung für alle an PISA Beteiligten mit sich, nämlich die Umstellung von Papier und Bleistift auf computerbasiertes Testen. Obwohl sich die grundlegenden Strukturen der Studie bereits etabliert hatten und zahlreiche Prozeduren zur Routine geworden waren, zieht diese Umstellung doch verschiedene Konsequenzen nach sich. Diese werden vorwiegend in Kapitel 12 dieses Berichtsbandes näher erläutert. In diesem einführenden Kapitel werden zunächst die Zielsetzungen und das Grunddesign der Studie vorgestellt. Der zweite Abschnitt wendet sich den Besonderheiten der aktuellen Studie und damit konkret PISA 2015 zu, der dritte Abschnitt beschreibt Aspekte der Durchführung von PISA 2015 in Deutschland. Darauf folgen Abschnitte, die beim Lesen der weiteren Kapitel in diesem Berichtsband hilfreich sein sollen. Abschließend werden die Organisationsstruktur sowie die an der Durchführung von PISA 2015 in Deutschland beteiligten Personen vorgestellt.

1.1 Ziele und Anlage des *Programme for International Student Assessment*

International vergleichende Schulleistungsstudien wie PISA sind ein Instrument des Bildungsmonitorings, dessen Ergebnisse überwiegend beschreibendes Wissen bereitstellen. Die OECD hat mit PISA ein eigenes Indikatorensystem für Bildungsergebnisse geschaffen, dessen Daten als empirischer Bezugspunkt für die Steuerung von Bildungssystemen relevant sind und entsprechend genutzt werden können. Weil diese Daten Auskunft über relative Stärken und Schwächen der teilnehmenden Bildungssysteme, über Probleme, Herausforderungen und Handlungsoptionen geben, hat sich die Teilnahme an internationalen Bildungsvergleichen mittlerweile zu Recht etabliert. In regelmäßigen Berichten verknüpft die OECD Strukturdaten von Bildungssystemen mit Ergebnissen aus PISA, sodass in diesen Berichten Zusammenhänge analysiert und Schlussfolgerungen über die Effektivität von Bildungssystemen gezogen werden können. Zugleich ist PISA ein Angebot an die OECD-Staaten sowie an interessierte OECD-Partnerstaaten, dieses Indikatorenprogramm für ein regelmäßiges Bildungsmonitoring zu nutzen. So lautet die zentrale Frage der PISA-Studie, wie gut es den teilnehmenden Staaten gelingt, ihre jungen Menschen in Schulen auf das Erwachsenenleben vorzubereiten. Dazu gehören explizit die Vorbereitung auf kontinuierliches Weiterlernen nach der Schule und über die gesamte Lebensspanne sowie eine aktive Teilhabe am gesellschaftlichen Leben. Mit diesem Anspruch verbinden sich Anforderungen an die Studie. So muss PISA selbstverständlich Aussagen über das Niveau und die Verteilung wichtiger Kompetenzen ermöglichen. Zugleich werden Indikatoren für Disparitäten benötigt, anhand derer systematische Zusammenhänge zwischen den Schülerleistungen und Merkmalen wie der sozialen Her-

© Waxmann Verlag GmbH

kunft, einem Zuwanderungshintergrund oder dem Geschlecht untersucht werden können. Dabei muss sich das in PISA verwendete Indikatorensystem mit der gesellschaftlichen Realität verändern, in der die untersuchten fünfzehnjährigen Schülerinnen und Schüler jeweils leben und gerade ihre Pflichtschulzeit beenden. In PISA 2000 sah diese Realität deutlich anders aus als in PISA 2015, denn allein die Verbreitung des Internets und die Verfügbarkeit digitaler Endgeräte haben in den 15 Jahren für spürbare Veränderungen gesorgt.

1.1.1 Funktionen eines internationalen Bildungsvergleichs

International vergleichende Schulleistungsstudien haben zwei Hauptfunktionen, und zwar Monitoring und Benchmarking (Seidel & Prenzel, 2008). Beides impliziert Vergleiche, zum einen mit bestimmten Standards (Monitoring) und zum anderen mit der Struktur, den Prozessen und Ergebnissen anderer Bildungssysteme (Benchmarking). Zentrale Erkenntnisse aus diesen Studien betreffen das Kompetenzniveau der teilnehmenden Schülerinnen und Schüler, dessen Verteilung sowie diverse Rahmenbedingungen, die Einfluss auf die Kompetenzentwicklung nehmen können. Der internationale Vergleich der Leistungsfähigkeit von Schülerinnen und Schülern zu bestimmten Zeitpunkten in der Schullaufbahn liefert Vergleichsperspektiven, die Trends im Sinne von Entwicklungen über die Zeit erfassen und abbilden können (Rutkowski & Prusinski, 2011). So können neben der internationalen Verankerung von Schülerleistungen Vergleiche dieser Leistungen über mehrere Erhebungsrunden angestellt, mögliche Wirkungen von Maßnahmen abgeschätzt und problematische Entwicklungen frühzeitig erkannt werden (Drechsel & Prenzel, 2008; Drechsel, Prenzel & Seidel, 2014).

Für die OECD als Auftraggeberin der PISA-Studien steht eine ökonomisch orientierte Frage im Mittelpunkt: Wie gut sind Schülerinnen und Schüler kurz vor dem Ende ihrer Pflichtschulzeit auf das vorbereitet, was sie nach der Schulpflicht erwartet? Je besser junge Menschen beispielsweise auf das erfolgreiche Durchlaufen einer Berufsausbildung oder einer weiterführenden Schulbildung vorbereitet sind, desto besser sind bisherige Investitionen in das Bildungswesen angelegt. Dementsprechend wurde als Zielpopulation eine Altersgruppe gewählt, die international normalerweise noch das Pflichtschulsystem besucht, nämlich fünfzehnjährige Schülerinnen und Schüler. Die Untersuchung einer Alterskohorte anstelle von Klassen- oder Jahrgangsstufen hat den Vorteil, dass Wirkungen von Bildungsprozessen und -systemen international auf dasselbe Lebensalter und damit auf eine bestimmte Zeitspanne bezogen werden. Selbstverständlich werden die internationalen Vergleiche an repräsentativen Stichproben durchgeführt und entsprechende Mindestquoten an Testbeteiligung, Ausschöpfung der Stichprobe und weitere Rahmenbedingungen definiert.

Inhaltlich ist es bei internationalen Vergleichsstudien nicht unbedingt kanonisch, welche Kompetenzbereiche (in PISA: Domänen) jeweils untersucht werden sollen. Für die Auswahl der Untersuchungsbereiche sind mehrere Kriterien zu beachten. Zum einen

© Waxmann Verlag GmbH

müssen die Domänen in Bezug auf ihre Lerngelegenheiten international vergleichbar sein, was bei Domänen wie Geschichte oder Sozialkunde in der Regel schwieriger ist als etwa bei Mathematik oder Naturwissenschaften. Zum anderen müssen die Bildungsergebnisse mit Testverfahren aussagekräftig erfasst, zuverlässig bewertet und sinnvoll interpretiert werden können, was etwa in Kunst oder Musik problematisch ist. Darüber hinaus sollen internationale Schulleistungsvergleiche verallgemeinerbare Einschätzungen über die Qualität von Schulen zulassen und Kompetenzen erfassen, die für kontinuierliches Lernen über die Lebensspanne relevant sind. Auf der Suche nach möglichst großen internationalen Gemeinsamkeiten in den Curricula zeichnete sich bereits in Vorläuferstudien zu PISA (etwa der *Third International Science and Mathematics Study*, TIMSS, initiiert von der *International Association for the Evaluation of Educational Achievement* – IEA) ab, dass diese in den Bereichen Mathematik und Naturwissenschaften weitgehend vorhanden sind und aussagekräftige Testungen auch bei umfangreichen Stichproben realisiert werden können. Das Spektrum der getesteten Domänen wurde erweitert durch die Lesekompetenz, die in einer ganzen Reihe von Schulfächern eine grundlegende Rolle spielt und für weiterführendes Lernen über die Lebensspanne unerlässlich ist.

1.1.2 Das Grundbildungskonzept in PISA: *Literacy*

In PISA basieren die Leistungstests auf dem angelsächsischen *Literacy*-Konzept, das im Deutschen am besten unter dem Begriff der Grundbildung zu fassen ist. Mit Blick auf die Grundbildung von Fünfzehnjährigen stellt PISA die Frage, inwieweit Schülerinnen und Schüler gegen Ende ihrer Pflichtschulzeit Kenntnisse und Fähigkeiten erworben haben, die für eine erfolgreiche Teilhabe an modernen Gesellschaften als unerlässlich angesehen werden (vgl. etwa OECD, 2016a).

Die OECD verwendet in PISA den Begriff der *Literacy*, wie er im englischsprachigen Raum als Bezeichnung einer funktionalen Grundbildung gebraucht wird. Im Kontext von PISA umfasst „funktional“ im Wesentlichen zwei Aspekte, nämlich die *Anwendbarkeit* für die jetzige und die spätere, nachschulische Teilhabe an einer Kultur sowie die *Anschlussfähigkeit* im Sinne kontinuierlichen Weiterlernens über die Lebensspanne. Die Konzentration auf diese beiden Aspekte ist entscheidend für das Ziel, Bildungsergebnisse gegen Ende der Pflichtschulzeit zu messen. Einerseits sollen Wissen und Können in den untersuchten Domänen erfasst werden, die als Grundlagen für eine lernende Weiterentwicklung in diesen Bereichen vorausgesetzt werden. Andererseits geht es um die Frage, inwieweit bis zum Testzeitpunkt schulische und außerschulische Lerngelegenheiten wahrgenommen worden sind, um in bestimmten Bereichen teilhabe- und handlungsfähig zu werden (im Englischen „literate“). Dass die Lesekompetenz für beide Aspekte besonders wichtig ist, ist unmittelbar einsichtig, denn in beinahe allen Lebensbereichen wird Wissen hauptsächlich in Form von Texten gespeichert, weitergegeben und angeeignet. Doch auch die Mathematik und die Naturwissenschaften weisen in

© Waxmann Verlag GmbH

vielfältigen Fach- und Bildungskontexten eine Anwendungsbreite auf, die deutlich über die jeweiligen Schulfächer hinausreicht. Insofern erfasst PISA relevante Aspekte einer Grundbildung, die durchaus im Sinne einer Allgemeinbildung zu verstehen sind (vgl. Tenorth, 2004; 2005). Zugleich ist jedoch zu berücksichtigen, dass es sich bei dieser Auswahl an Domänen stets nur um einen Ausschnitt sowohl aus dem Spektrum allgemeiner Bildung als auch allgemeinbildender Schulfächer handelt und die in PISA eingesetzten Tests insbesondere kognitive Aspekte fokussieren. Bisher wurden in PISA darüber hinausgehende Merkmale der Schülerinnen und Schüler sowie ihrer häuslichen und ihrer Lernumgebung relativ begrenzt erfasst. In PISA 2015 wird nicht kognitiven Merkmalen allerdings ein breiterer Raum als zuvor zugestanden und damit ihre Bedeutung für erfolgreiche Lernprozesse gewürdigt (vgl. Kapitel 3).

1.1.3 Grundmerkmale von PISA

PISA konzentriert sich mit Lesen, Mathematik und Naturwissenschaften auf ausgewählte Inhalts- und Kompetenzbereiche, die weltweit in Schulen vermittelt werden und die als grundlegend für die persönliche Bildungsbiografie, das Berufsleben und die Partizipation an Gesellschaft und Kultur gelten. Das Wissen und Können fünfzehnjähriger Jugendlicher in diesen Bereichen wird mit geeigneten Testverfahren gemessen. Darüber hinaus werden Informationen über die Herkunft und das Umfeld der Fünfzehnjährigen sowie ihre Einstellungen, Lern- und Freizeitaktivitäten und Erfahrungen in Schule und Unterricht gesammelt. Alle drei Jahre erfolgt eine PISA-Erhebungsrunde an repräsentativen Stichproben dieser Altersgruppe, die einen Vergleich der Leistungen der unterschiedlichen Kohorten im Zeitverlauf ermöglicht. Der Bezugspunkt für Vergleiche der Staaten untereinander ist jeweils das mittlere Kompetenzniveau der OECD-Staaten.

Die zentralen Elemente der PISA-Studie sind in den theoretischen Rahmenkonzeptionen (*Assessment Frameworks)* sowie den *Technical Reports* zu jeder Erhebungsrunde detailliert beschrieben (vgl. etwa OECD, 2013; 2014; 2016a). Die drei untersuchten Domänen wurden dabei von Beginn an unterschiedlich gewichtet, sodass jeweils eine als Hauptdomäne differenziert erfasst wird (und etwa die Hälfte aller Testaufgaben ausmacht) und die beiden anderen Nebendomänen mit entsprechend weniger Testaufgaben sind. Beginnend mit PISA 2015 soll die Gewichtung der drei Domänen einander zunehmend angeglichen werden, wobei die Unterteilung in Haupt- und Nebendomäne erhalten bleibt. In PISA 2000 bis PISA 2012 bestand der Hauptunterschied zwischen der Haupt- und den Nebendomänen in der unterschiedlichen Anzahl der Testaufgaben. In PISA 2015 werden erstmals in den beiden Nebendomänen insgesamt mehr Testaufgaben eingesetzt, die relativ gesehen von weniger Schülerinnen und Schülern bearbeitet werden. In PISA 2015 stellen nach PISA 2006 zum zweiten Mal die Naturwissenschaften die Hauptdomäne dar. So kann alle neun Jahre die jeweilige Hauptdomäne umfassend und differenziert erfasst und dementsprechend detailliert im Trend betrachtet werden. Dieses Design bietet zudem die Möglichkeit, für die Hauptdomäne einer Erhebungsrunde die

© Waxmann Verlag GmbH

theoretische Rahmenkonzeption zu aktualisieren und den Aufgabenpool entsprechend zu erneuern. In PISA 2015 erfolgte beides für die Domäne Naturwissenschaften.

PISA ist eine Querschnittsuntersuchung. Gleichwohl ist die Abbildung und Beschreibung von Entwicklungen über die Zeit ein etablierter Aspekt der PISA-Studie. Pro untersuchter Kohorte gibt es nur einen Messzeitpunkt, es ist prinzipiell aber möglich, anhand von Trends über mehrere Erhebungsrunden hinweg eine Entwicklung nachzuzeichnen. Eine wichtige Voraussetzung sind Aufgaben (sogenannte Anker- oder *Link-Units*), die in jeder Runde verwendet und entsprechend nicht öffentlich gemacht werden. Diese Link-Units ermöglichen leistungsbezogene Vergleiche der Schülerkompetenzen über die Erhebungszeiträume. Veraltete Aufgaben, die der aktuellen Lebenswelt heutiger Fünfzehnjähriger nicht mehr gerecht werden, werden mit der Zeit ausgetauscht. Indem die PISA-Tests mit Ansätzen der *Item Response Theory* (IRT, vgl. z. B. Rost, 2004 und Kapitel 2 dieses Berichtsbands) skaliert und ausgewertet werden, können einige Vorteile genutzt werden: So muss nicht jede Schülerin und jeder Schüler alle vorhandenen Aufgaben bearbeiten – dies würde mehrere Stunden dauern und die Konzentrationsspanne bei weitem übersteigen. Vielmehr werden die Aufgaben auf mehrere unterschiedlich zusammengesetzte Testversionen verteilt, wobei eine Schülerin oder ein Schüler nur jeweils einen Teil aller PISA-Aufgaben bearbeitet. Ein praktischer Nebeneffekt ist, dass so ein Abschreiben verhindert wird, da sich die Testversionen voneinander unterscheiden. Durch die Rotation der Testaufgaben, d. h. die Verteilung von Aufgabenblöcken auf unterschiedliche Testformen bei den 2015 eingesetzten Aufgaben am PC kann eine große Menge an Aufgabenmaterial in der begrenzten Testzeit von 120 Minuten eingesetzt und zur Schätzung der Leistungsfähigkeit auf Aggregatebenen wie Bildungssystem oder Staat verwendet werden (vgl. Sälzer, 2016).

Die Konstruktion und Auswertung der PISA-Tests auf der Grundlage der *Item Response Theory* ermöglicht die Differenzierung und Beschreibung von Kompetenzstufen. Eine Kompetenzstufe entspricht einem bestimmten Bereich von Leistungswerten, der durch entsprechende Aufgabenanforderungen inhaltlich beschrieben werden kann. Jede Kompetenzstufe ist durch Anforderungen definiert, die von Schülerinnen und Schülern auf dieser Kompetenzstufe typischerweise bewältigt werden. Damit veranschaulichen die Kompetenzstufen, was die Punktwerte auf der PISA-Skala inhaltlich in Bezug auf die jeweils untersuchte Domäne bedeuten. Anhand dieser Veranschaulichung der Schülerkompetenzen kann abgeschätzt werden, ob bestimmte Bildungsziele erreicht und damit grundlegende Voraussetzungen für weiterführendes Lernen (etwa für den Besuch einer weiterführenden Schule oder einer beruflichen Ausbildung) entwickelt wurden. Auf diese Weise können umgekehrt aber auch Jugendliche am Ende der Pflichtschulzeit identifiziert werden, denen diese Entwicklung (noch) nicht gelungen ist und deren Nachholbedarf so auch auf inhaltlicher Ebene detailliert beschrieben werden kann.

Die Stärke von PISA als internationaler Bildungsvergleichsstudie zeigt sich dort, wo systemübergreifende Informationen gegeben werden können. Das gilt etwa für das *Assessment*, das auf drei Domänen konzentriert und jeweils um eine übergreifende Kompetenz ergänzt wird. Die Tests in PISA werden anhand differenziert entwickelter Kon-

© Waxmann Verlag GmbH

zeptionen konstruiert und erfassen sowohl lehrplan- als auch alltagsrelevantes Wissen sowie fachliche Voraussetzungen für weiterführendes Lernen. Gefordert wird dabei weniger die simple Reproduktion, sondern vielmehr das flexible Anwenden des Wissens und das Lösen domänenspezifischer Probleme. Inhaltlich können unterschiedliche Facetten durch Kompetenzstufen beschrieben werden, die eine Prognose über die Anschlussfähigkeit für weiterführendes Lernen erlauben. Anhand der Trends über die Zeit können neben Vergleichen mit anderen Staaten auch Verbesserungen oder Verschlechterungen fachlicher Kompetenz im eigenen Land beschrieben werden, was letztlich der bildungspolitischen Steuerung zugutekommt. Die aus den Kontextfragebögen gewonnenen Daten zu Merkmalen von Schülerinnen und Schülern sowie ihrem häuslichen und schulischen Umfeld liefern hauptsächlich Erkenntnisse zu möglichen Disparitäten bei Bildungschancen, etwa in Bezug auf Geschlecht, Zuwanderungsstatus oder soziale Herkunft.

Für eine angemessene Interpretation der Befunde ist es zentral, auch die Grenzen der Studie zu kennen. Dem Ziel der OECD folgend, mit PISA in erster Linie Wissen für die Steuerung von Bildungssystemen bereitzustellen, ist das Erhebungsdesign der PISA-Studie stark auf die Ebenen der Schule und der Bildungssysteme ausgerichtet. Befunde zur individuellen Schülerdiagnostik werden prinzipiell nicht aus der Studie gewonnen (vgl. auch Sälzer, 2016). So erfordert das internationale Stichprobendesign mit Blick auf die Aggregationsebene „Schule“ die Ziehung einer Zufallsstichprobe in der Größenordnung von $n = 30$ fünfzehnjährigen Schülerinnen und Schülern pro Schule. Üblicherweise sind Schulen der Sekundarstufe mehrzügig, sodass sich die Stichprobe der Schülerinnen und Schüler auf mehrere Klassen und wegen der altersbasierten Definition auch auf mehrere Klassenstufen verteilt; in Deutschland finden sich fünfzehnjährige Schülerinnen und Schüler in den Klassenstufen 7 bis 11. Demnach bietet das internationale Grunddesign keine Möglichkeit, etwa Unterrichtsmerkmale auf der Klassenebene aus der Sicht ganzer Klassen oder zumindest mehrerer Schülerinnen und Schüler einer Klasse zu erfassen. Damit ist auch eine analytische Trennung der Klassen- und Schulebene nicht möglich, weil die zufällig gezogenen Fünfzehnjährigen unterschiedliche Klassen und Klassenstufen besuchen. Äußerst begrenzt sind entsprechend die Möglichkeiten von PISA, systematisches und objektivierbares Wissen über Unterrichtsmuster und die Qualität des Unterrichts zu gewinnen (Prenzel & Lankes, 2013). Aussagen sind allenfalls auf einer Systemebene möglich, die beispielsweise den Schultyp in den Fokus nimmt, aber auch hier voraussetzt, dass eine hinreichende Zahl von Teilnehmern in der Grundgesamtheit ist (vgl. Kapitel 2).

Aufgrund des querschnittlichen Designs sind Zusammenhänge zwischen Leistungsmaßen und weiteren Merkmalen, die beispielsweise zur Kompetenzentwicklung beigetragen haben könnten, auf der Basis von Daten aus PISA *nicht kausal* zu interpretieren. Das internationale Design von PISA sieht neben den verbindlichen Tests zwei weitere verpflichtende Elemente vor, und zwar einen Schüler- und einen Schulfragebogen (vgl. auch Kapitel 3 und 5). Der Fragebogen für Schülerinnen und Schüler umfasst Fragen zu Merkmalen der Jugendlichen und ihrer Bildungsbiografie, zu ihrem Elternhaus, Frei-

© Waxmann Verlag GmbH

zeitaktivitäten, Unterricht und Schule. Der genaue Fokus der Fragen richtet sich in der Regel nach der jeweils untersuchten Hauptdomäne einer Erhebungsrunde. Aufgrund der international vorgegebenen alters- und nicht klassenbasierten Schülerstichprobe gelangen die erfassten Einschätzungen der Schülerinnen und Schüler an Grenzen, wenn etwa Unterrichtsmuster beschrieben werden sollen. Der Schulfragebogen ist an die Schulleitung gerichtet und zielt auf Merkmale der Schule, etwa Größe, Kollegium, Profil oder auch Zusammenarbeit mit Eltern. Neben diesen beiden verpflichtenden Fragebögen können optional auch ein Elternfragebogen und seit 2015 ein Lehrerfragebogen eingesetzt werden. Letzterer wurde in Deutschland bereits in früheren PISA-Erhebungsrunden verwendet, mit PISA 2015 sind nun auch vereinzelt internationale Vergleiche möglich (vgl. Kapitel 5).

1.2 Besonderheiten von PISA 2015

Die Erhebungsrunde PISA 2015 nimmt in zweierlei Hinsicht eine Sonderrolle innerhalb der Studie ein. So fand dieses Mal die Datenerhebung (zumindest in den Staaten, die es technisch umsetzen konnten) erstmals vollständig am Computer statt. Es wurde also ein Moduswechsel vollzogen, um einerseits nah an der Lebenswelt der Fünfzehnjährigen zu bleiben und andererseits innovative Test- und Auswertungsverfahren nutzen zu können. Außerdem schließt PISA 2015 den zweiten vollständigen Zyklus ab, indem nun mit den Naturwissenschaften auch die dritte Domäne bereits zum zweiten Mal den Schwerpunkt bildet. Im Folgenden werden diese Besonderheiten im Detail vorgestellt.

1.2.1 Teilnehmende Staaten

An PISA 2015 nahmen insgesamt 72 Staaten teil, im definierten Erhebungszeitraum vom 1. März bis 30. Juni 2015 wurden damit weltweit ungefähr 530.000 Fünfzehnjährige getestet (OECD, 2016b). Neben allen 35 OECD-Staaten beteiligten sich weitere 37 sogenannte OECD-Partnerstaaten an PISA 2015 (vgl. Tabelle 1.1). Damit hat sich im Vergleich mit der ersten Erhebungsrunde im Jahr 2000 der Kreis der Teilnehmer (damals 32 Staaten) mehr als verdoppelt. Die wachsende Zahl der Partnerstaaten zeigt das zunehmende Interesse von Staaten außerhalb der OECD, sich mit PISA an einem internationalen Bildungsmonitoring zu beteiligen.

© Waxmann Verlag GmbH

Tabelle 1.1: An PISA 2015 teilnehmende Staaten

Albanien*	Italien	Portugal
Algerien*	Japan	Republik Mazedonien*
Argentinien*	Jordanien*	Republik Moldau*
Australien	Kanada	Republik Montenegro*
Belgien	Kasachstan*	Republik Serbien*
Brasilien*	Katar*	Republik Trinidad und Tobago*
Bulgarien*	Kolumbien*	Rumänien*
B-S-J-G China*[1]	Korea	Russische Föderation*
Chile	Kosovo*	Schweden
Chinesisch Taipeh*	Kroatien*	Schweiz
Costa Rica*	Lettland	Singapur*
Dänemark	Libanesische Republik*	Slowakische Republik
Deutschland	Litauen*	Slowenien
Dominikanische Republik*	Luxemburg	Spanien
Estland	Macau (China)*	Thailand*
Finnland	Malaysia*	Tschechische Republik
Frankreich	Malta*	Tunesien*
Georgien*	Mexiko	Türkei
Griechenland	Neuseeland	Ungarn
Hongkong (China)*	Niederlande	Uruguay*
Indonesien*	Norwegen	Vereinigte Arabische Emirate*
Irland	Österreich	Vereinigte Staaten
Island	Peru*	Vereinigtes Königreich
Israel	Polen	Vietnam*

Anmerkung: * OECD-Partnerstaaten [1] An PISA 2015 teilgenommen haben die vier Provinzen Peking, Shanghai, Jiangsu und Guangdong

1.2.2 Testkonzeption und Testdesign

Nach PISA 2006 waren in PISA 2015 zum zweiten Mal die Naturwissenschaften Hauptdomäne. Die Nebendomäne Lesen war in PISA 2000 und 2009 der Schwerpunkt, die Mathematik in PISA 2003 und 2012. Die Testkonzeption, die bei PISA 2006 der Aufgabenentwicklung für die Naturwissenschaften als Hauptdomäne zugrunde lag, wurde für die Erhebung 2015 überprüft und an einigen Stellen überarbeitet und erweitert (vgl. Kapitel 2 in Bezug auf die Naturwissenschaften). Inhaltlich ist die für PISA 2015 modifizierte Testkonzeption so angelegt, dass Vergleiche zu PISA 2006, als die Naturwissenschaften zum ersten Mal den Erhebungsschwerpunkt bildeten, und über die weiteren Erhebungsrunden hinweg prinzipiell für die Gesamtskala möglich sind (OECD, 2016a), auch wenn verschiedene Änderungen (z. B. Modus der Erhebung, Präsentation der Aufgaben, Bewertung nicht bearbeiteter Aufgaben, Veränderung des Skalierungsmodells) hier einschränkend wirken. Die Konzeptionen der beiden Nebendomänen wurden

© Waxmann Verlag GmbH

Tabelle 1.2: Organisation der theoretischen Rahmenkonzeptionen der drei kognitiven Domänen in PISA 2015

	Mathematik	Lesen	Naturwissenschaften
Definition	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.	Die Fähigkeit einer Person, geschriebene Texte zu verstehen, zu nutzen 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.	Die Fähigkeit, die charakteristischen Eigenschaften sowie die Bedeutung der Naturwissenschaften in unserer heutigen Welt zu verstehen, naturwissenschaftliches Wissen anzuwenden, um Fragestellungen zu erkennen, naturwissenschaftliche Phänomene zu beschreiben und aus Belegen Schlussfolgerungen zu ziehen, sowie die Bereitschaft, sich reflektierend mit naturwissenschaftlichen Ideen und Themen auseinanderzusetzen.
Inhalte	Vier Inhaltsbereiche: – Quantität – Raum und Form – Veränderung und Beziehungen – Unsicherheit und Daten	Textformate: – kontinuierliche Texte, z. B. Zeitungsartikel, Aufsätze, Romane und Briefe – nichtkontinuierliche Texte, z. B. Tabellen, Grafiken, Diagramme, Stundenpläne, Fahrpläne	Bereiche *naturwissenschaftlichen Wissens*: – physikalische Systeme – lebende Systeme – Erd- und Weltraumsysteme Bereiche des (Meta-)*Wissens über die Naturwissenschaften*: – naturwissenschaftliches Forschen – naturwissenschaftliche Erklärungen
Kontexte bzw. Situationen	Die Rahmenkonzeption unterscheidet folgende Kontexte: – persönliche – berufliche – gesellschaftsbezogene – wissenschaftliche	Die Rahmenkonzeption unterscheidet folgende Situationen bzw. Kontexte: – private Situationen (z. B. persönliche Briefe, E-Mails) – öffentliche Situationen (z. B. offizielle Dokumente, Zeitungsartikel) – bildungsbezogene Situationen (z. B. Lehrbuchtexte, Lernsoftware) – berufsbezogene Situationen (z. B. Stellenausschreibungen in der Zeitung oder im Internet)	Die Rahmenkonzeption unterscheidet fünf Kontexte, die jeweils einen persönlichen, sozialen oder globalen Bezug haben können: – Gesundheit – natürliche Ressourcen – Umwelt – Gefahren – Grenzen von Naturwissenschaft und Technik
Prozesse	– Situationen mathematisch formulieren – anwenden – interpretieren	– Informationen suchen und extrahieren – kombinieren und interpretieren – reflektieren und bewerten	– Phänomene naturwissenschaftlich erkären – naturwissenschaftliche Forschung bewerten und Untersuchungen planen – Daten und Evidenz naturwissenschaftlich interpretieren

© Waxmann Verlag GmbH

gegenüber den jeweils letzten Erhebungsrunden mit ihrem Schwerpunkt nicht verändert, allerdings wurden auch hier das Testmedium (Computer anstelle von Papier-und-Bleistift-Test) und das Skalierungsmodell geändert.

Tabelle 1.2 zeigt, wie die drei kognitiven Domänen in PISA 2015 in der theoretischen Rahmenkonzeption organisiert sind. Ausgehend von einer klaren Definition der Kompetenzen, die gemessen werden sollen, sind die drei Domänen anhand von Inhalten, Prozessen und Kontexten strukturiert:

- *Inhalte*, deren Kenntnis als Bestandteil der Grundbildung in den drei Domänen vorausgesetzt wird
- *Prozesse,* welche die Jugendlichen bei der Lösung der PISA-Aufgaben unter Beweis stellen sollen
- *Kontexte,* in denen die Jugendlichen ihre Kompetenz für eine adäquate Lösung der Aufgabe anwenden können sollen.

Für PISA 2015 wurden 26 Aufgaben für die Domäne Naturwissenschaften mit insgesamt 99 Teilaufgaben (*Items*) neu entwickelt. Davon waren fünf Aufgaben interaktiv (24 Teilaufgaben) und 21 waren Standard-Aufgaben (75 Teilaufgaben). Neben diesen neuen Aufgaben kamen 30 Aufgaben mit 85 Teilaufgaben aus früheren Erhebungsrunden zum Einsatz, anhand derer die Trends berechnet werden konnten. Alle neu entwickelten Aufgaben wurden einem mehrstufigen Begutachtungsverfahren unterzogen, wobei internationale Experten für naturwissenschaftliche Grundbildung ebenso wie nationale Projektmanager und Fachdidaktiker die Aufgaben prüften und Verbesserungsvorschläge machten. Die Aufgaben wurden dann zunächst mit kleinen Schülergruppen erprobt und diskutiert, ehe sie einer erneuten Begutachtung unterzogen wurden. Die Begutachtung erfolgte nach mehreren Kriterien, etwa fachliche Richtigkeit, Aufgabenschwierigkeit, Bezüge zum Lehrplan oder Relevanz für die Lebenswelt heutiger Fünfzehnjähriger. Auf die Vermeidung von kultureller oder geschlechtsbezogener Benachteiligung wurde dabei selbstverständlich geachtet. Alle neu entwickelten Aufgaben wurden in einem sogenannten Feldtest, der 2014 in allen teilnehmenden Staaten durchgeführt wurde, systematisch erprobt.

Für die Erfassung der Nebendomänen Lesen und Mathematik wurden wie bisher keine neuen Aufgaben eingesetzt, sondern ausschließlich Aufgaben, die für frühere PISA-Erhebungsrunden entwickelt wurden (vgl. OECD, 2016a). Entsprechend ist der Bezugspunkt für die Lesekompetenz PISA 2009 (OECD, 2009) und für die mathematische Kompetenz PISA 2012 (OECD, 2013). Für die Tests im Bereich Lesen wurden in PISA 2015 insgesamt 25 Aufgaben (88 Teilaufgaben) und im Bereich Mathematik 45 Aufgaben (81 Teilaufgaben) eingesetzt.

Bei PISA 2015 wurde für den Bereich Naturwissenschaften Aufgabenmaterial für eine Testzeit von insgesamt 360 Minuten verwendet. Die Aufgaben wurden in sogenannte *Cluster* gruppiert, von denen jedes Aufgaben im Umfang von etwa 30 Minuten enthielt. Die Cluster sind jeweils sortenrein, d. h. sie umfassen ausschließlich Aufgaben aus einer der drei Domänen. Die einzelnen Aufgaben erscheinen dabei in mehre-

© Waxmann Verlag GmbH

ren Clustern und dort an unterschiedlichen Positionen. Ein Cluster besteht jeweils entweder ausschließlich aus bereits früher eingesetzten Aufgaben (*Link-Aufgaben*) oder aus neu entwickelten Aufgaben im Bereich Naturwissenschaften. Der PISA-Test wird dann für jede Schülerin und jeden Schüler so zusammengestellt, dass vier Cluster zur Bearbeitung vorgelegt werden (4 x 30 Minuten). Entsprechend war eine Bearbeitungszeit von 120 Minuten zuzüglich einer Pause nach 60 Minuten vorgesehen. Dieser in PISA gewählte Ansatz zur Zusammenstellung der Aufgaben in Clustern (und diese dann in sogenannte Testhefte bzw. Testformen bei Bearbeitung am PC) wird als *Multi-Matrix-Design* bezeichnet. In PISA 2015 kamen insgesamt 66 Testformen zum Einsatz. Diese wurden systematisch aus insgesamt sechs Trend-Clustern (also aus Aufgaben aus früheren PISA-Erhebungsrunden) der drei Domänen sowie sechs Clustern mit neuen naturwissenschaftlichen Testaufgaben zusammengesetzt. Die Testformen wurden den Schülerinnen und Schülern zufallsbasiert zugeteilt. Weil dabei jede Schülerin und jeder Schüler nur einen kleinen Teil der in PISA eingesetzten Aufgaben bearbeitet, ist die Messung der Kompetenz für keinen der teilnehmenden Schüler vollständig. Deshalb ist es auch nicht sinnvoll, auf individueller Ebene ein bestimmtes Kompetenzniveau anzunehmen – denn dies wurde mit dem Multi-Matrix-Design nicht gemessen. Vielmehr wurden die Aufgaben so auf verschiedene Testformen verteilt, dass jede Aufgabe von einer ausreichend großen Stichprobe Fünfzehnjähriger bearbeitet wurde. Die teilnehmenden Schülerinnen und Schüler werden bei der Auswertung ebenso wie alle eingesetzten Testaufgaben als Gesamtgruppe betrachtet, sodass alle PISA-Teilnehmer pro Staat gemeinsam alle eingesetzten Testaufgaben bearbeitet haben (aber eben nicht jeder einzelne Jugendliche jede Aufgabe). Jede Aufgabe wurde demnach von einem bestimmten Anteil der Schülerstichprobe bearbeitet (z. B. von zwei Dritteln aller teilnehmenden Schülerinnen und Schüler). Auf dieser Basis wird nun aufgrund aller Aufgaben zu einer Domäne auf die durchschnittliche Kompetenz aller Schülerinnen und Schüler der PISA-Stichprobe eines Staates geschlossen – und damit deren Kompetenz letztlich geschätzt. Dieser Ansatz funktioniert explizit nur auf der Ebene der Gesamtstichprobe und nicht auf der Ebene des einzelnen Jugendlichen (Sälzer, 2016). Die Punktwerte, die mithilfe des Multi-Matrix-Designs ermittelt werden, können mit speziellen Auswertungsverfahren der *Item Response Theory* (vgl. Kapitel 12) ausgewertet und direkt miteinander verglichen und interpretiert werden.

Im Anschluss an den Aufgabenteil im PISA-Test bearbeiteten die Schülerinnen und Schüler einen Fragebogen, der entsprechend der theoretischen Rahmenkonzeption (OECD, 2016a) zusammengestellt wurde (vgl. auch Kapitel 13). Ebenfalls zum Pflichtprogramm bei PISA 2015 gehörte ein Fragebogen für die Schulleitung. Als internationale Optionen wurden ein Fragebogen für Eltern sowie ein Fragebogen für Lehrkräfte angeboten.

© Waxmann Verlag GmbH

1.2.3 Computerbasiertes Testen

PISA 2015 hatte im Vergleich zu früheren Erhebungsrunden ein besonders innovatives Element: Erstmals fand die Erhebung in Deutschland und der Mehrheit der anderen Staaten vollständig am Computer statt und nicht mehr als Papier-und-Bleistift-Test. Bereits seit PISA 2003 waren immer wieder einzelne Komponenten auch am Computer administriert worden, teilweise auch nur in einzelnen Staaten. PISA 2009 als erste Runde des zweiten Zyklus umfasste mit dem *Electronic Reading Assessment* (ERA) eine umfangreiche digitale Erhebung, die in PISA 2012 durch das *Computer-Based Reading* und *Computer-Based Mathematics* sowie einen Test im *Computer-Based Problem-Solving* noch erweitert wurde. Mit der Umstellung bei PISA 2015 vom papierbasierten Testmodus auf die computergestützte Erfassung von Kompetenzen sind Implikationen verbunden, die in Kapitel 12 näher ausgeführt werden. So müssen beispielsweise Aufgaben, die bereits in früheren PISA-Erhebungsrunden zum Einsatz kamen und der Abbildung von Trends dienen, von einer papierbasierten Form auf die Darstellung am Computerbildschirm übertragen werden. Dabei kann es vorkommen, dass Aufgaben oder Teile daraus für die Darstellung am Bildschirm nicht geeignet sind. Dies kommt etwa vor, wenn Streckenlängen geschätzt werden sollen, für die man auf Papier bestimmte Punkte und Abstände nutzen konnte und deren Lage und Betrag am Computer von der Bildschirmauflösung und -konfiguration abhängen; sie sind damit nicht mehr eindeutig. Außerdem können Aufgaben am Computer systematisch leichter oder schwieriger sein, als sie dies in papierbasierter Form waren. Für neu entwickelte Aufgaben bietet der Wechsel des Testmodus aber viele Vorteile, beispielsweise die Möglichkeit interaktiver Aufgaben.

Um mögliche Auswirkungen des Moduswechsels von Kompetenztests in Papierform hin zu computerbasiertem Testen zu untersuchen, führte das internationale Konsortium für PISA 2015 unter Federführung des *Educational Testing Service* (ETS) eine sogenannte Moduseffekt-Studie durch (vgl. Kapitel 12). Diese fand im Rahmen des Feldtests zu PISA 2015 im Frühjahr 2014 statt. Dabei wurde geprüft, ob die Schwierigkeiten und Trennschärfen der Aufgaben je nach gewähltem Modus variieren, also ob die Aufgaben durch den Wechsel von Papier-und-Bleistift auf Computer leichter oder schwieriger werden – oder ihre Schwierigkeit unverändert bleibt. Auf internationaler Ebene, das heißt bei Betrachtung aller PISA-Teilnehmerstaaten, waren keine Moduseffekte zu erkennen. Items, die in Form des Papier-und-Bleistift-Tests administriert wurden, hatten also gleiche Itemparameter wie in dem Fall, in dem sie computerbasiert administriert wurden. Dies schließt jedoch nicht aus, dass es in einzelnen Staaten durchaus Veränderungen in den Ergebnissen gibt, die zumindest zum Teil mit dem Wechsel des Testmodus zusammenhängen. Allein die unterschiedlich häufige Nutzung von Computern im Unterricht in verschiedenen Bildungssystemen kann bedeuten, dass für eher ungeübte Schülerinnen und Schüler die Bearbeitung eines Tests am Computer schwieriger ist als auf Papier. Für Deutschland liefert der Feldtest 2014 Hinweise darauf, dass die PISA-Aufgaben im Mittel am Computer schwieriger waren als auf Papier. Dabei scheinen die Moduseffekte durch den Computer bei den Naturwissenschaftsitems am größten, bei den Leseitems

© Waxmann Verlag GmbH

am geringsten zu sein. Entsprechende weiterführende wissenschaftliche Publikationen sind seitens des Zentrums für Internationale Bildungsvergleichsstudien in Vorbereitung.

Logistisch brachte die Umstellung auf computerbasiertes Testen ebenfalls einige Neuerungen mit sich. So nahmen Testleiterinnen und Testleiter Kontakt zur Schulkoordinatorin bzw. zum Schulkoordinator auf und vereinbarten einen Termin für die sogenannte Systemdiagnose. Diese Systemdiagnose war die Überprüfung, ob die Schule über die für eine computerbasierte PISA-Testung benötigte Infrastruktur verfügte. Dazu zählte etwa die Anzahl verfügbarer Computer mit der definierten Ausstattung, die Anzahl der Computerräume und der administrativen Rahmenbedingungen. Internetzugang war nicht erforderlich, da der Test anhand eines USB-Sticks administriert wurde. Schulen, welche die Mindestanforderungen an die Computer-Infrastruktur nicht erfüllten, wurden mit Laptops versorgt.

Im Zuge der Umstellung auf eine rein computergestützte Erhebung (siehe dazu auch Kapitel 12) wurden den Jugendlichen im Rahmen von PISA 2015 auch Fragen zur Verfügbarkeit von Computern in der Schule, zur Nutzung des Computers innerhalb und außerhalb der Schule sowie zur Rolle von Computern in ihrem Alltag gestellt. Der entsprechende Fragebogenabschnitt erfasst die Vertrautheit der Jugendlichen mit ICT (*Information and Communications Technology)* und kam teilweise bereits bei PISA 2012 zum Einsatz.

Wenn die Ergebnisse aus PISA 2015 mit den Befunden aus früheren PISA-Erhebungsrunden verglichen werden sollen, muss der Wechsel des Testmodus von Papier-und-Bleistift auf Computer bei der Interpretation berücksichtigt werden. Um die in diesem Band beschriebenen Ergebnisse auch vor dem Hintergrund der Nutzungsgewohnheiten und Einstellungen der Jugendlichen interpretieren zu können, werden einige Angaben zur Computernutzung vorgestellt.

Tabelle 1.3: Computernutzung Fünfzehnjähriger bei PISA 2015 in Deutschland

	Gesamt		Mädchen		Jungen	
Computernutzung	*M**	*SE*	*M**	*SE*	*M**	*SE*
ICT-Nutzung in der Schule allgemein	*-0.42*	(0.02)	-0.48[a]	(0.02)	-0.35[a]	(0.02)
ICT-Nutzung außerhalb der Schule für Hausaufgaben	*-0.38*	(0.01)	-0.38	(0.02)	-0.38	(0.02)
ICT-Nutzung außerhalb der Schule als Freizeitbeschäftigung	*-0.29*	(0.01)	-0.51[a]	(0.02)	-0.07[a]	(0.02)
ICT-Ressourcen	*0.06*	(0.01)	0.02[a]	(0.01)	0.10[a]	(0.02)
ICT-Interesse	*0.05*	(0.01)	-0.04[a]	(0.02)	0.13[a]	(0.02)
Wahrgenommene ICT-Kompetenz	*-0.05*	(0.01)	-0.35[a]	(0.02)	0.26[a]	(0.02)
Wahrgenommene ICT-Autonomie	*0.20*	(0.02)	-0.19[a]	(0.02)	0.60[a]	(0.02)
ICT in sozialer Interaktion	*-0.19*	(0.01)	-0.54[a]	(0.02)	0.17[a]	(0.02)

* Die Ausprägungen aller ICT-Konstrukte sind immer am OECD-Mittelwert von 0,00 und einer Standardabweichung von 1,00 standardisiert. Ergebnisse müssen daher immer im Vergleich zum OECD-Mittelwert interpretiert werden.
kursiv: Signifikante Unterschiede zum OECD-Mittelwert
[a]**:** Signifikante Unterschiede zwischen Mädchen und Jungen

© Waxmann Verlag GmbH

Die in Tabelle 1.3 dargestellten Werte sind auf einer Logit-Skala abgebildet und beziehen sich auf den auf 0 normierten OECD-Mittelwert (*SD* = 1.0). Sie sind daher stets in Bezug auf den OECD-Mittelwert zu lesen, sodass positive Werte eine überdurchschnittliche Ausprägung bedeuten und negative Werte eine unterdurchschnittliche. Wie in Tabelle 1.3 abgebildet, nutzen die Fünfzehnjährigen in Deutschland Computer sowohl innerhalb der Schule als auch außerhalb der Schule seltener, als dies im Mittel in den OECD-Staaten der Fall ist. Jungen geben dabei deutlich häufiger an, Computer außerhalb der Schule als Freizeitbeschäftigung zu nutzen, als Mädchen.

In Bezug auf ihre Vertrautheit mit Computern unterscheiden sich die Fünfzehnjährigen in Deutschland ebenfalls signifikant vom OECD-Durchschnitt. Während die verfügbaren ICT-Ressourcen und das Interesse der Fünfzehnjährigen in Deutschland an ICT vergleichsweise nahe am OECD-Mittelwert liegen, schätzen sich fünfzehnjährige Schülerinnen und Schüler in Deutschland als autonomer im Umgang mit dem Computer ein als im OECD-Durchschnitt. Fünfzehnjährige in Deutschland erleben sich also als selbstständig, wenn es um das Bedienen von Computern und ähnlicher Geräte geht, und trauen sich beispielsweise zu, Probleme selbst zu lösen. Unterschiede zwischen Mädchen und Jungen zeigen sich in Deutschland in Bezug darauf, inwieweit sie sich aus ihrer Sicht aus eigenem Antrieb mit dem Computer beschäftigen und wie gut sie in ihrer Wahrnehmung mit Computern umgehen können. Mädchen schätzen sich hier als deutlich weniger kompetent im Umgang mit dem Computer ein, als die Jungen es tun. Ein weiterer Geschlechterunterschied findet sich bei der Rolle von ICT in Form von sozialen Interaktionen, also etwa gemeinsamem Computerspielen mit Freundinnen und Freunden oder Informationsaustausch über Probleme mit digitalen Geräten. Die Mädchen nutzen Computer deutlich seltener für soziale Interaktionen als der OECD-Mittelwert, aber auch deutlich seltener als Jungen in Deutschland (diese liegen mit ihrer Selbsteinschätzung über den OECD-Durchschnitt). Auch in dieser Hinsicht geben die Jungen also an, das Thema der Informations- und Kommunikationstechnologien deutlich mehr für zwischenmenschliche Kommunikation und Interaktion zu nutzen als die Mädchen.

Insgesamt lässt dieser aktuelle Stand darauf schließen, dass der Einsatz von Computern in Schulen in Deutschland einerseits nach wie vor nicht selbstverständlich ist und Computer andererseits auch deutlich seltener als in den OECD-Staaten für Freizeitaktivitäten genutzt werden. Allerdings treten auch Unterschiede im Nutzungsverhalten zwischen Mädchen und Jungen zutage, die bei der Interpretation der in diesem Band berichteten Ergebnisse berücksichtigt werden sollten.

1.3 Durchführung von PISA 2015 in Deutschland

Eine vollständige PISA-Erhebungsrunde dauert insgesamt etwa vier Jahre von der Entwicklung der theoretischen Rahmenkonzeptionen bis hin zur Berichterstattung. In allen Phasen des Projekts arbeiten international besetzte Expertengruppen und Institutionen mit den nationalen Projektmanagern zusammen, wobei auf internationaler Ebene über-

© Waxmann Verlag GmbH

greifende Pläne und Arbeitspakete definiert werden und seitens der nationalen Projektmanager eine Begutachtung der Passung für das jeweilige Land erfolgt. Auch gezielte inhaltliche Beiträge wie Vorschläge für Testaufgaben oder deren Optimierung werden von den nationalen Projektmanagern eingebracht. Um ein komplexes und umfangreiches Projekt wie PISA erfolgreich durchführen zu können, müssen die Arbeitsschritte international verbindlich festgelegt und terminiert werden. So können methodische Standards sowie die internationale Vergleichbarkeit gesichert werden. Im Folgenden werden die zentralen Arbeitsschritte und Anforderungen mit besonderem Fokus auf die Erhebung in Deutschland beschrieben.

1.3.1 Population und Stichprobe: Wer nimmt an PISA teil?

PISA hat das Anliegen, den Kompetenzstand fünfzehnjähriger Schülerinnen und Schüler in den teilnehmenden Staaten abzubilden. Die untersuchte Population besteht demnach aus fünfzehnjährigen Schülerinnen und Schülern, die in ihrem jeweiligen Staat oder Bildungssystem eine Schule besuchen. Bei der Stichprobenziehung muss sichergestellt werden, dass möglichst alle Schulen, Schülerinnen und Schüler, die in die Stichprobe gezogen wurden, auch tatsächlich an der Erhebung teilnehmen. Nur so kann die Repräsentativität gesichert werden. In den folgenden Abschnitten wird skizziert, wie dies in PISA 2015 geschah. Details zur Stichprobenziehung finden sich in Kapitel 12 in diesem Band sowie in diversen *Technical Reports* zu PISA (u. a. OECD, 2012; 2014).

Ziehung der repräsentativen Stichprobe

Das Schließen auf die in PISA untersuchte Population fünfzehnjähriger Schülerinnen und Schüler in den beteiligten Staaten erfordert die Ziehung einer national repräsentativen Stichprobe. Entsprechend ist auch die PISA-Stichprobe in Deutschland definiert. Sie erlaubt allerdings keine Vergleiche etwa zwischen den Ländern innerhalb Deutschlands. Vergleiche zwischen den Bundesländern sind den standardbezogenen Leistungsvergleichen vorbehalten, die vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) regelmäßig durchgeführt werden. Der letzte bundesweite Ländervergleich Sprachen zum Erreichen von Bildungsstandards der Sekundarstufe wurde im Herbst 2016 veröffentlicht (Stanat, Böhme, Schipolowski & Haag, 2016).

In Deutschland haben an PISA 2015 insgesamt 6504 Fünfzehnjährige an allen Formen allgemeinbildender Schulen der Sekundarstufe I sowie an beruflichen und an Förderschulen (insgesamt 253 Schulen, vgl. Tabelle 1.4) teilgenommen. Diese Stichprobe wurde in Form einer nationalen Option erweitert, um gezielt Forschungsfragen beantworten zu können, die über das übliche PISA-Pflichtprogramm hinausreichen. Während die Standardstichprobe in PISA altersbasiert ist, um auf die Population der Fünfzehnjährigen schließen zu können, besteht die Möglichkeit, eine klassenbasierte Zusatzstichprobe zu ziehen. Dies hat den Vorteil, dass gegenüber der altersbasierten Stichprobe, die

© Waxmann Verlag GmbH

sich auf unterschiedliche Klassen und Klassenstufen verteilt, bei einer klassenbasierten Stichprobe auch geteilte Wahrnehmungen von Schülerinnen und Schülern erfasst werden können. Anhand dieser Stichprobe können Skalen etwa zur Unterrichtswahrnehmung weiterführend und mit Bezug auf die ganze Klasse ausgewertet werden, da sich Schülerinnen und Schüler derselben Klasse bei der Frage nach Elementen ihres Unterrichts in Physik, Chemie oder Biologie auf dieselbe Lehrkraft und gemeinsam erlebte Unterrichtssituationen beziehen können. Aus jeder ausgewählten Schule wurde für die optionale Klassenstichprobe zufallsbasiert eine 9. Klasse gezogen. Die Wahl fiel auf diese Klassenstufe, weil die Mehrzahl der fünfzehnjährigen Schülerinnen und Schüler in Deutschland diese Stufe besucht. Aus Gründen der Testökonomie wurde aus dieser gezogenen 9. Klasse pro Schule eine Zufallsauswahl von 15 Schülerinnen und Schülern getroffen, was bei einer durchschnittlichen Klassengröße von 25 Schülerinnen und Schülern als verlässliches, repräsentatives Abbild der geteilten Wahrnehmung des schulischen Umfelds gelten darf. Diese 15 Schülerinnen und Schüler wurden als eine separate Testgruppe gemeinsam getestet und befragt.

Tabelle 1.4: Zusammensetzung der Stichprobe in den an PISA 2015 teilnehmenden Schulen

Allgemeinbildende Schulen (ohne Sonder- und Förderschulen)	Sonder- und Förderschulen	Berufliche Schulen
230	12	11
Max. 30 Fünfzehnjährige + 15 Schüler aus einer 9. Klasse	Max. 30 Fünfzehnjährige	Max. 30 Fünfzehnjährige

Die Repräsentativität der Stichprobe wird im Rahmen von PISA durch ein genau beschriebenes, verbindliches und mehrstufiges Ziehungsverfahren gewährleistet. Gezogen wird eine mehrfach geschichtete (stratifizierte) Wahrscheinlichkeitsstichprobe, die in Deutschland wie in jedem Teilnehmerstaat einige Besonderheiten (z. B. Unterteilung in Bundesländer, mehrere Schulformen der Sekundarstufe) berücksichtigen muss. Vorbereitend wird unter Bezug auf Daten der statistischen Landesämter zunächst eine vollständige Liste aller Schulen in Deutschland erstellt, an denen sich potenziell fünfzehnjährige Schülerinnen und Schüler finden lassen. Als „fünfzehnjährig“ gilt, wer im definierten Geburtszeitraum vom 1. Januar 1999 bis zum 31. Dezember 1999 geboren wurde. Diese Schulliste repräsentiert die Population (Grundgesamtheit), aus der die an PISA teilnehmenden Schulen zufallsbasiert gezogen werden. Um Schulen aus allen Bundesländern in der Stichprobe zu finden, wird die Schulliste in die 16 Bundesländer unterteilt. Ausgehend von der aktuell vorhandenen Anzahl fünfzehnjähriger Schülerinnen und Schüler sowie der Schularten der Sekundarstufe wird bestimmt, wie viele Schulen und wie viele Schülerinnen und Schüler pro Bundesland gezogen werden müssen. Diese Unterteilung wird Stratifizierung (Schichtung) genannt, wobei jedes Bundesland ein Stratum

© Waxmann Verlag GmbH

(eine Schicht) bildet. Neben den Bundesländern werden in PISA für Deutschland zwei weitere Strata bestimmt, und zwar berufliche Schulen sowie Sonder- und Förderschulen. Durch diese Stratifizierung wird sichergestellt, dass genau so viele Schulen und Schülerinnen und Schüler aus allen Bundesländern und Schulformen, die von Fünfzehnjährigen besucht werden können, in die Stichprobe gezogen werden. Indem für jede gezogene Schule nach demselben Prinzip zwei Ersatzschulen bestimmt werden, die im Falle alternativloser Nichtteilnahme von Schulen (etwa weil diese keine fünfzehnjährigen Schülerinnen und Schüler hat) nachrücken und statt der ursprünglich gezogenen Schule an PISA teilnehmen, kann der Anspruch der Repräsentativität auch bei Ausfall gezogener Schulen umgesetzt werden. Nach erfolgter Ziehung der Schulen werden pro ausgewählter Schule zufallsbasiert 30 fünfzehnjährige Schülerinnen und Schüler zur Teilnahme an PISA ausgewählt. Nachdem diese altersbasierte Stichprobe gezogen ist, erfolgt die jahrgangsbasierte Ziehung der optionalen 9. Klasse und daraus dann ebenfalls zufallsbasiert die Ziehung von 15 Schülerinnen und Schülern. Ursprünglich waren 267 Schulen für die Teilnahme an PISA 2015 gezogen worden. Von diesen Schulen mussten zehn aus der Stichprobe ausgeschlossen werden, da sie keine fünfzehnjährigen Schülerinnen und Schüler aufwiesen. Eine weitere Schule war zwischenzeitlich geschlossen worden und konnte daher ebenfalls nicht am PISA-Test teilnehmen. Drei private Schulen sagten ihre Teilnahme ab; private Schulen können seitens der KMK nicht zur Teilnahme an PISA verpflichtet werden, weshalb ihre Beteiligung freiwillig ist. Wiederum drei Schulen, die zwar an PISA teilgenommen hatten, wiesen eine zu geringe Beteiligung auf Schülerebene auf (weniger als 50 Prozent der gezogenen Schülerinnen und Schüler) und wurden aus der Datenanalyse ausgeschlossen. Fünf Schulen schließlich konnten aufgrund schwerwiegender struktureller Belastungen wie Umzug oder Gebäudeabriss den PISA-Test nicht ermöglichen. Gemäß den Regeln der Stichprobenziehung wurden acht ursprünglich gezogene Schulen, die nicht an der Erhebung teilnehmen konnten, durch ihre erste bzw. zweite Ersatzschule ersetzt. Tabelle 1.5 zeigt Angaben zur Zielpopulation und zum Ausschöpfungsgrad der PISA-Stichprobe in Deutschland.

Tabelle 1.5: Zielpopulation und Ausschöpfungsgrad bei PISA 2015 in Deutschland

	Ausschlüsse auf Schulebene		**Ausschlüsse in den Schulen**		**Erreichte Population**	
15-Jährige in Schulausbildung [a] (PISA-Population)	absolut [b]	in Prozent der Zielpopulation	absolut	in Prozent der Zielpopulation	absolut	in Prozent der Zielpopulation
774 149	11 150	1.44	628	0.71	762 999	98.6%

a Quelle: Statistisches Bundesamt, Fachserie 11
b Quelle: Sampling Frame

© Waxmann Verlag GmbH

Damit verblieben letztlich 253 Schulen, an denen PISA 2015 in Deutschland durchgeführt wurde. Die gewichtete Teilnahmequote auf Schulebene beträgt demnach 98.9 Prozent und drückt aus, wie viele der gezogenen Schulen tatsächlich an PISA 2015 teilgenommen haben. Auf Schülerebene betrug die gewichtete Teilnahmequote 93.2 Prozent. Beide Werte erfüllen die von der OECD gesetzten Mindeststandards. An beruflichen Schulen sowie an Sonder- und Förderschulen wurde jeweils eine Vollerhebung aller fünfzehnjährigen Schülerinnen und Schüler durchgeführt; an den allgemeinbildenden Schulen mit mehr als 30 Fünfzehnjährigen erfolgte eine zufallsbasierte Stichprobenziehung.

Zum Zeitpunkt der Stichprobenziehung und auch der Datenerhebung in PISA 2015 befand sich die Sekundarschullandschaft in mehreren deutschen Bundesländern im Umbruch. Das Gymnasium ist dabei die einzige Schulart, die nach wie vor in allen 16 Bundesländern existiert und eine gemeinsame Entstehungsgeschichte hat. Andere Schulen der Sekundarstufe wurden jedoch in einigen Ländern in neue oder teilweise neue Schularten überführt, womit auch eine neue Bezeichnung einherging. In einem Großteil der Bundesländer wurden eigenständige Haupt- und Realschulen abgeschafft und durch kombinierte Schularten ersetzt, die mehrere Bildungsgänge und Abschlüsse anbieten. Diese kombinierten Schularten sind landesspezifisch sehr unterschiedlich ausgestaltet, sodass auf der Basis der Schulstatistik keine zuverlässige Aufgliederung in Schularten mit mehreren Bildungsgängen oder Gesamtschulen erfolgen kann (Autorengruppe Bildungsberichterstattung, 2016). Entsprechend werden im vorliegenden nationalen Berichtsband zwei Gruppen von Schularten unterschieden, nämlich Gymnasien und nicht gymnasiale Schularten. Unter den nicht gymnasialen Schularten sind folglich Hauptschulen, Realschulen, Integrierte Gesamtschulen sowie Schulen mit mehreren Bildungsgängen zusammengefasst. Förderschulen sowie berufliche Schulen gehen ausschließlich in die Analysen der Gesamtstichprobe ein. Tabelle 1.6 zeigt die Untersu-

Tabelle 1.6: Untersuchungsbeteilung der Fünfzehnjährigen nach Schulart und Klassenstufe

Klassenstufe	Hauptschule	Schule mit mehreren Bildungsgängen	Integrierte Gesamtschule	Realschule	Gymnasium	Berufsschule	Sonder- und Förderschule	Summe
insgesamt gezogen	808	924	910	1 699	2 428	214	206	7 189
7	12	3	1	10	1		1	
8	138	77	53	145	55		32	
9	414	477	394	787	905		127	
10 und höher	148	297	354	634	1 305		46	
keine Zuordnung						160		
insgesamt getestet	712	854	802	1 576	2 266	160	134	6 504
Ausschöpfung in %	88.1	92.4	88.1	92.8	93.3	74.8	65.0	92.2

Anmerkungen. Die Tabelle berichtet die realisierten Stichproben ungewichtet. Die gewichtete Ausschöpfung beträgt 93.2 Prozent.

© Waxmann Verlag GmbH

chungsbeteiligung der Fünfzehnjährigen differenziert nach den Schularten, die der Listung zum Zeitpunkt der Stichprobenziehung entspricht. Die als nicht gymnasiale Schularten berichteten Schulen wurden also durchaus noch beispielsweise als Haupt- oder Realschule in die Stichprobe gezogen und für die Berichterstattung zusammengefasst.

Je nachdem, ob Fünfzehnjährige eine große Schule mit einer vergleichsweise hohen Zahl von Gleichaltrigen oder eine kleine Schule besuchen, konkurrieren sie sozusagen mit mehr oder weniger anderen Fünfzehnjährigen um die Zufallsauswahl in die PISA-Stichprobe. Je mehr Fünfzehnjährige an einer Schule sind, desto geringer ist die Ziehungswahrscheinlichkeit für die einzelne Schülerin oder den einzelnen Schüler. Diese unterschiedliche Ziehungswahrscheinlichkeit muss vor der Auswertung der Daten ausgeglichen werden, um Verzerrungen auszuschließen (vgl. auch Kapitel 12). Ein solcher Ausgleich erfolgt über eine Gewichtung der Daten, die dafür sorgt, dass jede tatsächlich getestete Schülerin und jeder tatsächlich getestete Schüler eine bestimmte Anzahl von Schülerinnen und Schülern der Population repräsentiert. Analog zur Ziehung der Stichprobe erfolgt die Berechnung der Gewichte in mehreren Stufen. Als Erstes wird jeder gezogenen Schule ein Gewicht zugewiesen, das sie umgekehrt proportional zu ihrer Ziehungswahrscheinlichkeit höher oder niedriger gewichtet. Eine Schule, die höher gewichtet wird, repräsentiert mehr andere Schulen und eine Schule, die niedriger gewichtet wird, entsprechend weniger andere Schulen. Gibt es beispielsweise in einem Bundesland sehr viele Gymnasien, so hat bei einer Zufallsziehung jedes einzelne Gymnasium eine relativ geringe Chance, in die Stichprobe zu gelangen. Eine Realschule im selben Bundesland kann eine deutlich höhere Ziehungswahrscheinlichkeit haben, wenn es dort nur recht wenige Realschulen gibt. Im ersten Schritt der Gewichtung wird also zunächst diese unterschiedliche Ziehungswahrscheinlichkeit ausgeglichen. So entsteht das sogenannte Schulgewicht. Die zweite Stufe der Gewichtung findet innerhalb der Schulen, nämlich auf Schülerebene statt. In jeder gezogenen Schule werden zufallsbasiert 30 fünfzehnjährige Schülerinnen und Schüler aus allen dort registrierten Fünfzehnjährigen gezogen. Abhängig von der Gesamtzahl der Schülerinnen und Schüler, die die PISA-Teilnahmekriterien erfüllen, hat jede und jeder Einzelne eine unterschiedliche Ziehungswahrscheinlichkeit, die mithilfe des sogenannten Schülergewichts ausgeglichen wird. Zusammen mit dem Schulgewicht und unter Berücksichtigung von Schülerinnen und Schülern, die in die PISA-Stichprobe gekommen sind, jedoch am Testtag nicht anwesend waren, wird im letzten Schritt das Schülergesamtgewicht errechnet. Erst dieses Schülergesamtgewicht macht es möglich, dass jeder an PISA teilnehmende Staat repräsentativ abgebildet wird und darüber hinaus unabhängig von seiner Größe, Bevölkerungsdichte und der Struktur seines Schulsystems in gleichem Maße in die Auswertungen eingeht.

© Waxmann Verlag GmbH

1.3.2 Teilnahmeverpflichtung und Ausschlussgründe

Deutschland beteiligt sich auf Beschluss der KMK und des BMBF an PISA. Demgemäß sind alle staatlichen Schulen, die in die PISA-Stichprobe gezogen wurden, zur Teilnahme verpflichtet. Innerhalb der Schulen sind die Schülerinnen und Schüler verpflichtet, die Kompetenztests zu bearbeiten. Unterschiede zwischen den 16 deutschen Bundesländern bestanden bei PISA 2015 – wie auch in früheren Erhebungsrunden – lediglich in Bezug auf die Verpflichtung zur Bearbeitung des Schülerfragebogens. In sieben Bundesländern war auch das Ausfüllen des Schülerfragebogens obligatorisch, nämlich in Brandenburg, Bremen, Hessen, Mecklenburg-Vorpommern, Niedersachsen, Sachsen-Anhalt und Thüringen. In acht Bundesländern erfolgte die Bearbeitung des Schülerfragebogens auf freiwilliger Basis, wofür vorab eine schriftliche Genehmigung der Erziehungsberechtigten (sogenannte Elterngenehmigung) eingeholt werden musste. Ohne diese Genehmigung durften die betreffenden Schülerinnen und Schüler keinen Fragebogen erhalten. Die acht Bundesländer mit freiwilliger Bearbeitung des Schülerfragebogens waren Baden-Württemberg, Bayern, Hamburg, Nordrhein-Westfalen, Rheinland-Pfalz, Saarland, Sachsen und Schleswig-Holstein. In Berlin galt eine Teilverpflichtung, sodass Teile des Fragebogens zwingend auszufüllen waren und andere Teile auf freiwilliger Basis. Entsprechend unterschiedlich sind die Beteiligungsquoten am Schülerfragebogen im Vergleich der Bundesländer. In den sieben Bundesländern, die eine Bearbeitung des Schülerfragebogens als verpflichtend festgelegt hatten, füllten zwischen 82 und 97 Prozent der Schülerinnen und Schüler den Fragebogen aus. In den acht Ländern, in denen eine Elterngenehmigung notwendig war, lag die Beteiligung zwischen 62 und 91 Prozent. In Berlin wurde ein Rücklauf von 92 Prozent erzielt. Insgesamt lag die Beteiligung am Schülerfragebogen in Deutschland bei 85 Prozent.

Unter Berücksichtigung der generellen Verpflichtung zur Teilnahme an den PISA-Kompetenztests wurden zur Sicherung der internationalen Vergleichbarkeit auch mehrere sogenannte Ausschlussgründe definiert, bei deren Vorliegen eine gezogene Schülerin oder ein gezogener Schüler nicht an PISA teilnehmen darf. So muss etwa ein Jugendlicher mit Zuwanderungshintergrund mindestens ein Jahr vor dem PISA-Test eine Schule in dem Staat besucht haben, in dem er für die PISA-Stichprobe ausgewählt wurde. So kann der Anspruch, dass der PISA-Test in der jeweiligen Unterrichtssprache bearbeitet werden muss, fair und vergleichbar umgesetzt werden. Darüber hinaus muss eine Schülerin oder ein Schüler den PISA-Test selbstständig und ohne fremde Hilfe bearbeiten können, was insbesondere bei Schülerinnen und Schülern mit besonderem Förderbedarf immer wieder zu Ausschlüssen von der PISA-Erhebung führt, etwa weil ihr Hör- oder Sehvermögen eingeschränkt sind und es derzeit noch keine Varianten wie Vorlesen oder visuelle Einführungen (die vom Testleiter vor der Testgruppe vorgelesen werden) gibt. Solche Hilfestellungen sind allerdings für künftige PISA-Erhebungsrunden geplant. Ein dritter Ausschlussgrund ist die unzureichende Beherrschung der Unterrichtssprache, die von der Schulkoordinatorin oder dem Schulkoordinator eingeschätzt wird. Teilweise treffen mehrere Ausschlussgründe auf einzelne Fünfzehnjährige zu, was

© Waxmann Verlag GmbH

detailliert dokumentiert und auf internationaler Ebene genau überprüft wird, um etwa leistungsbezogene Ausschlüsse zu verhindern.

1.3.3 Ablauf eines PISA-Testtags

Nachdem die Schulen im Herbst 2014 über ihre Teilnahme an PISA 2015 informiert worden waren, wurde die Ziehung der Schülerstichprobe vorbereitet (vgl. Kapitel 12). Jeder Schule wurden gezielt geschulte Testleiterinnen und Testleiter zugewiesen, die meist Studierende höherer Semester in den Bereichen Lehramt, Psychologie oder Bildungs- und Sozialwissenschaften sind. In aller Regel war jede Testleiterin und jeder Testleiter für eine Testgruppe zuständig, wobei einer der Testleiter die Kommunikation mit der Schule übernahm. Hauptansprechperson an den Schulen war die Schulkoordinatorin beziehungsweise der Schulkoordinator, üblicherweise ein Mitglied des Lehrerkollegiums. Alle Absprachen zum zeitlichen Ablauf sowie zur Vorbereitung und Durchführung des PISA-Tests erfolgten seitens der Testleiterinnen und Testleiter mit der Schulkoordinatorin beziehungsweise dem Schulkoordinator. Am Testtag selbst trafen die Testleiterinnen und Testleiter zur vereinbarten Uhrzeit mit allen benötigten Materialien für die Testsitzung an der Schule ein. Wo nicht ausreichend Computer in der geforderten Ausstattung vorhanden waren, brachten die Testleiter portable Geräte mit, die vor Beginn der PISA-Testsitzung aufgebaut und gestartet wurden. Der Schulkoordinator hatte für jede Testgruppe eine Liste mit Namen und Ordnungsnummern vorbereitet, anhand derer jeder Schülerin und jedem Schüler ein Zugangscode zur Testsoftware zugewiesen werden konnte. Zur Wahrung der Datenschutzgrundsätze wurden die Namen nach erfolgter Zuweisung von den Ordnungsnummern getrennt. Nach der Vorbereitung des Testraumes durch die Testleiterin oder den Testleiter wurden die Schülerinnen und Schüler begrüßt und in die Testsitzung eingeführt (ca. 15 Minuten). Dann folgte die Bearbeitung der Testaufgaben, wobei die Schüler zunächst zwei Aufgabencluster (insgesamt 60 Minuten) dargeboten bekamen und nach einer kurzen Pause zwei weitere Cluster (insgesamt 60 Minuten). Es folgte eine Pause von ca. 15 Minuten, ehe der Schülerfragebogen zur Bearbeitung vorgelegt wurde. Dieser Teil dauerte etwa 35 Minuten. Abschließend wurde die Testsitzung beendet, die Daten gesichert und die Testleiterin beziehungsweise der Testleiter verstaute alle Materialien. Insgesamt belief sich der Zeitaufwand für die Schülerinnen und Schüler auf ca. dreieinhalb Stunden.

1.3.4 Sicherung der internationalen Vergleichbarkeit

Eine der größten Herausforderungen in international vergleichenden Schulleistungsstudien ist die Sicherstellung der Vergleichbarkeit und Fairness der Messinstrumente für unterschiedliche Teilnehmerstaaten. Die Vergleichbarkeit kann durch vielerlei Aspekte eingeschränkt sein, sodass entsprechende Maßnahmen ergriffen werden müssen, um

© Waxmann Verlag GmbH

die Ergebnisse möglichst gut international vergleichbar zu machen. Dementsprechend durchzieht der Anspruch der Vergleichbarkeit die Arbeit an PISA von der Vorbereitung der Testaufgaben über die Übersetzung in die jeweilige Unterrichtssprache bis hin zur Testdurchführung und Kodierung (Auswertung) der Lösungen. Eine systematische Bevor- oder Benachteiligung von Schülergruppen, etwa wegen einer besonders hohen oder geringen Vertrautheit mit Inhalten oder Formaten von Aufgaben, muss ausgeschlossen werden. Aus diesem Grund beginnt der Prozess der Aufgabenentwicklung und damit die Arbeit an einer PISA-Erhebungsrunde mit der Entwicklung einer theoretischen Rahmenkonzeption, in welcher für die Testbereiche relevante Anforderungen aus einer internationalen Perspektive festgelegt, begründet und strukturiert werden. In den folgenden Abschnitten werden wichtige Aspekte der Aufgabenentwicklung und des Testdesigns beschrieben, die gemeinsam zur Sicherstellung internationaler Vergleichbarkeit beitragen.

Aufgabenentwicklung

In jeder PISA-Erhebungsrunde werden neue Aufgaben nur für die Hauptdomäne entwickelt. Für die Kompetenzmessung in den beiden Nebendomänen kommen bereits früher eingesetzte Aufgaben zum Einsatz. In PISA 2015 wurden also für die Naturwissenschaften neue Aufgaben entwickelt, während für die Erfassung der Lesekompetenz und der mathematischen Kompetenz bereits eingesetzte Aufgaben verwendet wurden. Zu den bereits früher eingesetzten Aufgaben liegen sowohl Erfahrungen als auch psychometrische Kennwerte vor, die eine hohe Qualität der Aufgaben in Bezug auf ihre Fairness und damit die Vergleichbarkeit ihrer Ergebnisse garantieren. Auch im Schwerpunktbereich Naturwissenschaften kamen solche bewährten Aufgaben zum Einsatz, die bereits von früheren PISA-Kohorten bearbeitet worden waren und anhand derer sich Trends gut abbilden lassen. Der Aufgabenpool wurde durch zahlreiche neu entwickelte Aufgaben aufgefüllt, sodass alle Inhalte, Prozesse und Kontexte der theoretischen Rahmenkonzeption (OECD, 2016a) angemessen erfasst und umgesetzt werden konnten.

Testsicherheit und Datenschutz

Die Sicherheit und Geheimhaltung der Testaufgaben ist in so umfangreichen internationalen Bildungsvergleichsstudien wie PISA von größter Bedeutung. Die Aufgaben, die möglicherweise in späteren Erhebungsrunden nochmals eingesetzt werden sollen, dürfen nicht öffentlich bekannt werden, um die Ergebnisse nicht zu beeinflussen. Das wird durch eine Reihe von Maßnahmen sichergestellt.

In Deutschland wurden die Prozeduren der Datenerhebung und -verarbeitung in den 16 Bundesländern entsprechend den jeweiligen gesetzlichen Vorgaben gestaltet. Das nationale Projektmanagement am Zentrum für Internationale Bildungsvergleichsstudien (ZIB) an der TUM School of Education arbeitete hier eng mit den Datenschutzbeauftragten aller Länder zusammen. Zentral war die Wahrung der Anonymität der

© Waxmann Verlag GmbH

gesammelten Informationen aus Tests und Fragebögen, die durch aufwändige Verfahren sichergestellt wurde. Sowohl die Testhefte bzw. Testformen als auch die Fragebögen erhielten eine eindeutige Identifikationsnummer, die genau einer Schülerin oder einem Schüler zugeordnet werden konnte. Diese Identifikationsnummern wurden einmalig kurz vor Beginn der Testsitzung mit den Namen der Schülerinnen und Schüler zusammengeführt, damit jeder Schülerin und jedem Schüler die individuellen Zugangsdaten zum Einloggen am Computer und damit eine bestimmte Testform und ein Fragebogen zugewiesen werden konnten. Im Anschluss wurden die Identifikationsnummern wieder von den Namen der getesteten Fünfzehnjährigen entfernt. Die Namensliste hat zu keinem Zeitpunkt die Schule verlassen. Damit erfolgt die Verarbeitung der Daten vollständig pseudonymisiert, das heißt ausschließlich auf Basis der Identifikationsnummern und nicht von Namen oder Schulen. Da der Fragebogen für die Eltern der Schülerinnen und Schüler als einziges Erhebungsinstrument papierbasiert administriert wird, erhält dieser neben der internationalen, beim Sampling erzeugten Identifikationsnummer auch eine nationale Kennzahl. Anhand dieser nationalen Kennzahl können die Fragebögen lückenlos verfolgt werden, vom Druck über den Versand an die Testleiter, das Ausfüllen, die Rückgabe an die Schule beziehungsweise den Rückversand an das IEA *Data Processing and Research Center* (DPC), die dortige Logistik, Erfassung und Verarbeitung bis hin zum Archiv. Dort werden alle Unterlagen zehn Jahre gelagert beziehungsweise gespeichert und anschließend vernichtet oder gelöscht.

Antwortformate und Kodierung

Die in PISA 2015 eingesetzten Testaufgaben weisen drei unterschiedliche Antwortformate auf: Mehrfachwahl (dabei *Simple* oder *Complex Multiple Choice*) oder offenes Antwortformat. Etwa die Hälfte der Aufgaben gab den Schülerinnen und Schüler mehrere Antwortmöglichkeiten vor, aus denen sie eine oder mehrere auswählen mussten. Die andere Hälfte der Aufgaben erforderte eine selbst formulierte, mehr oder weniger ausführliche Antwort in offenem Format. Unterschieden wurden dabei Kurzantworten und numerische Antworten sowie Antworten, die aus mehreren Sätzen, einer Skizze oder einer Zeichnung bestehen konnten. Die Auswertung der Multiple-Choice-Antworten konnte aufgrund der computerbasierten Datenerfassung automatisiert und digital erfolgen, während die offenen Antworten von intensiv geschulten und hoch qualifizierten Kodiererinnen und Kodierern ausgewertet wurden. Für PISA 2015 stand hierfür erstmals eine eigens entwickelte Software zur Verfügung, das *Open Ended Coding System* (OECS, vgl. Kapitel 12). Die Fairness der Auswertung und die internationale Vergleichbarkeit wurden durch standardisierte, akribisch einzuhaltende Kodiervorschriften gestützt. Die Beherrschung dieser Vorschriften war Voraussetzung für die Zulassung zur Kodierung. Die Kodiererinnen und Kodierer mussten jeweils entscheiden, ob die Antworten der Schülerinnen und Schüler „vollständig richtig“, „teilweise richtig“ oder „falsch“ waren und dementsprechende Punktzahlen vergeben. Jede Antwort wurde dabei von mehreren Kodiererinnen und Kodierern bewertet, sodass ein Mehr-Augen-Prinzip

© Waxmann Verlag GmbH

die Qualität der Datenauswertung absichert. Der Großteil der Aufgaben war auch in PISA 2015 dichotom, das heißt, es wurden lediglich „richtige“ von „falschen“ Antworten unterschieden, und die Kategorie „teilweise richtig“ kam nur in Ausnahmefällen zum Einsatz. Der gesamte Kodiervorgang erfolgt international nach standardisierten Richtlinien, die zum Zweck der Qualitätssicherung in jedem Teilnehmerstaat kontrolliert werden. Dazu wird eine Stichprobe von Schülerantworten von vier unabhängigen Kodierern eingeschätzt und bewertet und die Übereinstimmung dieser Einschätzungen international überprüft. Für jede einzelne PISA-Aufgabe wurde vom internationalen Konsortium bei einer zu geringen Übereinstimmung der Kodierereinschätzungen mit den betreffenden Staaten Kontakt aufgenommen und geklärt, woher die zu großen Diskrepanzen kommen könnten. In Deutschland war dies bei keiner Aufgabe der Fall. Nach erfolgter Auswertung sowohl der automatisch als auch der von Menschen zu kodierenden Schülerantworten wurde ein Datensatz erstellt, der für alle bearbeiteten Testaufgaben die Angabe enthält, welche Schülerinnen und Schüler sie lösen konnten und welche nicht. Darüber hinaus enthält der Datensatz Angaben aus den Fragebögen sowie demografische Daten. Namen und Adressen sind im Sinne der Pseudonymisierung zu keinem Zeitpunkt Teil des Datensatzes.

Übersetzung und Anpassung der Aufgaben und Fragebögen

Nach abgeschlossener Aufgabenentwicklung wurden die Tests und Fragebögen in die jeweilige Unterrichtssprache der PISA-Teilnehmerstaaten übersetzt. Als Grundlage für die Übersetzung erhalten die nationalen Projektmanager zwei sogenannte Quellversionen der Testaufgaben, jeweils eine auf Englisch und eine auf Französisch. Eine der beiden Versionen wird in den Teilnehmerstaaten von zwei unabhängigen Übersetzungsteams in die jeweilige Testsprache übertragen und dann von einer dritten Übersetzerin oder einem Übersetzer unter Abgleich mit der zweiten Sprachversion in eine finale Version zusammengeführt. In Deutschland erfolgte die Übersetzung aufgrund der englischen Quellversion und die Zusammenführung unter Bezug auf die französische Quelle. Diese finale Version wurde in Deutschland von nationalen Expertinnen und Experten für Naturwissenschaften und ihre Didaktik begutachtet, wobei insbesondere auf inhaltliche Korrektheit, die Angemessenheit des sprachlichen Niveaus für die Zielgruppe fünfzehnjähriger Schülerinnen und Schüler sowie auf die Verwendung von Fachbegriffen gemäß dem schulischen Sprachgebrauch geachtet wurde. Da an PISA mehrere Staaten mit Unterrichtssprache Deutsch teilnehmen, müssen die Übersetzungen der Testaufgaben und Fragebögen ins Deutsche unter den deutschsprachigen Teilnehmerstaaten präzise abgestimmt werden. Die nationalen Projektmanager aus Österreich, der deutschsprachigen Schweiz, Luxemburg und Deutschland arbeiten hier eng zusammen. Deutschland übernahm dabei federführend die Übersetzung der neu entwickelten Naturwissenschaftsaufgaben. Die abgestimmten und von den nationalen Projektmanagern finalisierten Aufgaben und Fragebögen wurden im Anschluss von Übersetzungsexpertinnen und -experten auf ihre linguistische Äquivalenz mit den Quellversionen hin

© Waxmann Verlag GmbH

überprüft und schließlich für den Einsatz in PISA 2015 freigegeben. Damit die Schülerinnen und Schüler weltweit an derselben Stelle im Text in ihrer jeweiligen Sprache weiterklicken mussten und stets dieselben Informationen auf dem Bildschirm angezeigt bekamen, erfolgte die Übersetzung bereits im Format der später eingesetzten Testformen. Darin waren Abbildungen und Tabellen ebenso enthalten wie Texte und Umbrüche.

Durchführung der Datenerhebung und Qualitätskontrolle

Neben der genau vorgegebenen Ziehung einer repräsentativen Stichprobe, der Aufgabenentwicklung unter Beteiligung von Experten und nationalen Projektmanagern und der Übersetzung von Aufgaben und Fragebögen kommt auch der Standardisierung der Datenerhebung eine zentrale Rolle bei der Sicherung der internationalen Vergleichbarkeit zu. Die Kompetenz und Zuverlässigkeit der Testleiterinnen und Testleiter ist hier besonders wichtig. Die Rekrutierung und Schulung übernahm in Deutschland wie seit der ersten PISA-Erhebungsrunde im Jahr 2000 das IEA *Data Processing and Research Center* in Hamburg. Detaillierte und verbindliche Manuale beschreiben die Anforderungen der Testung und dienen den Testleiterinnen und Testleitern als Vorbereitung auf den PISA-Testtag. Darin enthalten sind die Aufgaben der Testleiterinnen und Testleiter, der Schulkoordinatorinnen und Schulkoordinatoren sowie Informationen zu bundeslandspezifischen Regelungen, etwa der Notwendigkeit einer Elterngenehmigung für den Schülerfragebogen. Durch den Testtag selbst führt ein sogenanntes Testleiterskript, das in der PISA-Testsitzung genau abgearbeitet wird. So wird gewährleistet, dass die PISA-Testsitzung in allen Schulen standardisiert abläuft und nicht durch die Testleiterin oder den Testleiter beeinflusst wird. Darüber hinaus wurde ein umfangreiches Qualitätsmonitoring durchgeführt. Dazu besuchten mehrere unabhängige, geschulte Beobachter einen Teil der PISA-Schulen unangekündigt, um sich zu vergewissern, dass die Testsitzung dem Skript entsprechend ablief und die Standardisierung nicht beeinträchtigt war. Es wurden keine Auffälligkeiten oder Abweichungen vom international vorgegebenen Vorgehen festgestellt.

1.4 Interpretation der Befunde

Befunde aus Large-Scale-Assessments wie PISA liefern empirische Hinweise auf Stärken und Schwächen von Bildungssystemen. Die in PISA gewonnenen Daten und deren Interpretation können als Bezugspunkt für die Begründung von Entscheidungen verwendet werden (Bieber, Martens, Niemann & Windzio, 2014; Bromme, Prenzel & Jäger, 2014). Dementsprechend liegt das Hauptaugenmerk dieses ersten Berichts zu PISA 2015 in Deutschland auf der Beschreibung bedeutsamer Bildungsergebnisse. Darüber hinaus ist es ein zentrales Anliegen von PISA, relevante Zusammenhänge zwischen Kompetenzmaßen und weiteren Merkmalen von Bildungssystemen und Schülerinnen und Schülern

© Waxmann Verlag GmbH

zu identifizieren und darzustellen. Merkmale, die aus theoretischer Sicht als Voraussetzungen oder Bedingungen für den Erwerb von Bildung oder die Entwicklung von Kompetenzen gelten können, werden in PISA erfasst und können vergleichend beschrieben werden. Solche Befunde über das Zusammenspiel verschiedener Merkmale mit Kompetenzen dienen als Diskussionsgrundlage und können im jeweiligen Kontext eines Bildungssystems verortet werden. Die Beschreibung solcher Zusammenhänge anhand der PISA-Daten erlaubt jedoch keine Aussagen darüber, wie die erfassten Merkmale kausal miteinander zusammenhängen. Vielmehr benötigt die Prüfung der kausalen Relevanz von Einflussfaktoren eine Reihe von Bedingungen wie etwa mehrere Messzeitpunkte, die systematische Variation von Untersuchungsbedingungen oder die Kontrolle weiterer Einflüsse. All dies ist durch das Design von PISA nicht gegeben und auch nicht angestrebt. Aussagen über Ursachen von Unterschieden – etwa zwischen Bildungssystemen – sind daher problematisch. Die in diesem Berichtsband vorgestellten statistischen Auswertungen beziehen sich folglich ausschließlich auf gegenseitige Beziehungen von Kompetenzen und bestimmten Hintergrundmerkmalen der untersuchten Fünfzehnjährigen. Gefundene Zusammenhänge zwischen hoher oder niedriger Kompetenz bei den Jugendlichen mit einzelnen ihrer Merkmale sind zwar eine notwendige, nicht aber eine hinreichende Bedingung für kausale Relevanz.

Mit PISA 2015 wurden von der OECD als Auftraggeberin der Studie mehrere Veränderungen gegenüber früheren Erhebungsrunden eingeführt. Die wichtigsten Neuerungen sind die Durchführung der Tests am Computer sowie ein verändertes statistisches Modell bei der Schätzung der Schülerkompetenzen. Beides bringt Folgen mit sich, die bei der Interpretation der Befunde berücksichtigt werden müssen, dies aber durch adäquate Methoden und Schlussfolgerungen auch können. Wenn mehrere Veränderungen zur gleichen Zeit vorgenommen werden, ist nicht auszuschließen, dass diese miteinander interagieren. Eine saubere Trennung, welche Veränderung schlussendlich welchen Effekt gehabt haben könnte, ist demnach kaum möglich. Die Ergebnisse aus PISA 2015 sind entsprechend – wie aber auch in früheren Erhebungsrunden – mit der üblichen Vorsicht und unter Berücksichtigung der Grenzen der Studie zu lesen. Genauso werden sie im vorliegenden Bericht präsentiert. In den einzelnen Kapiteln werden die Ergebnisse entsprechend dem aktuellen Stand der Bildungsforschung eingeordnet und plausible Erklärungsansätze für ihr Zustandekommen diskutiert.

1.5 Nationale und internationale Organisation der PISA-Studie

PISA ist das Ergebnis der Zusammenarbeit einer Vielzahl von Instituten und Organisationen. Als international vergleichende Schulleistungsstudie im Auftrag der OECD ist PISA auf zwei Ebenen organisiert: national sowie international. Auf internationaler Ebene obliegt die politische Steuerung dem sogenannten PISA *Governing Board* (PGB), in welchem alle Teilnehmerstaaten mit mindestens einer Vertreterin oder einem Vertreter repräsentiert sind. Für Deutschland sind dies in PISA 2015 Ministerialdirigentin

© Waxmann Verlag GmbH

Elfriede Ohrnberger (Bayerisches Staatsministerium für Unterricht und Kultus), Dr. Annemarie Klemm (Freie Hansestadt Bremen, Referat 20 Qualitätsentwicklung und Standardsicherung) sowie Dr. Susanne von Below (Bundesministerium für Bildung und Forschung), die zum 1. Januar 2015 von Katharina Koufen (Bundesministerium für Bildung und Forschung) abgelöst wurde. Zum 1. April 2016 übernahm Dr. Martina Diedrich (Institut für Bildungsmonitoring und Qualitätsentwicklung Hamburg) das Mandat von Dr. Annemarie Klemm. Die Koordination auf internationaler Ebene obliegt der OECD in Paris, verantwortlich ist dort Andreas Schleicher. Ein internationales Konsortium aus mehreren Institutionen wurde mit der Durchführung von PISA 2015 beauftragt. Federführend dabei war *Educational Testing Service*, ETS, in Princeton (USA) und dort Claudia Tamassia. In diesem Konsortium kooperierten die Institute ACER (Australien), *Department of Experimental and Theoretical Pedagogy* an der Universität von Liège (SPe, Belgien), cApStAn (Belgien), *Center for Educational Technology* (CET, Israel), CRP Henri Tudor (Luxembourg), DIPF (Deutschland), EMACS (Luxembourg), ETS (USA), GESIS (Deutschland), Pearson (Großbritannien), *Statistics Canada* und WESTAT (USA). Die theoretischen Rahmenkonzeptionen wurden von international bestellten Expertengruppen zu den in PISA untersuchten Fachdomänen (Naturwissenschaften, Mathematik und Lesen) sowie zur Entwicklung der Fragebögen und zur Sicherung der technischen Standards erarbeitet.

Auf nationaler Ebene wurden in allen 72 PISA-Teilnehmerstaaten sogenannte *National Project Managers* (NPM) damit beauftragt, PISA 2015 sachgemäß vorzubereiten und durchzuführen. In Deutschland wurde das nationale Projektmanagement nach PISA 2012 zum zweiten Mal an das Zentrum für Internationale Bildungsvergleichsstudien (ZIB) vergeben. Das ZIB wurde im Jahr 2010 vom Bundesministerium für Bildung und Forschung (BMBF) und der Ständigen Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland (KMK) als An-Institut der Technischen Universität München (TUM) eingerichtet. Als gemeinnütziger Verein bündelt das ZIB die Expertise dreier Einrichtungen im Bereich empirischer Bildungsforschung, nämlich der Technischen Universität München (TUM), des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF) in Frankfurt am Main sowie des Leibniz Instituts für die Pädagogik der Naturwissenschaften und Mathematik (IPN) in Kiel. Das ZIB war in verschiedenen Rollen an allen Phasen des Projekts PISA 2015 aktiv beteiligt. Das nationale Projektmanagement für PISA 2015 ist an der School of Education der TUM angesiedelt und wird von Prof. Dr. Kristina Reiss (Vorstandsvorsitzende des ZIB und Dekanin der TUM School of Education, Nachfolgerin von Prof. Dr. Manfred Prenzel, der bis Juli 2014 die Rolle des nationalen Projektmanagers innehatte) und PD Dr. Christine Sälzer geführt. Die Entwicklung des Designs und der Inhalte für die Kontextfragebögen (vgl. Kapitel 13 in diesem Band) lag in der Federführung des DIPF in Frankfurt unter der Leitung von Prof. Dr. Eckhard Klieme. Das IPN in Kiel brachte unter der Leitung von Prof. Dr. Olaf Köller seine fachliche Expertise bei der Begutachtung und Optimierung der neu entwickelten Aufgaben zur naturwissenschaftlichen Kompetenz ein. Das ZIB stimmt seine Arbeit mit der Amtschefskommission „Qualitätssicherung in Schulen“

© Waxmann Verlag GmbH

ab (Vorsitz: Ministerialdirektor Herbert Püls, Bayern, sowie Staatsrat Dr. Michael Voges, Hamburg).

Der Arbeitsgruppe PISA 2015 an der TUM School of Education gehörten und gehören PD Dr. Christine Sälzer, Dr. Anja Schiepe-Tiska, Dr. Mirjam Weis, Dr. Fabian Zehner, Elisabeth González Rodríguez, Jörg-Henrik Heine, Katharina Kiemer, Julia Mang, Stefanie Schmidtner, Inga Simm und zahlreiche studentische Hilfskräfte an. Ihre Expertise zur Hauptdomäne Naturwissenschaften brachten Prof. Dr. Marcus Hammann (Universität Münster), Dr. Silke Rönnebeck und Dr. Katrin Schöps (beide IPN Kiel) ein. Darüber hinaus besteht eine enge Zusammenarbeit der PISA-Arbeitsgruppe mit dem IQB in Berlin (Prof. Dr. Petra Stanat, Dr. Stefan Schipolowski), dem IPN in Kiel (Prof. Dr. Oliver Lüdtke, Prof. Dr. Gabriel Nagy, Prof. Dr. Aiso Heinze, Prof. Dr. Knut Neumann, Prof. Dr. Ilka Parchmann, Dr. Alexander Robitzsch) und dem DIPF in Frankfurt (Prof. Dr. Frank Goldhammer, Dr. Nina Jude).

Die Vorbereitung und Durchführung der Datenerhebung und -erfassung lag wie in allen früheren PISA-Erhebungsrunden beim IEA DPC in Hamburg. Die Projektkoordination hatte Heiko Sibberns. Zuständig waren darüber hinaus Anja Waschk, Tobias H. Haring, Viktoria Böhm (Feldarbeit), Dr. Lars Borchert, Jens Gomolka (internationales Datenmanagement), Guido Martin (Kodierung), Wolfram Jarchow (nationales Datenmanagement) und Dr. Miriam Hellrung (Datenschutz).

Mit PISA 2015 konnte bereits die sechste Erhebungsrunde erfolgreich abgeschlossen werden, was in der Hauptsache dem großen Engagement aller beteiligten Schulen zu verdanken ist. Die Bereitschaft der Schulen, diese Vergleichsstudie durch ihre Teilnahme zu unterstützen und aus den Ergebnissen zu lernen, ist für das Gelingen eines so großen Projektes wie PISA unerlässlich. Das nationale Projektmanagement dankt allen Schülerinnen und Schülern, ihren Eltern und Lehrkräften sowie ihren Schulleitungen sehr herzlich für ihre Mitarbeit!

Literatur

Autorengruppe Bildungsberichterstattung. (2016). *Bildung 2016 in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: Bertelsmann.

Bieber, T., Martens, K., Niemann, D. & Windzio, M. (2014). Grenzenlose Bildungspolitik? Empirische Evidenz für PISA als weltweites Leitbild für nationale Bildungsreformen. *Zeitschrift für Erziehungswissenschaft, 17* (S4), 141–166. http://dx.doi.org/10.1007/s11618-014-0513-6

Bromme, R., Prenzel, M. & Jäger, M. (2014). Empirische Bildungsforschung und evidenzbasierte Bildungspolitik. Eine Analyse von Anforderungen an die Darstellung, Interpretation und Rezeption empirischer Befunde. *Zeitschrift für Erziehungswissenschaft, 17* (Sonderheft 27), 3–54.

Drechsel, B. & Prenzel, M. (2008). *Aus Vergleichsstudien lernen. Aufbau, Durchführung und Interpretation internationaler Vergleichsstudien*. München: Oldenbourg.

© Waxmann Verlag GmbH

Drechsel, B., Prenzel, M. & Seidel, T. (2014). Nationale und internationale Schulleistungsstudien. In E. Wild (Hrsg.), *Pädagogische Psychologie* (2. vollst. überarb. und aktual. Aufl., S. 343–368). Berlin: Springer.

Gräsel, C. (2011). Was ist empirische Bildungsforschung? In H. Reinders, H. Ditton, C. Gräsel & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung* (S. 13–27). Wiesbaden: VS-Verl.

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.

OECD. (2009). *PISA 2009 assessment framework: Key competencies in reading, mathematics and science*. Paris: OECD.

OECD. (2012). *PISA 2009 technical report*. Paris: OECD.

OECD. (2013). *PISA 2012 Assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. Paris: OECD.

OECD. (2014). *PISA 2012 technical report*. Paris: OECD.

OECD. (2016a). *PISA 2015 Assessment and analytical framework. Science, reading, mathematic and financial literacy*. Paris: OECD.

OECD. (2016b). *PISA 2015 results (Volume I): Excellence and equity in education.* Paris: OECD.

Prenzel, M. (2005). Zur Situation der empirischen Bildungsforschung. In H. Mandl & B. Kopp (Hrsg.), *Impulse für die Bildungsforschung: Stand und Perspektiven/Dokumentation eines Expertengesprächs* (S. 7–21). Berlin: De Gruyter.

Prenzel, M. & Lankes, E.-M. (2013). Was können Schülerinnen und Schüler über ihren Unterricht sagen? Ein Blick in die Schülerfragebogen von internationalen Vergleichsstudien. In N. McElvany & H. G. Holtappels (Hrsg.), *Empirische Bildungsforschung. Theorien, Methoden, Befunde und Perspektiven* (S. 93–107). Münster: Waxmann.

Roeder, P. M. (2003). TIMSS und PISA – Chancen eines neuen Anfangs in Bildungspolitik, -planung, -verwaltung und Unterricht. Endlich ein Schock mit Folgen? *Zeitschrift für Pädagogik, 49* (2), 180–197.

Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion* (2. Aufl.). Berlin: Hans Huber.

Rutkowski, D. J. & Prusinski, E. L. (2011). The limits and possibilities of international large-scale assessments. *Education Policy Briefs, 9* (2), 1–4.

Sälzer, C. (2016). *Studienbuch Schulleistungsstudien. Das Rasch-Modell in der Praxis*. Berlin: Springer.

Sälzer, C., Prenzel, M. & Klieme, E. (2013). Schulische Rahmenbedingungen der Kompetenzentwicklung. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 155–187). Münster: Waxmann.

Seidel, T. & Prenzel, M. (2008). Large scale assessment. In J. Hartig, E. Klieme & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 279–304). Göttingen: Hogrefe & Huber.

Stanat, P., Böhme, K., Schipolowski, S. & Haag, N. (Hrsg.). (2016). *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.

Statistisches Bundesamt. (2015). *Allgemeinbildende Schulen, Fachserie 11, Reihe 1 – Schuljahr 2014/15.* Zugriff am 20.07.2016. Verfügbar unter https://www.destatis.de/DE/Pu-

© Waxmann Verlag GmbH

blikationen/Thematisch/BildungForschungKultur/Schulen/AllgemeinbildendeSchulen2110100157004.pdf?__blob=publicationFile

Tenorth, H.-E. (2004). Bildungsstandards und Kerncurriculum. Systematischer Kontext, bildungstheoretische Probleme. *Zeitschrift für Pädagogik, 50* (5), 650–661.

Tenorth, H.-E. (2005). „Grundbildung", „Basiskompetenzen" und allgemeine Bildung. In A. Kauf, K. Liebers & A. Prengel (Hrsg.), *Länderübergreifende Curricula für die Grundschule* (S. 93–107). Bad Heilbrunn: Klinkhardt.

© Waxmann Verlag GmbH

2 Naturwissenschaftliche Kompetenz in PISA 2015 – Ergebnisse des internationalen Vergleichs mit einem modifizierten Testansatz

Anja Schiepe-Tiska, Silke Rönnebeck, Katrin Schöps, Knut Neumann, Stefanie Schmidtner, Ilka Parchmann & Manfred Prenzel

In Deutschland erreichen die Fünfzehnjährigen in PISA 2015 in den Naturwissenschaften 509 Punkte und liegen damit erneut signifikant (nämlich 16 Punkte) über dem OECD-Durchschnitt. Deutschland gehört damit zu einer Gruppe von Staaten, die auf einem ähnlichen Niveau den noch einmal signifikant leistungsstärkeren Spitzenstaaten Japan, Estland, Finnland und Kanada folgen. Die drei Teilkompetenzen *Phänomene naturwissenschaftlich erklären*, *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* sowie *Daten und Evidenz naturwissenschaftlich interpretieren* sind in Deutschland gleichmäßig hoch ausgeprägt. Eine leichte Stärke zeichnet sich für die Teilkompetenz *Phänomene naturwissenschaftlich erklären* ab. Im Vergleich zu früheren PISA-Erhebungsrunden fällt auf, dass erstmals sowohl im OECD-Durchschnitt als auch in Deutschland die Jungen eine signifikant höhere naturwissenschaftliche Kompetenz erzielen als die Mädchen. Eine differenzierte Betrachtung der naturwissenschaftlichen Kompetenz nach Schulart zeigt, dass Jugendliche am Gymnasium eine relative Stärke in der Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren* haben. Im Vergleich zur PISA-Erhebung 2006, als die Naturwissenschaften das letzte Mal als Schwerpunktdomäne untersucht wurden, zeichnet sich für Deutschland keine signifikante Veränderung ab. Deutschlands Rangposition innerhalb der OECD-Staaten bleibt im Wesentlichen unverändert. Das heißt, es ist in Deutschland gelungen, die naturwissenschaftliche Kompetenz auf einem im internationalen Vergleich hohen Niveau zu stabilisieren. Einige Befunde lassen Möglichkeiten einer weiteren Leistungssteigerung erkennen. So ist zum Beispiel an den Gymnasien im Vergleich zu PISA 2006 ein statistisch bedeutsamer Rückgang der naturwissenschaftlichen Kompetenz zu beobachten. Für die nicht gymnasialen Schularten zeigt sich diese Entwicklung nicht. Deutschland schöpft das Potenzial des differenzierten Schulsystems zur besseren Förderung von Talenten in der naturwissenschaftlichen Bildung also weiterhin nicht hinreichend aus.

© Waxmann Verlag GmbH

In der heutigen globalisierten und digitalisierten Welt ist das gesellschaftliche Leben zunehmend durch Naturwissenschaften und Technik geprägt. Damit kommt naturwissenschaftlicher Bildung für die Teilhabe an der Gesellschaft eine herausragende Bedeutung zu (vgl. Bybee & Fuchs, 2006). In der Folge besteht das übergeordnete Ziel des Unterrichts in den Naturwissenschaften heute nicht alleine in der Ausbildung zukünftiger Naturwissenschaftlerinnen und Naturwissenschaftler, sondern richtet sich gleichermaßen auf die naturwissenschaftliche Bildung aller Schülerinnen und Schüler (Roberts, 2007).

Fortschritte in Naturwissenschaften und Technik konfrontieren bereits Schülerinnen und Schüler mit immer neuen Herausforderungen. Diese können das alltägliche individuelle Leben betreffen, aber auch bestimmte Regionen oder die Welt als Ganzes. Naturwissenschaftliche Bildung soll Schülerinnen und Schüler auf die Begegnung mit diesen Herausforderungen vorbereiten (Millar, 2006). Sie soll es ihnen ermöglichen, naturwissenschaftliche Phänomene und technische Prozesse in ihrem Alltag zu erkennen, zu verstehen und zu erklären. Da ständig neue Erkenntnisse gewonnen werden und immer neue Technologien Einzug in den Alltag halten, müssen Jugendliche insbesondere lernen, sich neue Informationen zu erschließen und in Wissen umzusetzen. Diese Informationen können aus Quellen im Sinne naturwissenschaftlicher oder technischer Publikationen, aber auch aus eigenen kleinen naturwissenschaftlichen Untersuchungen stammen. Nicht zuletzt muss naturwissenschaftliche Bildung Schülerinnen und Schüler befähigen, die gewonnenen Informationen kritisch zu bewerten und daraus Schlussfolgerungen für das eigene Handeln zu ziehen (OECD, 2016a; vgl. auch NRC, 2012). Die Jugendlichen müssen darüber hinaus auch in der Lage sein, im Bereich der Naturwissenschaften mit ihren Wissenslücken umzugehen und lernen, sich neues Wissen eigenständig anzueignen. Dies setzt neben einem fundierten Inhaltswissen auch Prozess- und Überblickswissen voraus (vgl. Shavelson, Ruiz-Primo & Wiley, 2005) und fordert ein gehöriges Maß an Motivation. Naturwissenschaftliche Bildung, die im schulischen Kontext Jugendliche im beschriebenen Sinne auf die Begegnung mit aktuellen und zukünftigen Herausforderungen vorbereitet, wird auch als naturwissenschaftliche Grundbildung, *Scientific Literacy*, bezeichnet (vgl. deBoer, 2000).

Inwieweit es einem Staat gelingt, Schülerinnen und Schülern eine solche naturwissenschaftliche Grundbildung in der Schule zu vermitteln, ist ein wesentlicher Faktor für dessen Wohlstand und Wohlergehen (deBoer, 2011). Auch aus dieser Perspektive haben Large-Scale-Assessments wie PISA, die auf einen internationalen Vergleich des Leistungsstands der Schülerinnen und Schüler im Hinblick auf eine naturwissenschaftliche Grundbildung zielen, in den letzten zwei Jahrzehnten zunehmend an Aufmerksamkeit gewonnen. Für Deutschland zeigte sich in den ersten Vergleichsstudien, dass die Leistungen der Schülerinnen und Schüler im Bereich der Naturwissenschaften nicht den Erwartungen entsprachen (Beaton et al., 1996; Prenzel, Rost, Senkbeil, Häußler & Klopp, 2001). Bereits nach TIMSS (*Third International Mathematics and Science Study*) wurde als Grund für die Leistungsschwächen eine mangelnde Kohärenz des naturwissenschaftlichen Unterrichts ausgemacht, die entsprechend kumulatives Lernen eher verhindere als unterstütze (Baumert et al., 1997). Im Anschluss an TIMSS und die erste

© Waxmann Verlag GmbH

PISA-Erhebungsrunde im Jahr 2000, in der die Ergebnisse in den Naturwissenschaften für Deutschland unter dem OECD-Durchschnitt lagen, wurden mehrere Initiativen und Maßnahmen zur Verbesserung der naturwissenschaftlichen Bildung ergriffen. Beispiele sind SINUS, das Programm zur Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts (Prenzel, Friedrich & Stadler, 2009), oder die sogenannten Kontext-Projekte in der Biologie, Chemie und Physik (Bayrhuber et al., 2007; Demuth, Gräsel, Parchmann & Ralle, 2008; Mikelskis-Seifert & Duit, 2010). Als eine Konsequenz aus den Befunden der ersten Vergleichsstudien wurden außerdem Bildungsstandards in den naturwissenschaftlichen Fächern (für den Mittleren Schulabschluss) eingeführt, die in Deutschland ebenfalls eine naturwissenschaftliche Grundbildung als übergeordnetes Ziel des Unterrichts in den Naturwissenschaften formulieren (KMK, 2005a, 2005b, 2005c).

In den nachfolgenden Erhebungsrunden der PISA-Studie verbesserten sich die Leistungen der Schülerinnen und Schüler in Deutschland im Bereich der Naturwissenschaften deutlich. Seit 2006 liegen sie signifikant über dem OECD-Durchschnitt (Rönnebeck, Schöps, Prenzel, Mildner & Hochweber, 2010; Schiepe-Tiska, Schöps, Rönnebeck, Köller & Prenzel, 2013). Auch in den letzten Jahren gab es zahlreiche Initiativen und Maßnahmen, um den MINT-Bereich (Mathematik, Informatik, Naturwissenschaften, Technik) stärker zu fördern. Dazu gehören die Einrichtung von Schülerlaboren oder die Durchführung (regionaler) naturwissenschaftlicher Wettbewerbe genauso wie die Gründung von Netzwerken wie MINT-EC („Nationales Excellence-Schulnetzwerk“) oder das „Nationale MINT Forum“.

In der aktuellen PISA-Erhebungsrunde 2015 stehen die Naturwissenschaften zum zweiten Mal nach PISA 2006 im Mittelpunkt der Betrachtung. Dies wurde von der OECD zum Anlass genommen, die Rahmenkonzeption und den Test zur Erfassung der naturwissenschaftlichen Kompetenzen weiterzuentwickeln. Durch diese Anpassungen soll sichergestellt werden, dass die Rahmenkonzeption und der Test auch in Anbetracht der ständig voranschreitenden Entwicklungen in Naturwissenschaft und Technik ihre Aktualität und Relevanz behalten. Im Zuge der Bemühungen um eine Weiterentwicklung von PISA wurde auch der Testmodus verändert: Erstmals erfolgte die gesamte Erhebung am Computer (vgl. Kapitel 12). Mit der Umstellung auf computerbasiertes Testen kamen zudem neue Aufgabenformate mit interaktiven Elementen zum Einsatz. Mit diesen den technischen Entwicklungen folgenden Modifikationen nimmt die OECD allerdings in Kauf, dass die Vergleichbarkeit mit den Ergebnissen vorheriger PISA-Erhebungsrunden nur eingeschränkt möglich ist und Veränderungen gegenüber früheren Befunden entsprechend vorsichtig interpretiert werden müssen.

Das vorliegende Kapitel konzentriert sich zunächst auf den Bericht der Ergebnisse der neuesten Erhebungsrunde der PISA-Studie aus dem Jahr 2015. Dazu wird im folgenden Abschnitt die überarbeitete Rahmenkonzeption naturwissenschaftlicher Grundbildung beschrieben und der weiterentwickelte Test mit den neuen Aufgabenformaten beispielhaft erläutert. Anschließend wird der Leistungsstand der Jugendlichen in Deutschland im internationalen Vergleich berichtet. Dabei wird insbesondere auf die Frage eingegangen, inwieweit und in welchen Bereichen Schülerinnen und Schüler in

© Waxmann Verlag GmbH

Deutschland bereits sehr gute Leistungen erbringen und in welchen Bereichen eine verstärkte Förderung nötig erscheint. Im Anschluss daran werden Ergebnisse vertiefender Analysen in Bezug auf das Gymnasium und nicht gymnasiale Schularten vorgestellt. Ferner werden – unter Berücksichtigung des modifizierten Testansatzes und insbesondere der damit eingeschränkt möglichen Vergleichbarkeit – Befunde zur Veränderung der naturwissenschaftlichen Kompetenz im Vergleich zu PISA 2006 erörtert.

2.1 Die Rahmenkonzeption zur Erhebung der naturwissenschaftlichen Kompetenz in PISA 2015

PISA untersucht, in welchem Maße Jugendliche am Ende ihrer Pflichtschulzeit über naturwissenschaftliche Grundbildung (*Scientific Literacy*) verfügen (OECD, 2016a). Dabei geht es zum einen darum, dass alle Schülerinnen und Schüler die Bedeutung von Naturwissenschaften und Technik für ihr alltägliches Leben verstehen. Sie sollen in der Lage sein, naturwissenschaftliche Daten und Informationen mithilfe wissenschaftlicher Herangehensweisen bewerten zu können, um daraus folgend evidenzbasiert Entscheidungen zu treffen. Zum anderen geht es bei PISA aber auch darum, besonders stark ausgeprägte Leistungen identifizieren zu können, die gute Voraussetzungen für eine Ausbildung oder ein Studium in MINT-Berufen bieten.

Bei PISA stehen drei Fragen im Mittelpunkt: Erstens, was wissen und können Schülerinnen und Schüler, wenn sie sich mit naturwissenschaftlichen und technischen Fragen und Problemen auseinandersetzen? Sind sie zweitens in der Lage, dieses Wissen flexibel in unterschiedlichen Situationen anzuwenden? Welche Bedeutung schreiben sie drittens Naturwissenschaften und Technik zu (OECD, 2016a)?

Naturwissenschaftliche Grundbildung bezieht sich im Sinne dieser Konzeption nicht nur auf eine Reproduktion von naturwissenschaftlichem Wissen, das in der Schule vermittelt wird. Wichtig ist vielmehr, inwieweit die Schülerinnen und Schüler dieses Wissen in unterschiedlichen, alltagsnahen Kontexten sinnvoll und problemlösend anwenden können (Bybee & McCrae, 2011). Die so gefasste Grundbildung dient als Fundament für lebenslanges Lernen und befähigt Jugendliche zur aktiven Teilhabe an gesellschaftlichen Prozessen und zu einer Mitgestaltung ihrer Lebensumwelt (Prenzel et al., 2001). Das Konzept der naturwissenschaftlichen Kompetenz schließt explizit ein naturwissenschaftsbasiertes Verständnis von Technik mit ein (OECD, 2016a; vgl. auch Prenzel, Reiss & Hasselhorn, 2009).

In PISA 2006, als die Naturwissenschaften zum ersten Mal als Hauptdomäne im Mittelpunkt der Studie standen, wurde naturwissenschaftliche Grundbildung als die Fähigkeit einer Person beschrieben,

- naturwissenschaftliches Wissen anzuwenden, um Fragestellungen zu erkennen, sich neues Wissen anzueignen, naturwissenschaftliche Phänomene zu beschreiben und zu erklären sowie aus Belegen Schlussfolgerungen zu ziehen,

© Waxmann Verlag GmbH

- die charakteristischen Eigenschaften der Naturwissenschaften als eine Form menschlichen Wissens und Forschens zu verstehen,
- zu erkennen und sich darüber bewusst zu sein, wie Naturwissenschaften und Technik unsere materielle, intellektuelle und kulturelle Umwelt formen,
- sowie die Bereitschaft, sich mit naturwissenschaftlichen Ideen und Themen zu beschäftigen und sich reflektierend mit ihnen auseinanderzusetzen (OECD, 2006; Prenzel et al., 2007).

Diese Beschreibung naturwissenschaftlicher Grundbildung wurde bei PISA 2015 aufgegriffen. Entsprechend wurden sowohl die Rahmenkonzeption als auch der Naturwissenschaftstest von 2006 weiterentwickelt und mit Blick auf das neue computerbasierte Erhebungsverfahren aktualisiert. Wie auch in der Konzeption 2006 unterscheidet das theoretische Kompetenzmodell vier Elemente, die miteinander in Beziehung stehen: Kontexte, Teilkompetenzen, Wissensbereiche sowie motivationale Orientierungen und Einstellungen (Abbildung 2.1). Im Folgenden werden die einzelnen Elemente der Rahmenkonzeption sowie deren Weiterentwicklungen genauer beschrieben.

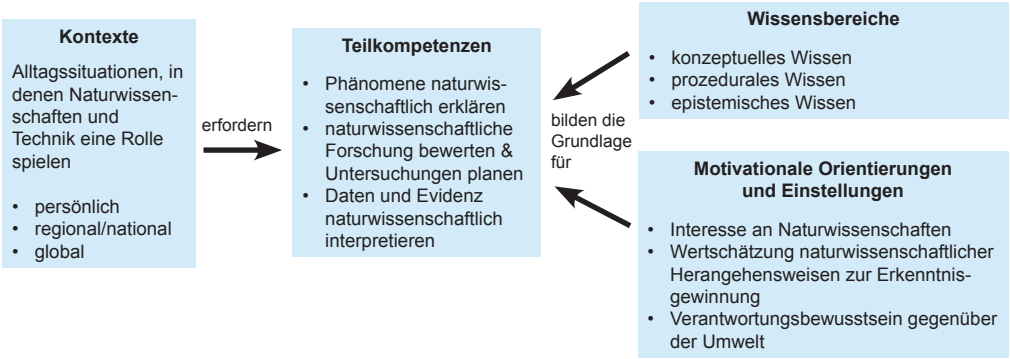


Abbildung 2.1: Die PISA-2015-Rahmenkonzeption naturwissenschaftlicher Grundbildung

2.1.1 Naturwissenschaftliche Kontexte

Wenn Schülerinnen und Schüler in unterschiedlichen Situationen (Kontexten) einer naturwissenschaftlichen Frage- oder Problemstellung begegnen, liegt ein entscheidender Aspekt naturwissenschaftlicher Grundbildung darin, dass sie nicht nur Auswendiggelerntes reproduzieren, sondern ihr Wissen flexibel und zweckmäßig anwenden können. Naturwissenschaftliche Kompetenz soll Jugendliche in die Lage versetzen, naturwissenschaftliche Fragestellungen und Probleme, denen sie im Alltag begegnen, aus unterschiedlichen naturwissenschaftlichen Perspektiven betrachten und bearbeiten zu können. Das kann – je nach Kompetenz der Jugendlichen – besser oder weniger gut gelingen. Damit die Aufgaben für Jugendliche bedeutungsvoll sind, werden sie bei PISA

© Waxmann Verlag GmbH

in Kontexte eingebettet, die relativ alltagsbezogen und den Jugendlichen einigermaßen vertraut sind. Dabei werden aus Gründen der Testfairness thematisch möglichst viele unterschiedliche Erfahrungsmöglichkeiten und Interessen berücksichtigt.

Die Kontexte werden formal danach unterschieden, ob sie sich auf die Jugendlichen selbst, ihre Familie oder Freunde beziehen (*persönlicher Kontext*), die sie umgebende Gesellschaft betreffen (*regionaler/nationaler Kontext*) oder von weltweiter Relevanz sind (*globaler Kontext*). Dabei werden (wie bei PISA 2006) jeweils fünf relativ breite Anwendungsbereiche differenziert: Gesundheit und Krankheit, natürliche Ressourcen, Umweltqualität, Risiken/Gefahren sowie Grenzen von Naturwissenschaften und Technik. Einige Anwendungsbereiche werden auch in einen historischen Kontext eingebettet, um zu überprüfen, inwieweit Schülerinnen und Schüler in der Lage sind, die Prozesse und Methoden zur Weiterentwicklung naturwissenschaftlichen Wissens zu verstehen. Bei der Auswahl der Kontexte sind außerdem kulturelle Unterschiede in den teilnehmenden Staaten zu berücksichtigen, da diese die Relevanz der Aufgaben für die Jugendlichen beeinflussen können. Tabelle 2.1 zeigt eine Übersicht der Kontexte mit konkreten Beispielen, die weitestgehend denen von PISA 2006 entsprechen (Prenzel et al., 2007).

Tabelle 2.1: Naturwissenschaftliche Kontexte bei PISA 2015

	Persönlich (Selbst, Familie und Freunde)	**Regional/National** (Gesellschaft)	**Global** (Leben weltweit)
Gesundheit und Krankheit	Erhaltung der Gesundheit, Unfälle, Ernährung	Kontrolle von Krankheiten, Wahl von Nahrungsmitteln, Gesundheitswesen	Epidemien, Ausbreitung infektiöser Krankheiten
Natürliche Ressourcen	persönlicher Verbrauch von Materialien und Energie	Erhaltung der menschlichen Bevölkerung, Lebensqualität, Sicherheit, Produktion und Verteilung von Lebensmitteln, Energieversorgung	erneuerbare und nicht erneuerbare natürliche Systeme, Bevölkerungswachstum
Umwelt	umweltfreundliches Verhalten, Benutzung und Entsorgung von Materialien	Bevölkerungsverteilung, Entsorgung von Abfällen, Umweltverträglichkeit, lokales Wetter	Artenvielfalt, Umweltverträglichkeit, Kontrolle der Umweltverschmutzung,Gewinnung und Verlust von Land
Risiken/ Gefahren	Einschätzen von Risiken bestimmter Lebensstile	schnelle Veränderungen (z. B. Erdbeben, Unwetter), langsam fortschreitende Veränderungen (z. B. Erosion, Sedimentation), Risikoeinschätzung	Klimaveränderung, Auswirkungen moderner Kommunikation
Grenzen von Naturwissenschaft und Technik	naturwissenschaftliche Aspekte von Hobbys, privat genutzter Technologie, Musik und Sport	neue Materialien, Geräte und Prozesse, Genmanipulation, Gesundheitstechnologie, Transport	Aussterben von Arten, Erforschung des Weltalls und des Ursprungs und der Struktur des Universums

© Waxmann Verlag GmbH

2.1.2 Naturwissenschaftliche Teilkompetenzen

Um erfolgreich mit naturwissenschaftlichen Fragestellungen umgehen und naturwissenschaftliche Probleme in alltagsnahen Situationen lösen zu können, müssen Schülerinnen und Schüler über bestimmte Kompetenzen verfügen. Bei PISA 2015 werden drei Teilkompetenzen naturwissenschaftlicher Grundbildung unterschieden. Diese wurden seit PISA 2006 wie folgt überarbeitet und weiterentwickelt (vgl. OECD, 2016a):

(1) *Phänomene naturwissenschaftlich erklären*,
(2) *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* sowie
(3) *Daten und Evidenz naturwissenschaftlich interpretieren*.

Diese Teilkompetenzen repräsentieren wichtige Handlungs- und Leistungspotenziale der Jugendlichen. *Phänomene naturwissenschaftlich erklären* meint, dass sie in der Lage sind,

Tabelle 2.2: Teilkompetenzen der naturwissenschaftlichen Grundbildung

Phänomene naturwissenschaftlich erklären
Erklärungen für naturwissenschaftliche und technische Phänomene erkennen, entwickeln und bewerten unter der Nutzung folgender Fähigkeiten: • Vorhandenes Wissen in den Naturwissenschaften in einer gegebenen Situation anwenden • Erkennen, Erstellen und Nutzen von Erklärungsmodellen und Repräsentationen • Angemessene Vorhersagen treffen und begründen • Erklärende Hypothesen bereitstellen • Mögliche Konsequenzen naturwissenschaftlichen Wissens für die Gesellschaft erklären
Naturwissenschaftliche Forschung bewerten und Untersuchungen planen
Naturwissenschaftliche Untersuchungen beschreiben und bewerten sowie Möglichkeiten zur Untersuchung naturwissenschaftlicher Fragestellungen vorschlagen unter der Nutzung folgender Fähigkeiten: • Fragestellungen erkennen, die in einer naturwissenschaftlichen Studie untersucht wurden • Fragestellungen erkennen, die mit naturwissenschaftlichen Methoden untersucht werden können • Möglichkeiten zur naturwissenschaftlichen Untersuchung einer Fragestellung vorschlagen und bewerten • Beschreiben und bewerten, wie in der Wissenschaft die Reliabilität von Daten sowie die Objektivität und Generalisierbarkeit von Ergebnissen gewährleistet wird
Daten und Evidenz naturwissenschaftlich interpretieren
Naturwissenschaftliche Daten, Behauptungen und Argumente in verschiedenen Repräsentationsformen analysieren, bewerten und daraus angemessene Schlussfolgerungen ziehen unter der Nutzung folgender Fähigkeiten: • Datentransformation von einer Repräsentation in eine andere • Daten analysieren und interpretieren sowie geeignete Schlussfolgerungen daraus ziehen • Annahmen, Evidenz und Argumentationen in naturwissenschaftlichen Texten erkennen • Argumente, die auf naturwissenschaftlicher Evidenz und naturwissenschaftlichen Theorien basieren, von solchen unterscheiden, die auf anderen Grundlagen beruhen • Naturwissenschaftliche Argumente und Evidenz verschiedener Quellen (z.B. Zeitungen, Zeitschriften, Internet) bewerten

© Waxmann Verlag GmbH

naturwissenschaftliche Erklärungen für natürliche und technische Phänomene zu erkennen, zu entwickeln und zu beurteilen. *Naturwissenschaftliche Forschung bewerten und Untersuchungen planen* bezieht sich auf die Fähigkeit, naturwissenschaftliche Fragestellungen zu beschreiben und zu bewerten sowie Vorschläge zu unterbreiten, wie diese naturwissenschaftlich untersucht werden können. *Daten und Evidenz naturwissenschaftlich interpretieren* beschreibt die Fähigkeit, unterschiedlich präsentierte naturwissenschaftliche Daten zu analysieren und zu interpretieren, angemessene Schlussfolgerungen zu ziehen sowie Argumente und Evidenz kritisch zu hinterfragen und zu bewerten.

Tabelle 2.2 fasst zusammen, über welches Wissen und welche Handlungsmöglichkeiten eine Person verfügt, die eine ausgeprägte naturwissenschaftliche Grundbildung in den drei Teilkompetenzen besitzt.

2.1.3 Naturwissenschaftliche Wissensbereiche

Die Anwendung der Teilkompetenzen in einer konkreten Situation setzt voraus, dass unterschiedliche Wissensbereiche bei den Fünfzehnjährigen vorhanden sind. Das Wissen, welches zur Beantwortung von Fragestellungen und zum Lösen von Problemen gebraucht wird, lässt sich bei PISA 2015 in drei Bereiche unterteilen: (a) *konzeptuelles Wissen*, (b) *prozedurales Wissen* und (c) *epistemisches Wissen*. Auch hier wurde die Rahmenkonzeption von 2006 weiterentwickelt: Die beiden letztgenannten Bereiche differenzieren das *Wissen über die Naturwissenschaften*, das 2006 betrachtet wurde, weiter aus. Alle Wissensbereiche repräsentieren anschlussfähiges Wissen, das für weiteres, lebenslanges Lernen notwendig ist (Duschl, 2008).

Konzeptuelles Wissen umfasst deklaratives Wissen („Wissen, dass“), das im naturwissenschaftlichen Unterricht in Deutschland zumeist schulfachbezogen vermittelt wird. Für die Frage- und Problemstellungen, denen Jugendliche in ihrem Alltag begegnen, benötigen sie jedoch häufig Wissen aus verschiedenen Disziplinen, welches sie angemessen verknüpfen müssen. Deshalb werden bei PISA 2015 – wie auch in den vorangegangenen Erhebungsrunden – nicht herkömmliche Schulfächer, sondern Wissenssysteme unterschieden: *physikalische Systeme*, *lebende Systeme* sowie *Erd- und Weltraumsysteme*. *Physikalische Systeme* betreffen im Zusammenhang mit der Struktur von Materie beispielsweise sowohl physikalische Modellvorstellungen wie das Teilchenmodell, aber auch chemische Modellvorstellungen zu Bindungen. Im Gegensatz zu PISA 2006 werden die Wissenssysteme nicht nur auf das *konzeptuelle Wissen* bezogen: Auch *prozedurales* und *epistemisches Wissen* werden nun den Wissenssystemen zugeordnet.

Tabelle 2.3 führt einige Beispiele an, welche die einzelnen Kategorien veranschaulichen. Das zur Bewältigung der Fragestellungen notwendige Wissen soll dabei erstens für alltagsbezogene Situationen relevant sein, zweitens ein wichtiges naturwissenschaftliches Konzept oder eine bedeutsame Theorie repräsentieren und drittens dem Entwicklungsstand von Fünfzehnjährigen angemessen sein.

© Waxmann Verlag GmbH

Tabelle 2.3: Drei Wissenssysteme

Physikalische Systeme
• Struktur von Materie (z. B. Teilchenmodell, Bindungen)
• Eigenschaften von Materie (z. B. Zustandsänderungen, thermische und elektrische Leitfähigkeit)
• Chemische Veränderungen von Materie (z. B. chemische Reaktionen, Energietransfer, Säuren/Basen)
• Bewegungen (z.B. Geschwindigkeit) und Kräfte (z.B. Reibungskräfte, magnetische, gravitative und elektrostatische Kräfte)
• Energie und Energieumwandlungen (z. B. Energieerhaltung, Dissipation, chemische Reaktionen)
• Wechselwirkungen von Energie mit Materie (z. B. Licht- und Radiowellen, Schall- und seismische Wellen)
Lebende Systeme
• Zellen (z. B. Strukturen und Funktionen, DNA, pflanzliche und tierische Zellen)
• Aufbau und Funktion von Organismen (z.B. Einzeller und Mehrzeller)
• Menschen (z. B. Gesundheit, Ernährung, Teilsysteme wie beispielsweise Verdauung, Atmung, Kreislauf, Fortpflanzung und deren Beziehung zueinander)
• Populationen (z. B. Arten, Evolution, Artenvielfalt, genetische Vielfalt)
• Ökosysteme (z. B. Nahrungsketten, Energie- und Materiefluss)
• Biosphäre (z. B. ökologische Ressourcen, Nachhaltigkeit)
Erd- und Weltraumsysteme
• Aufbau des Systems Erde (z. B. Lithosphäre, Atmosphäre, Hydrosphäre)
• Energie im System Erde (z. B. Energiequellen, globales Klima)
• Veränderungen im System Erde (z. B. Plattentektonik, geochemische Kreisläufe, konstruktive und destruktive Kräfte)
• Erdgeschichte (z. B. Fossilien, Entstehung der Erde, Evolution)
• Die Erde im Weltraum (z. B. Gravitation, Sonnensysteme, Galaxien)
• Entstehungsgeschichte, Dimensionen und Struktur des Universums (z.B. Lichtjahr, Urknalltheorie)

Prozedurales Wissen („Wissen, wie“) umfasst Prozess- und Methodenwissen, das erforderlich ist, um naturwissenschaftliche Fragestellungen bearbeiten und dabei reliable und valide Daten gewinnen zu können. Dazu gehören das Aufstellen von Hypothesen auf der Basis von Theorien und bisherigen Forschungsbefunden, deren empirisches Überprüfen sowie die Darstellung und Interpretation von Ergebnissen. Das Verständnis und die Durchführung eigener Untersuchungen erfordern Fähigkeiten im Bereich der isolierenden Variablenkontrolle, den Umgang mit und die Minimierung von Messfehlern, das Erkennen von Datenstrukturen oder Datenmustern sowie Methoden der Datenaufbereitung und Präsentation.

Den Schülerinnen und Schülern soll außerdem bewusst sein, dass naturwissenschaftliches Wissen – wissenschaftstheoretisch betrachtet – immer vorläufig ist und dass die Belastbarkeit naturwissenschaftlicher Erkenntnisse maßgeblich davon abhängt, wie sie generiert wurden. Tabelle 2.4 fasst die unterschiedlichen Aspekte zusammen und gibt jeweils Beispiele an.

© Waxmann Verlag GmbH

Tabelle 2.4: Prozedurales naturwissenschaftliches Wissen

Prozedurales naturwissenschaftliches Wissen
• Konzept von Variablen inklusive abhängige, unabhängige und Kontrollvariablen • Konzept von Messungen, z. B. quantitativ (Messwerte), qualitativ (Beobachtungen), Verwendung einer Skala, kategoriale und kontinuierliche Variablen • Möglichkeiten zur Beurteilung und Verringerung von Unsicherheiten, wie Messungen replizieren und Durchschnittswerte bilden • Mechanismen zum Sicherstellen der Replizierbarkeit (Güte der Übereinstimmung zwischen wiederholten Messungen der gleichen Messgröße) und Genauigkeit von Daten (Güte der Übereinstimmung zwischen einer Messgröße und dem wahren Wert eines Messwertes) • Möglichkeiten zur Abstrahierung und Darstellung von Daten mithilfe von Tabellen, Graphen und Abbildungen sowie deren angemessene Nutzung • Variablenkontrollstrategie und ihre Bedeutung in experimentellen Studiendesigns oder die Nutzung randomisierter Kontrollgruppen, um konfundierte Ergebnisse zu verhindern und mögliche kausale Zusammenhänge zu identifizieren • Eigenschaften geeigneter Designs für die Untersuchung bestimmter naturwissenschaftlicher Fragestellungen, z. B. Experiment, Feldstudie oder Mustererkennung

Epistemisches Wissen bezieht sich schließlich auf das Verständnis der Bedeutung spezifischer Konstrukte und wichtiger charakteristischer Eigenschaften der Naturwissenschaften sowie ihrer Funktion beim Generieren naturwissenschaftlichen Wissens (Duschl, 2008). Dazu gehört auch das Wissen über Merkmale der Naturwissenschaften, die den Prozess der Erkenntnisgewinnung leiten. Jugendliche mit einem ausgeprägten epistemischen Wissen sollten zum Beispiel erklären können, was den Unterschied zwischen einer naturwissenschaftlichen Theorie und einer Hypothese ausmacht oder warum es notwendig ist, Messungen und Experimente wiederholt durchzuführen. Sie wissen um die Bedeutung von Modellen in den Naturwissenschaften und verstehen diese als formale, abstrakte Repräsentationen und nicht als exakte Abbildungen der materiellen Welt. Sie sind in der Lage, die Bedeutung von Begriffen der Fach- und Umgangssprache zu unterscheiden. Darüber hinaus haben sie eine Vorstellung davon entwickelt, wie naturwissenschaftliches Wissen von Wissenschaftlerinnen und Wissenschaftlern generiert und formuliert wird und dass dabei dem wissenschaftlichen Austausch (Kooperation, Wettbewerb, gegenseitige Kontrolle durch Peer Reviews) eine zentrale Rolle zukommt. Tabelle 2.5 stellt Aspekte epistemischen Wissens mit Beispielen dar.

© Waxmann Verlag GmbH

Tabelle 2.5: Epistemisches naturwissenschaftliches Wissen

Epistemisches naturwissenschaftliches Wissen
Die Konstrukte und charakteristischen Eigenschaften von Naturwissenschaften:
• Eigenschaften naturwissenschaftlicher Beobachtungen, Fakten, Hypothesen, Modelle und Theorien • Ziele von Naturwissenschaften (Erklärungen für die natürliche Welt generieren) im Gegensatz zu Technik (eine optimale Lösung für menschliche Bedürfnisse finden) • Eigenschaften naturwissenschaftlicher und technischer Fragestellungen sowie geeigneter Daten • Wertvorstellungen in den Naturwissenschaften, z. B. Offenlegung von Ergebnissen, Objektivität, und das Ausschließen von Verzerrungen • Arten naturwissenschaftlichen Argumentierens, z. B. deduktiv, induktiv, Inferenz der besten Erklärung (abduktiv), analogisch, modellbasiert
Die Rolle, die diese Konstrukte und Eigenschaften spielen, um das generierte naturwissenschaftliche Wissen zu belegen:
• Wie naturwissenschaftliche Aussagen durch Daten und evidenzbasiertes Argumentieren unterstützt werden. • Welche verschiedenen Formen empirischer Untersuchungen zum Aufbau von Wissen es gibt, sowie die Ziele (Hypothesen testen oder Muster identifizieren) und Designs (Beobachtungen, kontrollierte Experimente, Korrelationsstudien) dieser Untersuchungen. • Welchen Einfluss Messfehler auf das Vertrauen in naturwissenschaftliche Ergebnisse und daraus generiertes Wissen haben. • Welchen Nutzen und welche Einschränkungen reale oder abstrakte Modelle sowie Modelle von Systemen haben. • Welche Bedeutung Zusammenarbeit und Kritik haben, und wie *Peer Reviews* dazu beitragen, das Vertrauen in naturwissenschaftliche Aussagen und Befunde zu stärken. • Welche Bedeutung naturwissenschaftliches Wissen (gemeinsam mit anderen Formen von Wissen) für das Erkennen und Bearbeiten gesellschaftlicher oder technischer Fragestellungen hat.

2.1.4 Motivationale Orientierungen und Einstellungen zu den Naturwissenschaften

Neben den Wissensbereichen wurde bei PISA 2006 erstmals auch die Bedeutung motivationaler Orientierungen und Einstellungen für die naturwissenschaftliche Grundbildung explizit hervorgehoben und in die Rahmenkonzeption aufgenommen (OECD, 2006; Prenzel et al., 2007). Ob und wie Jugendliche ihre Kompetenzen in einer spezifischen Problemsituation nutzen, wird neben dem Wissen auch durch motivationale Orientierungen und Einstellungen beeinflusst (vgl. Kapitel 3). Wissen allein hilft im Alltag wenig, wenn Schülerinnen und Schüler sich nicht für Naturwissenschaften interessieren oder ihre Bedeutung für den Alltag oder das Weltgeschehen nicht nachvollziehen können. Entsprechend mehrdimensionale Bildungsziele, die neben kognitiven Zielen in Form von Fachleistungen auch motivationale Orientierungen und Einstellungen umfassen, sind in Deutschland in vielen Schulgesetzen und Lehrplänen verankert. Diese Zielsetzungen sollten eine eng gefasste, rein kognitive Leistungsorientierung überwinden, geraten aber im schulischen Alltag leicht aus dem Blick (vgl. Aktionsrat Bildung, 2015).

© Waxmann Verlag GmbH

Wie bei PISA 2006 werden das *Interesse an Naturwissenschaften*, die *Wertschätzung naturwissenschaftlicher Herangehensweisen zur Erkenntnisgewinnung* sowie das *Verantwortungsbewusstsein gegenüber der Umwelt* betrachtet. Im Gegensatz zu PISA 2006 werden diese Orientierungen und Einstellungen jedoch nicht mehr über Fragen im Rahmen des kognitiven Tests erhoben. Bei PISA 2015 sind sie Bestandteil des Schülerfragebogens und werden zum Teil im Rahmen des Kapitels zu den Schülermerkmalen in diesem Band differenziert betrachtet (vgl. Kapitel 3).

2.2 Der Naturwissenschaftstest in PISA 2015

2.2.1 Entwicklung der Aufgaben und ihre Formate

Durch den Status als Hauptdomäne stand für die Erfassung der naturwissenschaftlichen Kompetenz in dieser Runde ein wesentlicher Anteil der Testzeit zur Verfügung. Entsprechend konnte eine große Zahl von Aufgaben eingesetzt werden, welche auch die Teilkompetenzen differenziert erfassten. Die Aufgaben wurden unter Bezugnahme auf die Rahmenkonzeption von verschiedenen Autorinnen und Autoren entwickelt (z. B. nationale Projektteams, Mitglieder der *Science Expert Group,* PISA-Projektkonsortium). Die Aufgaben wurden anschließend von Expertinnen und Experten aus den teilnehmenden Staaten eingeschätzt und im Feldtest pilotiert. Kriterien für die Auswahl einer Aufgabe für die Hauptstudie waren:

- Passung zur Rahmenkonzeption,
- Relevanz für Fünfzehnjährige,
- Angemessenheit für verschiedene kulturelle und curriculare Kontexte,
- Erfüllung technischer Qualitätsstandards und
- internationale Vergleichbarkeit.

Der Naturwissenschaftstest bei PISA 2015 setzte sich aus zwei Teilen zusammen, nämlich aus Aufgaben, die bereits 2006 zum Einsatz kamen (sogenannte Link-Aufgaben, die nicht veröffentlicht wurden) und aus neu entwickelten Aufgaben. Alle diese Aufgaben wurden erstmals computerbasiert administriert (vgl. Kapitel 12). Dabei wurden auch die zuvor auf Papier dargebotenen Link-Aufgaben in das neue Format übertragen. Die in diesem Kapitel präsentierten Aufgabenbeispiele *Völkerkollaps bei Bienen* (Abbildungen 2.2 bis 2.4) und *Energieeffiziente Häuser* (Abbildungen 2.5 bis 2.9) zeigen die Darbietung am Computer.

Die neu entwickelten Aufgaben nutzen zum Teil zusätzliche Möglichkeiten eines computerbasierten Tests. Manche Aufgaben weisen ein eher klassisches Format auf – ähnlich zu den Link-Aufgaben (siehe Aufgabenbeispiel *Völkerkollaps bei Bienen*). Andere, sogenannte interaktive Aufgaben, sind spezifisch auf die computerbasierte Testung abgestimmt und simulieren etwa die Einflüsse unterschiedlicher Variablen (siehe

© Waxmann Verlag GmbH

Aufgabenbeispiel *Energieeffiziente Häuser*). So wurde die Teilkompetenz *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* unter anderem darüber erfasst, dass Schülerinnen und Schüler selbstständig ein (simuliertes) Experiment planten und dessen Ergebnisse interpretierten. Die computerbasierte Erhebung ermöglicht außerdem das Erfassen von Prozessdaten. Damit kann man zum einen die Qualität der erhobenen Daten in der Testsituation und im Anschluss besser überprüfen. Zum anderen liefern diese Daten Einblicke in das Bearbeitungsverhalten der Schülerinnen und Schüler und können helfen, die Bearbeitung der Aufgaben besser zu verstehen (vgl. Kapitel 12).

Alle Aufgaben bestanden aus Stimulusmaterial und bis zu fünf Teilaufgaben, die üblicherweise als Items bezeichnet werden (Osterlind, 1990). Die Stimuli betten die Items in einen der Kontexte ein. Tabelle 2.6 zeigt die Zuordnung zu den Kontexten.

Tabelle 2.6: Verteilung der Items des Naturwissenschaftstests auf die Kontexte

Kontext	Persönlich	Regional/ National	Global	Gesamt
Gesundheit und Krankheit	3	13	6	22
Natürliche Ressourcen	0	35	11	46
Umwelt	0	23	11	34
Risiken/Gefahren	1	15	4	20
Grenzen von Naturwissenschaften und Technik	17	22	23	62
Gesamt	21	108	55	184

Insgesamt wurden 26 neu entwickelte (99 Items) und 30 Link-Aufgaben (85 Items) eingesetzt. Von den neu entwickelten Aufgaben waren fünf Aufgaben interaktiv (24 Items) und 21 Standard-Aufgaben (75 Items). Die Items lassen sich den Teilkompetenzen sowie den Wissensbereichen und Wissenssystemen zuordnen. Tabelle 2.7 spiegelt diese Zuordnung wider. Auch wenn aus theoretischer Sicht *prozedurales* und *epistemisches Wissen* voneinander getrennt werden können, werden diese (aufgrund der zu geringen Anzahl an Items) in einer Skala zusammengefasst und berichtet. Im Gegensatz zu PISA 2006 wurden in PISA 2015 nicht nur die Items des Wissensbereichs *konzeptuelles Wissen,* sondern auch die Items der Bereiche *prozedurales* und *epistemisches Wissen* den Wissenssystemen *physikalische Systeme*, *lebende Systeme* sowie *Erd- und Weltraumsysteme* zugeordnet.

© Waxmann Verlag GmbH

Tabelle 2.7: Verteilung der Items des Naturwissenschaftstests auf die drei Teilkompetenzen, Wissensbereiche und Wissenssysteme

Bereich der naturwissenschaftlichen Kompetenz	Anzahl Link-Items	Anzahl neue Items	Anzahl Items gesamt
Teilkompetenzen	85	99	184
Phänomene naturwissenschaftlich erklären	41	48	89
Naturwissenschaftliche Forschung bewerten und Untersuchungen planen	16	23	39
Daten und Evidenz naturwissenschaftlich interpretieren	28	28	56
Wissensbereiche	85	99	184
Konzeptuelles Wissen	51	47	98
Prozedurales Wissen	24	36	60
Epistemisches Wissen	10	16	26
Wissenssysteme	85	99	184
Physikalische Systeme	28	33	61
Lebende Systeme	39	35	74
Erd- und Weltraumsysteme	18	31	49

Neu in PISA 2015 ist, dass sich die Items auch nach dem Grad der kognitiven Anforderungen, die zur Lösung benötigt werden, einteilen lassen. Dieser Grad bestimmt die Schwierigkeit eines Items und wird durch vier Faktoren beeinflusst:

- Anzahl und Grad der Komplexität der Wissenselemente,
- Kenntnis des für die Aufgabe notwendigen inhaltlichen, prozeduralen und epistemischen Wissens,
- die kognitive Handlung, die das Item erfordert, z. B. Erinnern, Analysieren, Bewerten,
- das Ausmaß, in dem das Formulieren der Antwort von Modellen oder abstrakten naturwissenschaftlichen Ideen abhängig ist.

Es werden drei kognitive Anforderungsniveaus – niedrig, mittel, hoch – unterschieden, die in Tabelle 2.8 erläutert werden.

Bei der Bearbeitung der Aufgaben hatten die Fünfzehnjährigen drei verschiedene Antwortformate zu bewältigen. Das erste Format waren einfache Auswahl-Items (Multiple-Choice-Items; insgesamt 54), bei denen aus vier vorgegebenen Antwortoptionen die richtige Antwort ausgewählt werden musste (siehe Beispielaufgabe *Völkerkollaps bei Bienen*, Abbildung 2.4 oder *Energieeffiziente Häuser*, Abbildung 2.9). Bei den für PISA 2015 neu entwickelten interaktiven Items konnte als Antwortoption teilweise auch ein Element aus einer Grafik oder einem Text ausgewählt werden. Das zweite zu bewältigende

© Waxmann Verlag GmbH

Tabelle 2.8: Kognitive Anforderungsniveaus in PISA 2015

Kognitives Anfoderungsniveau	Beschreibung	Anzahl der Items
hoch	Die Lösung des Items erfordert die Analyse komplexer Informationen oder Daten, die Zusammenfassung oder Bewertung von Evidenz, das Belegen von Behauptungen, Argumentieren (anhand unterschiedlicher Quellen) oder das Entwickeln eines Vorgehens, wie man an ein Problem herangeht.	15
mittel	Die Lösung des Items erfordert die Nutzung und Anwendung konzeptuellen Wissens, um Phänomene zu beschreiben oder zu erklären, die Auswahl geeigneter Vorgänge/Prozeduren, die zwei oder mehr Schritte erfordern, die Organisation und Darstellung von Daten oder die Interpretation und Nutzung einfacher Datensätze oder Graphen.	113
niedrig	Die Lösung des Items erfordert einen einzelnen Vorgang/eine Prozedur, z.B. Fakten, Begriffe, Prinzipien oder Konzepte abzurufen oder eine einzelne Information aus einem Graphen oder einer Tabelle herauszulesen.	56

Antwortformat waren komplexe Mehrfachwahl-Antworten (66 Items). Wie bei PISA 2006 musste zum einen für jede der dargebotenen Aussagen beurteilt werden, ob diese richtig oder falsch ist. Zusätzlich konnten bei den neu entwickelten Items aus einer Liste von Antworten eine oder mehrere Antworten ausgewählt, (mehrere) Lücken eines Satzes durch das Auswählen einer Antwort aus einem Drop-down-Menü vervollständigt (siehe Beispielaufgabe *Völkerkollaps bei Bienen*, Item 3) oder Antworten durch *drag and drop* gegeben werden. Letzteres bedeutet, dass Elemente am Bildschirm bewegt werden müssen, um Antworten zuzuordnen, in eine Rangfolge zu bringen oder zu kategorisieren. Als drittes Antwortformat mussten offene Antworten gegeben werden (64 Items). Die meisten Items erforderten als Antwort eine Wortgruppe, einen Satz oder einen kurzen Absatz (z. B. zwei bis vier Sätze zur Erklärung; siehe Beispielaufgabe *Völkerkollaps bei Bienen*, Abb. 2.2 oder *Energieeffiziente Häuser*, Abb. 2.8). Bei einigen wenigen interaktiven Aufgaben musste etwas gezeichnet werden (Graph oder Diagramm). Die offenen Antwortformate wurden je nach Komplexität entweder computerbasiert oder von geschulten Kodiererinnen und Kodierern anhand ausführlicher Kodieranweisungen ausgewertet (vgl. auch Kapitel 12).

2.2.2 Aufgabenbeispiele

Im Folgenden werden eine traditionelle sowie eine interaktive Beispielaufgabe präsentiert, die im Feldtest zum Einsatz kamen. Ähnliche Aufgaben wurden in der Hauptstudie verwendet. Für diese wurden diejenigen Aufgaben ausgewählt, die den oben genannten Kriterien am besten entsprachen. Aufgaben, die nicht in der Hauptstudie zum Einsatz kamen, wurden zum Teil als Beispiele veröffentlicht.

© Waxmann Verlag GmbH

Die Aufgabe *Völkerkollaps bei Bienen* (Abbildungen 2.2 bis 2.4) ist eine im Prinzip traditionelle PISA-Aufgabe, die für den Einsatz am Computer neu entwickelt wurde. Bei dieser Aufgabe beschäftigen sich Schülerinnen und Schüler mit möglichen Ursachen dafür, dass Bienenvölker ihren Bienenstock aufgeben und verlassen. Die gesamte Aufgabe ist dem Kontext *Umwelt (regional/national)* zugeordnet (vgl. Tabelle 2.9). Bei Item 1 sollen die Fünfzehnjährigen in einem offenen Antwortfeld erklären, warum das Verschwinden von Bienen auch einen Einfluss auf die vorhandene Vogelpopulation haben kann (Abbildung 2.2). Um diese Frage richtig zu beantworten, müssen sie erläutern, dass viele Pflanzenarten, beispielsweise die Sonnenblume, ohne die Bestäubung ihrer Blüten durch Insekten keine Samen produzieren können. Werden also weniger Sonnenblumen bestäubt und weniger Samen produziert, kann sich das negativ auf Vogelpopulationen auswirken, die sich von Samen ernähren. Das Item ist der Teilkompetenz *Phänomene naturwissenschaftlich erklären* zugeordnet und erfasst *konzeptuelles Wissen* im Wissensbereich *lebende Systeme*. Das kognitive Anforderungsniveau wird als „mittel“ eingestuft.

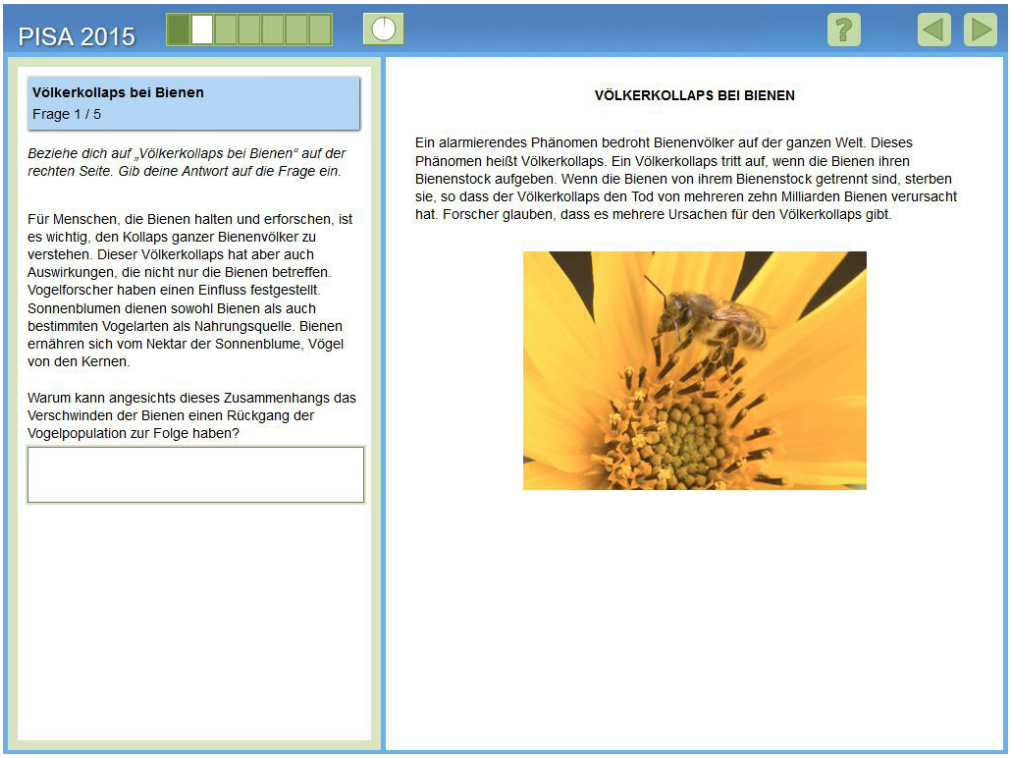


Abbildung 2.2: Beispielaufgabe *Völkerkollaps bei Bienen*, Item 1

© Waxmann Verlag GmbH

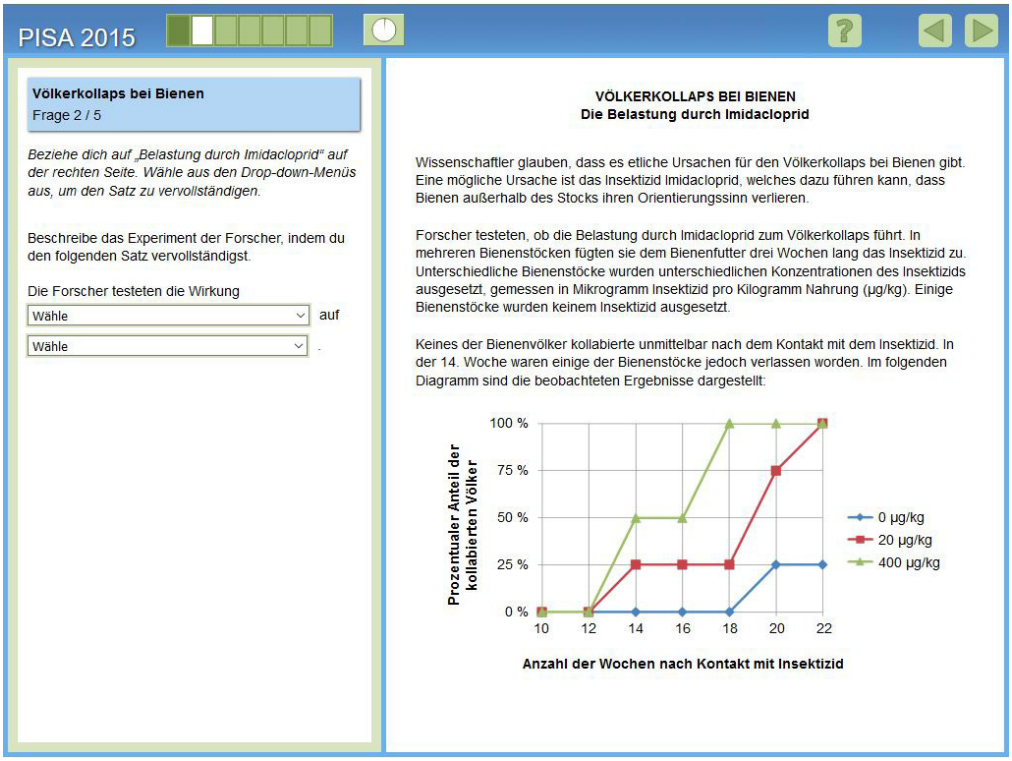


Abbildung 2.3: Beispielaufgabe *Völkerkollaps bei Bienen*, Item 2

Item 2 (Abbildung 2.3) beschreibt ein wissenschaftliches Experiment, das den Einfluss eines Insektizids auf den Völkerkollaps untersucht. Um die Frage zu beantworten, sollen die Schülerinnen und Schüler einen Satz mit zwei Lücken durch das Auswählen von Antworten aus zwei Drop-down-Menüs vervollständigen und damit die unabhängige und abhängige Variable in diesem Experiment identifizieren. Als Antwortoptionen stehen der Kollaps von Bienenvölkern, die Konzentration von Imidacloprid im Futter und die Immunität von Bienen gegenüber Imidacloprid zur Auswahl. Als richtig gewertet wird der Satz: „Die Forscher testeten die Wirkung der Konzentration von Imidacloprid im Futter auf den Kollaps von Bienenvölkern." Das Item ist der Teilkompetenz *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* zuzuordnen. Es erfasst *prozedurales Wissen* ebenfalls im Bereich *lebende Systeme* und entspricht einem mittleren kognitiven Anforderungsniveau.

© Waxmann Verlag GmbH

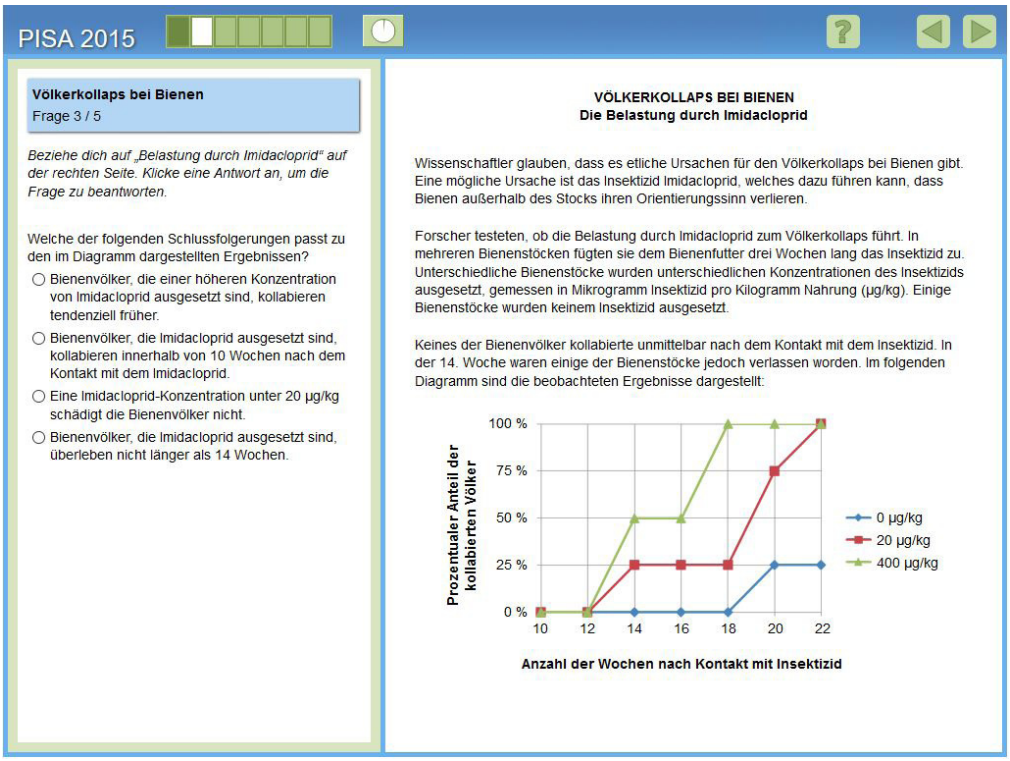


Abbildung 2.4: Beispielaufgabe *Völkerkollaps bei Bienen*, Item 3

Tabelle 2.9: Einordnung Beispielaufgabe *Völkerkollaps bei Bienen*

	Item 1	**Item 2**	**Item 3**
Teilkompetenz	Phänomene naturwissenschaftlich erklären	Naturwissenschaftliche Forschung bewerten & Untersuchungen planen	Daten und Evidenz naturwissenschaftlich interpretieren
Wissensbereich	Konzeptuelles Wissen	Prozedurales Wissen	Prozedurales Wissen
Wissenssystem	Lebende Systeme	Lebende Systeme	Lebende Systeme
Kontext	Regional/National – Umwelt	Regional/National – Umwelt	Regional/National – Umwelt
Kognitives Anforderungniveau	Mittel	Mittel	Mittel
Antwortformat	Offene Antwort	Komplexe Mehrfachwahl-Antwort	Einfache Auswahl-Antwort

© Waxmann Verlag GmbH

Item 3 (Abbildung 2.4) bezieht sich auf das gleiche Experiment wie Item 2. Die Fünfzehnjährigen sollen nun das Diagramm interpretieren, welches die Beziehung zwischen dem Anteil der kollabierten Bienenvölker und der Konzentration des Insektizids darstellt. Aus mehreren Antworten muss diejenige ausgewählt werden, welche die zum Diagramm passende Schlussfolgerung liefert (Antwortoption 1). Das Diagramm zeigt, dass der prozentuale Anteil der kollabierten Völker höher ist, wenn die Bienenstöcke im Zeitraum zwischen der 14. und 20. Woche des Experiments einer Konzentration von 400 µg/kg des Insektizids ausgesetzt wurden. Das Item erfasst *prozedurales Wissen* im Bereich *lebende Systeme* der Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren*. Auch diese Aufgabe entspricht einem mittleren Anforderungsniveau.

Die zweite Beispielaufgabe *Energieeffiziente Häuser* (Abbildungen 2.5 bis 2.9) ist ein interaktives Item, bei dem die Jugendlichen mithilfe einer Simulation untersuchen sollen, wie sich unterschiedliche Dachfarben in Abhängigkeit von der Außentemperatur auf den Energieverbrauch[1] eines Hauses auswirken. Zunächst werden in einer Einleitung der Kontext der Aufgabe (*natürliche Ressourcen – regional/national*) dargestellt (Abbildung 2.5) und die Simulation erläutert (Abbildung 2.6). Die Schülerinnen und Schüler haben die Möglichkeit, auszuprobieren und zu üben, wie diese funktioniert. Wenn sie nicht innerhalb einer Minute die angegebenen Schritte wie „Klicke auf eine Dachfarbe" ausführen, erscheint eine Hilfenachricht. Nach zwei Minuten Inaktivität werden die angegebenen Schritte automatisch ausgeführt und angezeigt. Anschließend bearbeiten die Fünfzehnjährigen mithilfe der Simulation die Items. Bei jedem Item gibt es die Schaltfläche „So führst du die Simulation aus". Wenn diese angeklickt wird, erscheint eine Erinnerung, wie die Simulation ausgeführt wird und wie eine Datenreihe ausgewählt oder gelöscht werden kann.

Bei Item 1 (Abbildung 2.7) soll ein Haus bei sehr heißen Außentemperaturen auf 23°C abgekühlt werden. Für die unterschiedlichen Dachfarben ist dafür unterschiedlich viel Energie erforderlich. Die Jugendlichen sollen die Dachfarben nach dem Energieverbrauch ordnen. Für die korrekte Antwort müssen sie eine Außentemperatur von 40°C auswählen und sich für jede Dachfarbe den Energieverbrauch anzeigen lassen. Anschließend müssen die Dachfarben in eine Reihenfolge gebracht werden. Die ausgewählte Reihenfolge soll dann mit Daten aus der Tabelle gestützt werden. Hierzu sollen die Schülerinnen und Schüler drei Zeilen aus der Tabelle durch Anklicken auswählen. Die richtige Reihenfolge ist schwarz (höchster Energieverbrauch bei 40°C), rot (mittel) und weiß (am geringsten). Das Item erfasst die Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren*. Sie ist dem *prozeduralen Wissen* im Bereich *physikalische Systeme* zuzuordnen (vgl. Tabelle 2.10).

1 Der Begriff „Energieverbrauch" wird in dieser Aufgabe, obwohl aus physikalischer Sicht nicht angemessen, bewusst verwendet. Es ist ein gängiger Begriff der Alltagssprache und 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 durch die Verwendung des Begriffs möglicherweise Alltagsvorstellungen von Energie ausgelöst werden, die Schülerinnen und Schüler in der korrekten Bearbeitung der Aufgabe behindern, soll mit erfasst werden.

© Waxmann Verlag GmbH

Abbildung 2.5: Beispielaufgabe *Energieeffiziente Häuser,* Einleitung

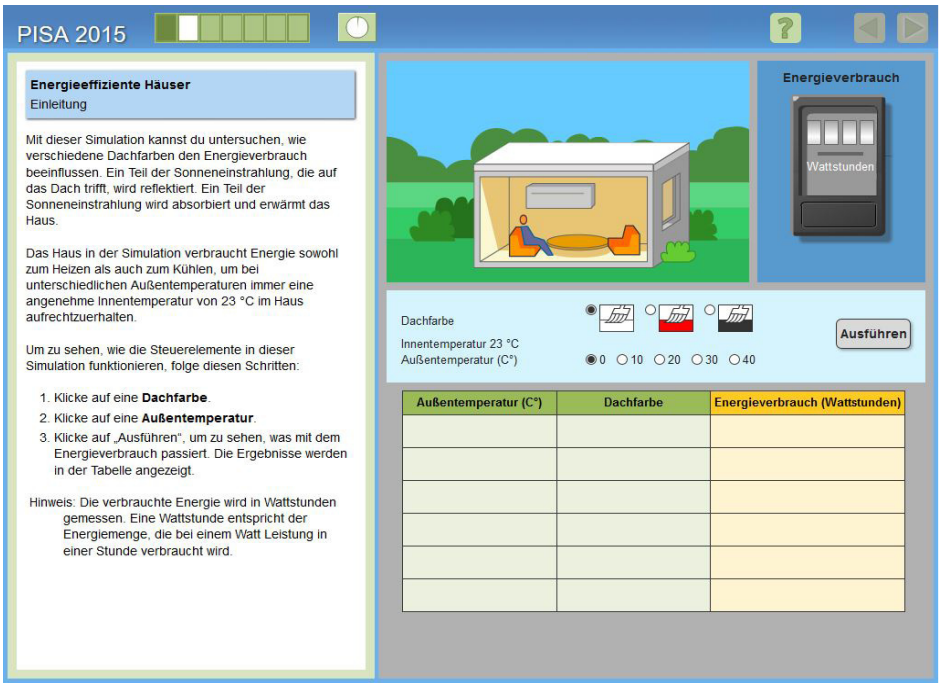


Abbildung 2.6: Beispielaufgabe *Energieeffiziente Häuser*, Erläuterung der Simulation

© Waxmann Verlag GmbH

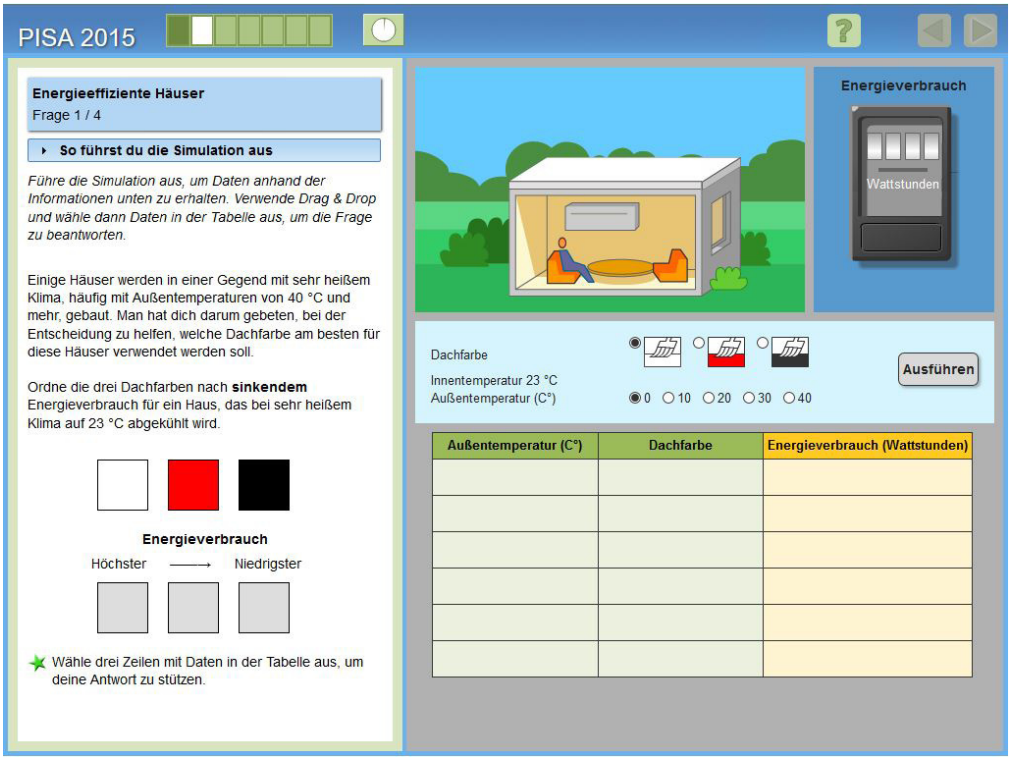


Abbildung 2.7: Beispielaufgabe *Energieeffiziente Häuser,* Item 1

Das Item spiegelt ein niedriges kognitives Anforderungsniveau wider, da im Wesentlichen Informationen aus einer Tabelle entnommen werden müssen.

Item 2 (Abbildung 2.8) besteht aus zwei Fragen mit jeweils unterschiedlichem Antwortformat. Zur Beantwortung sollen die Jugendlichen die Simulation verwenden, um den Energieverbrauch eines Hauses mit einem weißen Dach mit dem eines Hauses mit einem schwarzen Dach bei 10°C Außentemperatur und einer gewünschten Innentemperatur von 23 °C zu vergleichen. Zunächst müssen sie den Satz vervollständigen „Bei 10 °C verbraucht ein Haus mit einem weißen Dach ___ Energie als ein Haus mit einem schwarzen Dach". Das Drop-down-Menü bietet „mehr" und „weniger" zur Auswahl. Die richtige Antwort lautet „mehr Energie". Daten aus der Simulation, die diese Antwort belegen, wären zwei Zeilen mit der Außentemperatur 10°C und den Dachfarben weiß und schwarz. Diese Frage ist der Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren* zugeordnet. Sie erfasst *prozedurales Wissen* im Bereich *physikalische Systeme* und bildet ebenso wie Frage 2 ein mittleres kognitives Anforderungsniveau ab. Anschließend sollen die Jugendlichen den Unterschied im Energieverbrauch bei Verwendung der beiden Dachfarben mit eigenen Worten erklären. Sie sollen in ihrer Erklärung dabei insbesondere die Wirkung der Sonneneinstrahlung berücksichtigen. Eine korrekte Antwort bezieht sich auf die Strahlungsenergie der Sonne als Energiequelle,

© Waxmann Verlag GmbH

die von schwarzen Dächern in höherem Maße absorbiert wird als von weißen Dächern. Diese Frage erfasst *konzeptuelles Wissen* im Bereich *physikalische Systeme* sowie die Teilkompetenz *Phänomene naturwissenschaftlich erklären* (vgl. Tabelle 2.10).

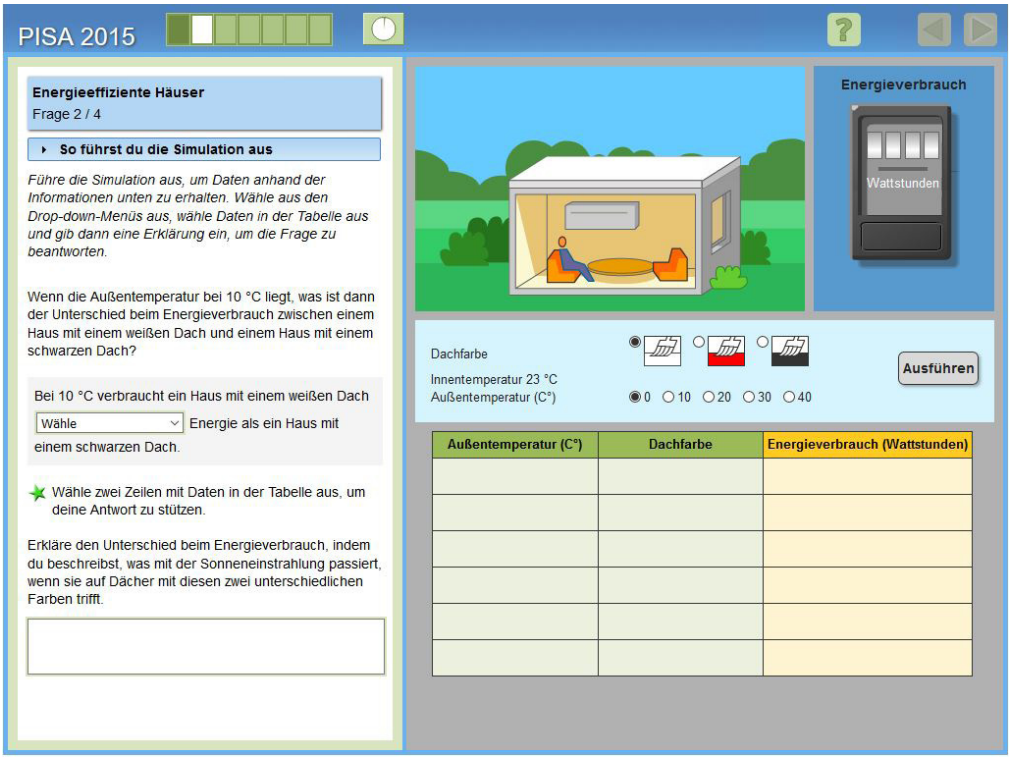


Abbildung 2.8: Beispielaufgabe *Energieeffiziente Häuser*, Item 2

Das dritte Item (Abbildung 2.9) fragt nach einer Verallgemeinerung des Befundes aus Item 2. Die Schülerinnen und Schüler sollen den Zusammenhang zwischen Außentemperatur und Energieverbrauch für den gesamten Temperaturbereich und alle Dachfarben, die durch die Simulation unterstützt werden, ermitteln. Dazu wählen die Jugendlichen aus vier möglichen Antworten eine aus. Die richtige Antwort ist Option 3: „Wenn sich der Unterschied zwischen der Außentemperatur und der Innentemperatur vergrößert, steigt der Energieverbrauch.“ Das Item erhebt *konzeptuelles Wissen* im Bereich *physikalische Systeme* und ist der Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren* zugeordnet. Da hier mehrere komplexe Informationen verarbeitet werden müssen, wird das kognitive Anforderungsniveau der Aufgabe als hoch eingeschätzt. Nur wenige Jugendliche sind in der Lage, dieses Item korrekt zu bearbeiten.

© Waxmann Verlag GmbH

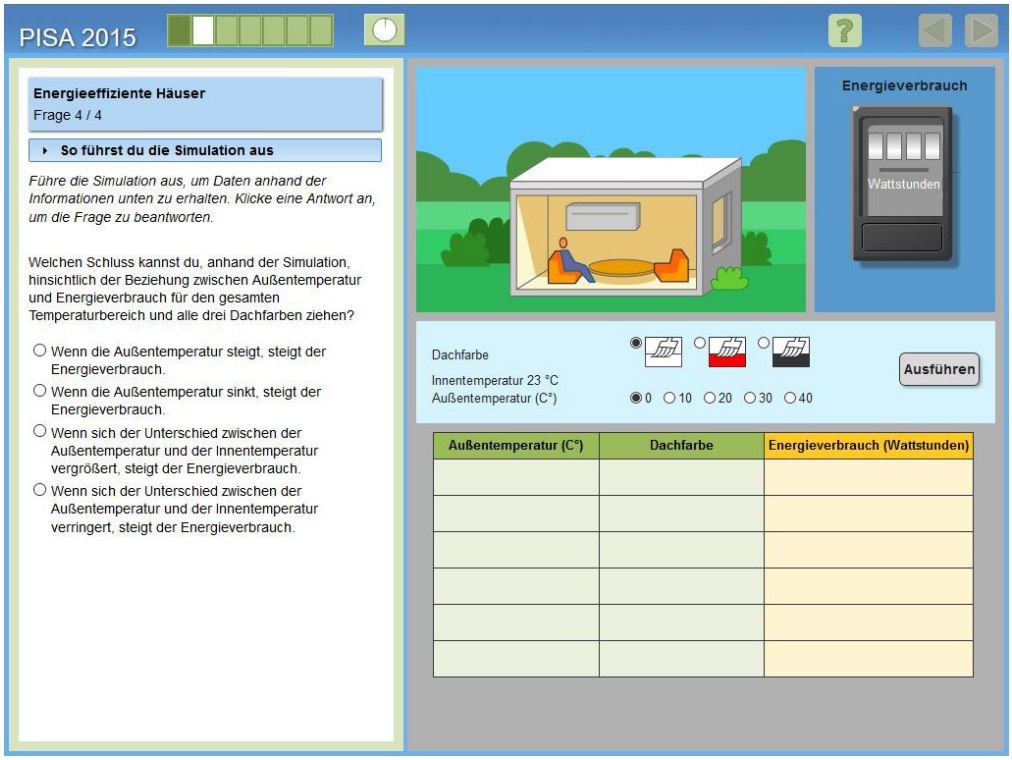


Abbildung 2.9: Beispielaufgabe *Energieeffiziente Häuser*, Item 3

Tabelle 2.10: Einordnung Beispielaufgabe *Energieeffiziente Häuser*

	Item 1	Item 2		Item 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
Wissenssystem	Physikalische Systeme	Physikalische Systeme	Physikalische Systeme	Physikalische Systeme
Kontext	Regional/National – Natürliche Ressourcen	Regional/National – Natürliche Ressourcen	Regional/National – Natürliche Ressourcen	Regional/National – Natürliche Ressourcen
Kognitives Anforderungniveau	Niedrig	Mittel	Mittel	Hoch
Antwortformat	Offene Antwort	Offene Antwort	Offene Antwort	Einfache Auswahl-Antwort

© Waxmann Verlag GmbH

2.2.3 Auswertung und Kompetenzstufen

PISA untersucht, auf welchem Niveau Schülerinnen und Schüler über naturwissenschaftliche Grundbildung verfügen. Ob eine Aufgabe für eine Person schwierig oder einfach ist, hängt von unterschiedlichen Einflussfaktoren ab, wie zum Beispiel dem Vorwissen, der Freude an Naturwissenschaften oder dem Kontext, in den die Aufgabe eingebettet ist. Der Schwierigkeitsgrad einer Aufgabe wird auf Basis der Antworten der Schülerinnen und Schüler und unter Nutzung psychometrischer Modelle der Item-Response-Theorie bestimmt (vgl. Kapitel 12). Diese Modelle erlauben es, die Schwierigkeit der Aufgaben und die Fähigkeit der Personen auf einer gemeinsamen Skala abzubilden. Diese Skala wird in Abschnitte unterteilt, die Items mit vergleichbaren Schwierigkeiten zusammenfassen. Die Abschnitte werden als Kompetenzstufen bezeichnet. Es werden wie in PISA 2006 sechs Kompetenzstufen unterschieden. Neu in PISA 2015 ist eine Unterteilung der untersten Kompetenzstufe I in Ia und Ib, um im unteren Leistungsbereich besser differenzieren zu können.

Die Kompetenzstufen werden inhaltlich definiert und erläutert (Tabelle 2.11). Die Definitionen erfolgen entsprechend der aktuellen, überarbeiteten Rahmenkonzeption. Mit ansteigender Kompetenzstufe werden die Anforderungen, welche die Aufgaben an die Fähigkeiten der Schülerinnen und Schüler stellen, komplexer. Einfache Items, die zum Beispiel grundlegendes Wissen in bekannten Kontexten erfordern, lassen sich den unteren Kompetenzstufen zuordnen. Im Gegensatz dazu müssen Schülerinnen und Schüler zur Lösung von Items, die im oberen Bereich der Skala liegen, auf abstraktere Konzepte und Ideen zurückgreifen, mit deren Hilfe sie auch unbekannte, komplexe naturwissenschaftliche Phänomene, Ereignisse und Prozesse erklären können. Je höher die naturwissenschaftliche Kompetenz einer oder eines Jugendlichen ausgeprägt ist, desto wahrscheinlicher wird sie oder er Items, die von ihrer Schwierigkeit her einer hohen Kompetenzstufe zugeordnet sind, lösen.

Neben der Gesamtskala naturwissenschaftlicher Kompetenz wurden aufgrund des Status als Hauptdomäne und des damit einhergehenden größeren Aufgabenpools auch Skalen für die Teilkompetenzen *Phänomene naturwissenschaftlich erklären*, *Naturwissenschaftliche Forschung bewerten und Untersuchungen planen* und *Daten und Evidenz naturwissenschaftlich interpretieren* gebildet. Außerdem wurden die Leistungen getrennt nach den Wissensbereichen skaliert. Hier wurden jedoch die Bereiche *prozedurales und epistemisches Wissen*, die aus theoretischer Sicht voneinander getrennt wurden, aufgrund der zu geringen Anzahl an Aufgaben in der Hauptstudie zu einer Skala zusammengefasst und berichtet. Auch für die Wissenssysteme *physikalische Systeme*, *lebende Systeme* sowie *Erd- und Weltraumsysteme* wurden Teilskalen gebildet. Diese Teilskalen wurden auch bei PISA 2006 berichtet. Allerdings wurden damals nur Items des *konzeptuellen Wissens* den Wissenssystemen zugeordnet. Bei PISA 2015 hat man sich für ein anderes Vorgehen entschieden und auch die Items des *prozeduralen und epistemischen Wissens* den Systemen zugeordnet, da die Bearbeitung aller Aufgaben Wissen über unterschiedliche Systeme erfordert.

© Waxmann Verlag GmbH

Tabelle 2.11: Stufen naturwissenschaftlicher Kompetenz in PISA 2015

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, zugreifen. 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 begrenzten 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 könnten.
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 grafische 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 bzw. 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.

© Waxmann Verlag GmbH

2.2.4 Bemerkungen zur Vergleichbarkeit der Ergebnisse von PISA 2015 und PISA 2006

Auch wenn die Modifikation der Rahmenkonzeption und die Umstellung des Darbietungsmodus der Testaufgaben von Papier auf den Computer sowie die damit verbundenen neu entwickelten Aufgabenformate in PISA 2015 sinnvolle Weiterentwicklungen sind, so werfen sie doch die Frage auf, inwieweit der Anschluss und die Vergleichbarkeit mit den Ergebnissen vorheriger Erhebungsrunden gewährleistet ist. Für einen möglichen Effekt der veränderten Darbietungsform wurde anhand der Feldtestdaten aller teilnehmenden Staaten (Frühjahr 2014) untersucht, ob die Darbietungsmodalität am Computer im Vergleich zu Papier-Bleistift die Schwierigkeit der Aufgaben beeinflusst. In der Hauptstudie sollten nur solche Aufgaben zum Einsatz kommen, für die international im Feldtest keine Effekte der Darbietungsform Computer versus Papier-Bleistift gefunden wurden (vgl. OECD, 2016b; OECD, in Vorbereitung).

Am Zentrum für Internationale Bildungsvergleichsstudien (ZIB) durchgeführte Analysen der in Deutschland erhobenen Feldtestdaten lieferten allerdings Hinweise darauf, dass für Schülerinnen und Schüler in Deutschland die naturwissenschaftlichen Aufgaben am Computer im Mittel schwieriger waren als in der Papier-Version (Robitzsch et al., im Druck). Ein Item wurde also mit geringerer Wahrscheinlichkeit gelöst, wenn es nicht in einem Testheft, sondern am Computer vorgegeben wurde. Hinsichtlich möglicher Geschlechterdifferenzen deuteten sich keine Unterschiede an. Diese Befunde sollten bei der Interpretation der nationalen Ergebnisse, die Veränderungen der naturwissenschaftlichen Kompetenz beschreiben (vgl. Abschnitt 2.4), berücksichtigt werden. An dieser Stelle sei auch noch einmal darauf hingewiesen, dass ein Vergleich der Teilkompetenzen, Wissensbereiche und Wissenssysteme mit denen von 2006 nicht möglich ist, da diese in der inhaltlichen Definition verändert und Aufgaben den Teilskalen verschiedentlich anders zugeordnet wurden (vgl. Abschnitt 2.1.3).

2.3 Ergebnisse des internationalen Vergleichs

Im Folgenden werden die Ergebnisse des internationalen Vergleichs vorgestellt und zunächst die Mittelwerte, Streuungen und Verteilungen auf die Kompetenzstufen sowie Geschlechterdifferenzen der Gesamtskala naturwissenschaftliche Kompetenz berichtet. Es wird der Frage nachgegangen, wie sich die Ergebnisse für Deutschland in das Gesamtbild der OECD-Staaten einordnen lassen. Da die Naturwissenschaften in PISA 2015 die Hauptdomäne sind, können darüber hinaus die verschiedenen Teilkompetenzen, Wissensbereiche und Wissenssysteme differenziert untersucht werden. Diese werden ebenfalls mit Blick auf mögliche Geschlechterunterschiede betrachtet.

© Waxmann Verlag GmbH

2.3.1 Die Gesamtskala naturwissenschaftliche Kompetenz im internationalen Vergleich

Mittelwerte der naturwissenschaftlichen Kompetenz

Einen Überblick über die Mittelwerte und die Verteilungen der naturwissenschaftlichen Kompetenz in den OECD-Staaten gibt Abbildung 2.10 (vgl. auch Tabelle A1 im Anhang inklusive OECD-Partnerstaaten). Bei PISA 2015 liegt der über die OECD-Staaten errechnete Mittelwert bei 493 Punkten. Dieser dient als Bezugspunkt für die Aufteilung der OECD-Staaten in drei Gruppen: Im oberen Teil der Abbildung 2.10 (oberhalb der fett gedruckten Linie) finden sich Staaten, deren Schülerinnen und Schüler mittlere Kompetenzwerte signifikant über dem Durchschnitt aller OECD-Staaten erreichen. Zu dieser Gruppe gehört auch Deutschland mit einem Mittelwert von 509 Punkten. Staaten im mittleren Bereich der Tabelle unterscheiden sich in ihren Mittelwerten nicht signifikant vom OECD-Durchschnitt, während die Staaten im unteren Tabellenbereich (unterhalb der fett gedruckten Linie) signifikant niedrigere Mittelwerte in der naturwissenschaftlichen Kompetenz zeigen.

In PISA 2015 erreichen insgesamt 18 Staaten mittlere Kompetenzwerte, die signifikant über dem OECD-Durchschnitt liegen. Bei der Interpretation der Punktwerte ist zu beachten, dass die Darstellung als Rangfolge nur begrenzt aussagekräftig ist, da sich die Mittelwerte der Staaten teilweise nicht statistisch signifikant voneinander unterscheiden (vgl. OECD, 2016b). Sinnvoller ist es daher, Gruppierungen von Staaten zu betrachten, die statistisch nicht signifikant voneinander abweichende Mittelwerte aufweisen. Innerhalb der Gruppe von Staaten, die signifikant über dem OECD-Mittelwert liegen, bilden Japan (538 Punkte) und Estland (534 Punkte) die Spitzengruppierung. Es folgt eine weitere Konstellation mit Finnland (531 Punkte) und Kanada (528 Punkte). Die dritte Konstellation von Staaten beginnt mit Korea (516 Punkte) und endet mit Irland (503 Punkte), also bereits mit einem beträchtlichen Abstand zu Kanada. In dieser relativ homogenen Gruppierung von neun Staaten (Korea, Neuseeland, Slowenien, Australien, Vereinigtes Königreich, Deutschland, Niederlande, Schweiz und Irland) befindet sich auch Deutschland mit einem mittleren Kompetenzwert von 509 Punkten. Die Gruppe der Staaten, die Kompetenzmittelwerte signifikant über dem OECD-Durchschnitt erreichten, beschließen Portugal (501 Punkte) und Norwegen (498 Punkte). Der Unterschied zwischen dem leistungsstärksten und leistungsschwächsten OECD-Staat, die über dem OECD-Mittelwert liegen, beträgt damit 40 Punkte – ein Abstand, der in etwa dem Unterschied einer halben Kompetenzstufe entspricht.

Die Ergebnisse zeigen erneut, dass die naturwissenschaftliche Kompetenz von Schülerinnen und Schülern in Deutschland im internationalen Vergleich anschlussfähig ist. Den Abstand zu den Spitzenstaaten Japan, Estland, Finnland und Kanada zu verringern, bleibt jedoch eine klare Herausforderung. Betrachtet man zusätzlich die Ergebnisse der OECD-Partnerstaaten (vgl. Tabelle A1 im Anhang), so zeigt sich, dass gerade in den dort versammelten asiatischen Staaten ein noch höheres Niveau naturwissenschaftlicher Kompetenz erreicht wird. Singapur als kompetenzstärkstes Bildungssystem liegt mit 556 Punk-

© Waxmann Verlag GmbH

ten noch einmal 18 Punkte vor dem OECD-Spitzenreiter Japan. Die nachfolgenden asiatischen Staaten bilden eine relativ homogene Gruppe (Chinesisch Taipeh, Macau [China], Vietnam, Hongkong [China] sowie ein Zusammenschluss der vier Provinzen Peking, Shanghai, Jiangsu, Guangdong [China]), die zwischen 532 und 518 Punkte erreichen.

Die Gruppe der Staaten, deren Mittelwerte *nicht signifikant* vom OECD-Mittelwert verschieden sind, beginnt mit den Vereinigten Staaten (496 Punkte) und endet mit Lettland (490 Punkte). Außerdem finden sich in diesem Bereich Österreich, Frankreich, Schweden, die Tschechische Republik und Spanien. Kompetenzwerte signifikant *unterhalb* des OECD-Durchschnitts werden in zehn Staaten erreicht, unter anderem in Ungarn (477 Punkte), Island (473 Punkte), Israel (467 Punkte) und der Slowakischen Republik (461 Punkte). Das untere Ende der Kompetenzverteilung wird von Mexiko mit 416 Punkten markiert. Damit liegt zwischen dem kompetenzstärksten OECD-Staat Japan und Mexiko ein Unterschied von fast zwei Kompetenzstufen. Auch ein Großteil der OECD-Partnerstaaten befindet sich deutlich unterhalb des OECD-Durchschnitts (vgl. Tabelle A1 im Anhang). Schlusslichter sind hier der Kosovo (378 Punkte) und Algerien (376 Punkte) sowie mit noch einmal 44 Punkten weniger die Dominikanische Republik (332 Punkte).

Streuung der naturwissenschaftlichen Kompetenz

Der Kennwert der Standardabweichung zeigt die Streuung der naturwissenschaftlichen Kompetenz an. Er gibt an, wie homogen die Kompetenzverteilung innerhalb eines Staates ist. Die im internationalen Vergleich größten Streubreiten finden sich in Israel (SD = 106), Neuseeland (SD = 104) sowie in Australien, Frankreich und Schweden (jeweils SD = 102). Die Standardabweichung in Deutschland (SD = 99 Punkte) ist ähnlich wie in den Niederlanden (SD = 101), der Schweiz und dem Vereinigten Königreich (jeweils SD = 100). In all diesen genannten Staaten liegt die Standardabweichung signifikant über dem entsprechenden OECD-Mittel (SD = 94 Punkte). Bemerkenswert ist der Befund, dass in der Gruppe der leistungsstarken Staaten (signifikant über dem mittleren OECD-Kompetenzwert) mehrere Bildungssysteme eine deutlich geringere Leistungsstreuung aufweisen als Deutschland. Zu nennen sind hier vor allem Estland (SD = 89), Kanada (SD = 92) und Japan (SD = 93). In diesen Staaten scheint eine Breitenförderung im Sinne von *Naturwissenschaften für alle* (vgl. Fensham, 1985) auf hohem Niveau besser zu gelingen. In der Gruppe über dem OECD-Kompetenzmittelwert fallen weitere Staaten mit einer relativ homogenen Kompetenzverteilung auf, nämlich Irland, Dänemark, Polen und Portugal. Singapur als kompetenzstärkster Partnerstaat (außerhalb der OECD; vgl. Tabelle A1 im Anhang) zeigt hingegen eine eher heterogene Kompetenzverteilung (SD = 104).

Eine grafische Darstellung der Streuung liefern die Perzentilbänder in Abbildung 2.10 (die Punktwerte finden sich in Anhang A1) Die Perzentilpunktwerte teilen die Leistungsbandbreiten in den jeweiligen Stichproben in definierte Prozentbereiche auf. Der Punktwert für das 95. Perzentil kennzeichnet beispielsweise jenen Kompetenzwert, bei

© Waxmann Verlag GmbH

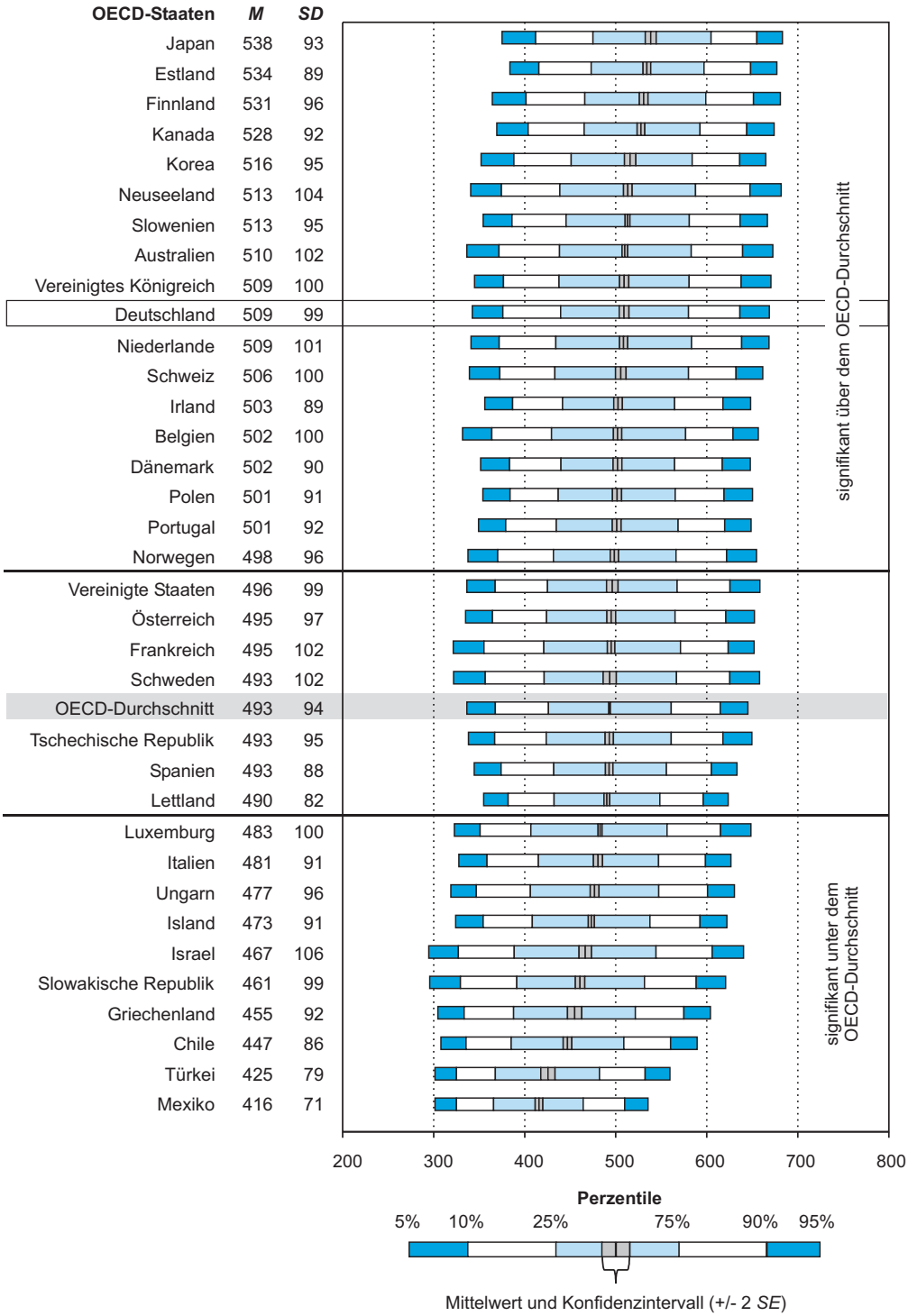


Abbildung 2.10: Mittelwerte, Streuungen und Perzentilbänder der Gesamtskala naturwissenschaftliche Kompetenz der OECD-Staaten

© Waxmann Verlag GmbH

dem die Gruppe der 5 Prozent leistungsstärksten Jugendlichen in einem Staat beginnt. Ein solcher Kennwert liefert Auskunft darüber, wie erfolgreich die Spitzenförderung in einem Staat gelingt. Betrachtet man die kompetenzstarken Staaten von Japan bis Irland, so bietet sich ein ähnliches Bild. Die höchsten Werte werden in Japan (hier beginnt die Spitzengruppe der kompetenzstärksten Jugendlichen bei 683 Punkten), Finnland (681 Punkte) und Neuseeland (682 Punkte) erreicht. Die Spitzengruppe der Jugendlichen in Deutschland beginnt bei 669 Punkten und erreicht damit ein etwas niedrigeres Niveau. Entsprechend dürfte bei der Förderung hochkompetenter Schülerinnen und Schüler in Deutschland das Potenzial noch nicht ausgeschöpft sein. Im unteren Bereich der Kompetenzverteilung, gegeben durch die Werte für das 5-Prozent-Perzentil, ist die Varianz hingegen deutlich größer. Dieser Perzentilpunktwert gibt den Testwert an, den 95 Prozent der Fünfzehnjährigen erreichen. Besonders die Spitzenstaaten Japan (375 Punkte) und Estland (384 Punkte), aber auch Irland (356 Punkte) zeigen hier Stärken. Ihre Kennwerte für das 5-Prozent-Perzentil sind relativ hoch und zeigen an, dass die leistungsschwächste Gruppe über ein vergleichsweise hohes Niveau naturwissenschaftlicher Kompetenz verfügt. Deutschland (342 Punkte) erreicht sehr ähnliche Punktwerte wie die in der Tabelle direkt benachbarten Staaten Niederlande (341 Punkte) und Vereinigtes Königreich (345 Punkte). Vergleichsweise niedrige Punktwerte für das 5-Prozent-Perzentil sind für Australien (336 Punkte) und Neuseeland (341 Punkte) zu beobachten. Betrachtet man die Ergebnisse für die Streuung und die Perzentilbänder gemeinsam, lässt sich feststellen, dass insbesondere in Japan und Estland eine erfolgreiche Breitenförderung stattfindet, die hochkompetente Jugendliche nicht vernachlässigt. In Deutschland scheinen weitere Anstrengungen nötig, um das Ziel einer bestmöglichen Förderung in den *Naturwissenschaften für alle* zu erreichen, das heißt eine Breitenförderung, ohne die Spitzenförderung zu vernachlässigen.

Verteilung auf Stufen der naturwissenschaftlichen Kompetenz

Die Kompetenzstufen können anhand der kognitiven Anforderungen der ihnen zugeordneten Aufgaben inhaltlich beschrieben werden und unterstützen damit eine kriteriumsorientierte Interpretation der im Test erzielten Kompetenzwerte (vgl. Abschnitt 2.2.3). Darüber hinaus erlaubt insbesondere die Betrachtung der Extrembereiche der Verteilung – also der prozentualen Anteile auf den Kompetenzstufen V und VI sowie auf und unter den Kompetenzstufen Ia und Ib – Aussagen zu den Anteilen besonders leistungsstarker bzw. leistungsschwacher Jugendlicher in den Naturwissenschaften. Während erstere einen potenziellen Nachwuchspool für Karrieren im naturwissenschaftlich-technischen Bereich stellen, ist bei letzteren zu fürchten, dass ihre Kompetenzen nicht ausreichen werden, den heutigen naturwissenschaftlichen und technischen Anforderungen in Beruf und Alltag zu begegnen. Schülerinnen und Schüler unterhalb der Kompetenzstufe II sind praktisch nicht in der Lage, auf grundlegendes, alltägliches Wissen zurückzugreifen, um einfache Daten zu interpretieren, Erklärungen zu identifizieren oder valide Schlussfolgerungen zu erkennen. Somit können diese Schülerinnen und

© Waxmann Verlag GmbH

Schüler weder aktuelle naturwissenschaftliche Themen verfolgen und bewerten noch die Konsequenzen bestimmter Handlungen einschätzen.

Abbildung 2.11 zeigt die Verteilung der Schülerinnen und Schüler auf die Stufen der naturwissenschaftlichen Kompetenz für die OECD-Staaten. Betrachtet man zunächst

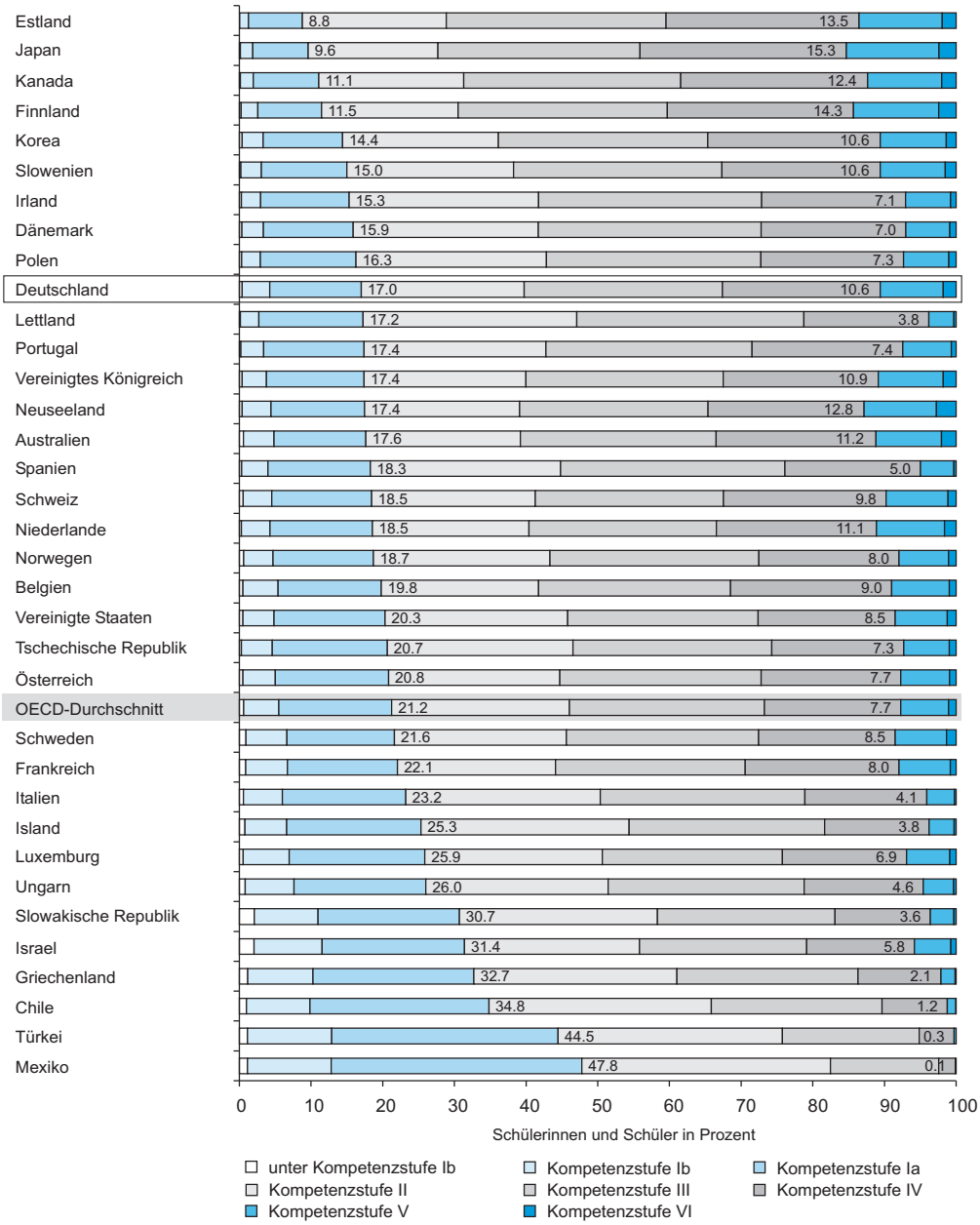


Abbildung 2.11: Prozentuale Anteile von Schülerinnen und Schülern auf Kompetenzstufe I oder darunter beziehungsweise auf Kompetenzstufe V oder VI für die Gesamtskala Naturwissenschaften

© Waxmann Verlag GmbH

den unteren Bereich der Kompetenzverteilung, so befinden sich im OECD-Durchschnitt 21.2 Prozent der Fünfzehnjährigen auf den Kompetenzstufen Ia und Ib bzw. darunter. In Deutschland ist dieser Anteil mit 17.0 Prozent signifikant niedriger. Für die von Deutschland auf Mittelwertebene nicht signifikant verschiedenen Staaten variiert der Anteil zwischen 14.4 Prozent in Korea und 18.5 Prozent in den Niederlanden. Korea weist damit – ebenso wie die vier Staaten der Spitzengruppe – einen signifikant niedrigeren Anteil kompetenzschwacher Jugendlicher auf. Besonders erfolgreich scheint die Förderung kompetenzschwacher Fünfzehnjähriger in Estland (8.8 Prozent) und Japan (9.6 Prozent), aber auch in Kanada (11.1 Prozent) und Finnland (11.5 Prozent) zu gelingen.

Im oberen Leistungsbereich, d. h. auf den Kompetenzstufen V und VI, befinden sich im OECD-Durchschnitt 7.7 Prozent der Jugendlichen. Der Unterschied zwischen Kompetenzstufe V und IV liegt im Wesentlichen in der Komplexität der Anforderungen (die komplexe Wissensbestände und tieferes Verstehen erfordern) sowie im Ausmaß der geforderten Transferleistung auf unvertraute Kontexte begründet. Schülerinnen und Schüler auf den Kompetenzstufen V und VI zeichnen sich besonders dadurch aus, dass ihnen ein ausgeprägtes *konzeptuelles*, *prozedurales* und *epistemisches Wissen* zur Verfügung steht. Dieses erlaubt ihnen, auf vielfältige, miteinander in Beziehung stehende naturwissenschaftliche Ideen und Konzepte zurückzugreifen, um unbekannte und komplexere Phänomene und Prozesse, die mehrere kausale Verknüpfungen beinhalten, zu erklären sowie unterschiedliche naturwissenschaftliche Herangehensweisen zu bewerten. Darüber hinaus können sie relevante von irrelevanten Informationen unterscheiden sowie evidenzbasierte Argumente und die Grenzen der Interpretierbarkeit von Daten erkennen (vgl. Abschnitt 2.2.3).

Verglichen mit dem OECD-Durchschnitt (7.7 Prozent) ist der Anteil an hochkompetenten Jugendlichen (Stufen V und IV) in Deutschland mit 10.6 Prozent signifikant größer. Betrachtet man erneut die Gruppe der Staaten, deren Mittelwerte der naturwissenschaftlichen Kompetenz sich von Deutschland nicht signifikant unterscheiden, so reichen die Anteile von 7.1 Prozent in Irland bis 12.8 Prozent in Neuseeland. Neben letzterem weisen lediglich die vier Spitzenstaaten signifikant höhere Anteile als Deutschland auf den oberen beiden Kompetenzstufen auf. Die höchsten Anteile hochkompetenter Fünfzehnjähriger beobachtet man in Japan mit 15.3 Prozent, gefolgt von Finnland mit 14.3 Prozent. Der Unterschied zu Estland (13.5 Prozent) ist allerdings nicht signifikant, lediglich in Kanada (12.4 Prozent) ist der Anteil signifikant niedriger als in Japan und Finnland.

Im internationalen Vergleich hat sich Deutschland damit im Kreise der Staaten, die naturwissenschaftlich-technische Nachwuchsförderung auf hohem Niveau betreiben, etabliert. Dennoch zeigen Bildungssysteme wie Japan oder Finnland, dass Deutschland gerade im Spitzenbereich noch Potenzial ausschöpfen könnte. Insbesondere die Fähigkeit, mit hochkomplexen Anforderungen in unvertrauten Kontexten umzugehen – ein Charakteristikum der Kompetenzstufen V und VI –, ist in der heutigen Zeit unverzichtbar und sollte als Anreiz für eine fortlaufende Förderung in diesem Bereich dienen.

© Waxmann Verlag GmbH

Geschlechterdifferenzen in der naturwissenschaftlichen Kompetenz

Im OECD-Durchschnitt zeigt sich bei PISA 2015 erstmals eine statistisch signifikante Differenz der naturwissenschaftlichen Kompetenz von vier Punkten zugunsten der männlichen Jugendlichen (Abbildung 2.12). Auch in einem Großteil der OECD-Staaten (15) erreichen Jungen signifikant höhere mittlere Kompetenzwerte als Mädchen. Dieser Befund gilt ebenfalls für Deutschland. Hier beträgt die Differenz 10 Punkte. In Japan, einem Staat der Spitzengruppe, sind die Jungen sogar 14 Punkte besser als die Mädchen. In vier Staaten (Finnland, Lettland, Griechenland und Slowenien) erreichen hingegen die Mädchen signifikant höhere mittlere Kompetenzwerte als die Jungen. In Finnland liegen die Leistungen der Mädchen zum Beispiel 19 Punkte (gerundet) höher als die der Jungen, in Slowenien sechs Punkte. In Estland und Kanada lassen sich keine signifikanten Geschlechterunterschiede beobachten, ebenso wenig in den anderen Bildungssystemen, die insgesamt ähnliche mittlere Kompetenzwerte aufweisen wie Deutschland (Korea, Neuseeland, Australien, Vereinigtes Königreich und die Niederlande). Diese in den OECD-Staaten beobachteten unterschiedlichen Ergebnisse weisen darauf hin, dass Geschlechterdifferenzen in den Naturwissenschaften nicht unabhängig von unterrichtlichen, schulischen und gesellschaftlichen Strukturen ausgeprägt sind. Für Deutschland stellt sich daher auch weiterhin die Aufgabe, die Ursachen der aufgezeigten Differenzen zu erkennen und Maßnahmen für eine Förderung von Jungen und Mädchen forschungsbasiert zu entwickeln und zu implementieren, die insbesondere das Ausschöpfen der Potenziale im Blick haben.

Betrachtet man die unteren Kompetenzstufen (Stufe Ia und darunter), so finden sich dort im OECD-Mittel signifikant weniger Mädchen (20.7 Prozent) als Jungen (21.8 Prozent). Ein entsprechendes Bild zeigt sich auch bei einem Großteil der OECD-Staaten. In der Spitzengruppe der OECD-Staaten sind diese Unterschiede für Estland, Finnland, Kanada und Korea statistisch signifikant. In Finnland und Korea sind sie mit über 6 Prozent Unterschied im OECD-Vergleich beträchtlich. In Japan hingegen findet sich kein bedeutsamer Unterschied des Anteils von Mädchen und Jungen auf Kompetenzstufe Ia und darunter. Neben Chile ist Deutschland eines der beiden Bildungssysteme, in denen der Anteil der Mädchen auf den unteren Kompetenzstufen (Stufe Ia und darunter) mit 18.1 Prozent signifikant größer ist als derjenige der Jungen mit 15.9 Prozent.

Betrachtet man die Schülerinnen und Schüler, die im Test besonders gut abgeschnitten und die Kompetenzstufe V und darüber erreicht haben, ist der Anteil der Jungen an dieser Gruppe in 26 OECD-Staaten statistisch signifikant höher als der der Mädchen. Auch in Deutschland ist der Anteil der Jungen, welche die Kompetenzstufen V und VI erreichen, mit 12.4 Prozent signifikant höher als der Anteil der Mädchen (8.7 Prozent). Besonders groß ist dieser Unterschied in Japan, hier sind 5.5 Prozent mehr Jungen als Mädchen in der Spitzengruppe vertreten. Einzig in Finnland machen die Mädchen den statistisch signifikant größeren Anteil an dieser Gruppe aus. Zusammenfassend lässt sich feststellen, dass die Mädchen in der Gruppe der kompetenzstärksten Jugendlichen in fast allen OECD-Staaten unterrepräsentiert sind, während sich in der Regel mehr Jungen als

© Waxmann Verlag GmbH

OECD-Staaten	Mädchen		Jungen		Differenz
	M	(*SE*)	*M*	(*SE*)	Jungen – Mädchen
Finnland	541	(2.6)	521	(2.7)	
Lettland	496	(2.2)	485	(2.0)	
Korea	521	(3.3)	511	(4.6)	
Griechenland	459	(3.9)	451	(4.6)	
Türkei	429	(4.4)	422	(4.5)	
Slowenien	516	(1.9)	510	(1.9)	
Schweden	496	(3.7)	491	(4.1)	
Island	475	(2.1)	472	(2.6)	
Slowakische Republik	461	(3.3)	460	(3.0)	
Vereinigtes Königreich	509	(3.3)	510	(2.9)	
Kanada	527	(2.3)	528	(2.5)	
Frankreich	494	(2.7)	496	(2.7)	
Australien	509	(1.7)	511	(2.1)	
Norwegen	497	(2.7)	500	(2.7)	
Ungarn	475	(2.9)	478	(3.4)	
Estland	533	(2.3)	536	(2.7)	
OECD-Durchschnitt	491	(0.5)	495	(0.5)	
Niederlande	507	(2.5)	511	(2.9)	
Israel	464	(4.1)	469	(4.7)	
Neuseeland	511	(2.7)	516	(3.2)	
Dänemark	499	(3.2)	505	(2.6)	
Schweiz	502	(3.5)	508	(3.1)	
Polen	498	(2.8)	504	(2.9)	
Spanien	489	(2.5)	496	(2.5)	
Vereinigte Staaten	493	(3.4)	500	(3.7)	
Luxemburg	479	(1.5)	487	(1.7)	
Mexiko	412	(2.3)	420	(2.6)	
Tschechische Republik	488	(2.5)	497	(3.3)	
Portugal	496	(2.6)	506	(2.9)	
Deutschland	504	(2.8)	514	(3.2)	
Irland	497	(2.6)	508	(3.2)	
Belgien	496	(2.7)	508	(3.1)	
Japan	532	(2.9)	545	(4.1)	
Chile	440	(2.7)	454	(3.1)	
Italien	472	(3.6)	489	(3.1)	
Österreich	486	(3.1)	504	(3.6)	

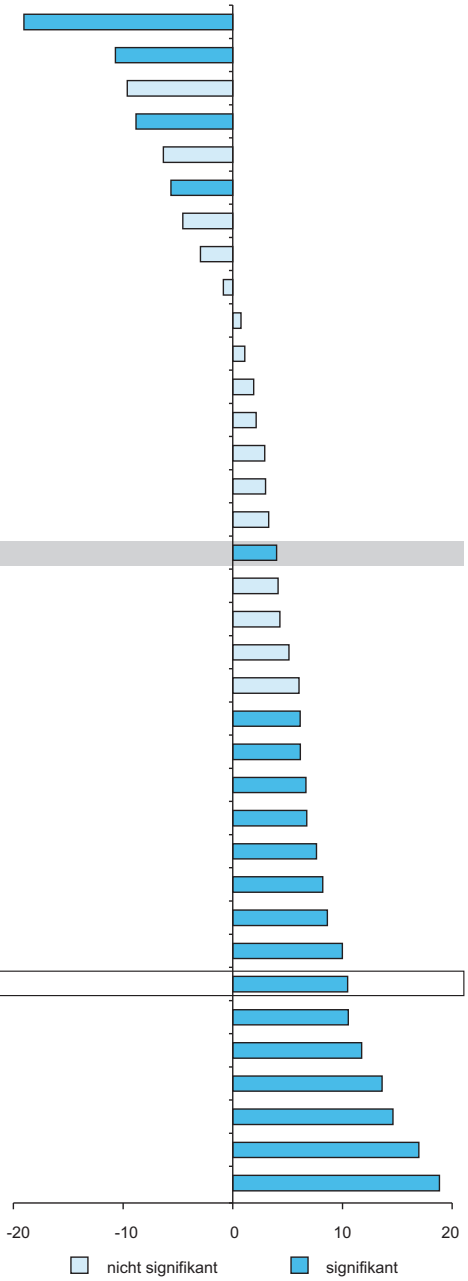


Abbildung 2.12: Mittelwerte naturwissenschaftlicher Kompetenz von Jungen und Mädchen in den OECD-Staaten

© Waxmann Verlag GmbH

Mädchen auf den unteren Kompetenzstufen befinden. In Deutschland finden sich hingegen auf den unteren Kompetenzstufen mehr Mädchen als Jungen. Die Notwendigkeit einer differenziellen Förderung von Mädchen im unteren und oberen Leistungsbereich wird damit auch von diesen Ergebnissen unterstützt.

2.3.2 Naturwissenschaftliche Teilkompetenzen, Wissensbereiche und Wissenssysteme im internationalen Vergleich

Aufgrund des Status als Hauptdomäne konnten bei PISA 2015 auch die drei Teilkompetenzen, die beiden Wissensbereiche sowie die Wissenssysteme der naturwissenschaftlichen Kompetenz jeweils auf einer eigenen Teilskala abgebildet werden. Die Ergebnisse dieser Teilskalen erlauben einen differenzierten Blick auf die relativen Stärken und Schwächen der Jugendlichen in den Naturwissenschaften. Die drei Teilkompetenzen *Phänomene naturwissenschaftlich erklären*, *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* und *Daten und Evidenz naturwissenschaftlich interpretieren* als Kern der Definition naturwissenschaftlicher Kompetenz bei PISA 2015 werden im Folgenden ausführlicher – auch unter Berücksichtigung von Geschlechterunterschieden – berichtet. Für die Wissensbereiche und Wissenssysteme werden ausgewählte Ergebnisse dargestellt. Für eine ausführliche Darstellung sei auf den internationalen OECD-Bericht verwiesen (OECD, 2016b).

Mittelwerte, Streuungen und Geschlechterdifferenzen der naturwissenschaftlichen Teilkompetenzen

Die Mittelwerte, Streuungen und Geschlechterunterschiede für die drei naturwissenschaftlichen Teilkompetenzen sind in den Abbildungen 2.13, 2.14 und 2.15 dargestellt (vgl. auch die Tabellen 2 bis 4 im Anhang A für die OECD-Partnerstaaten). Die mittleren Kompetenzwerte der Staaten wie auch die Rangfolgen unterscheiden sich nicht wesentlich von der Gesamtskala Naturwissenschaften. Trotzdem gibt es einige interessante Unterschiede.

Für Deutschland liegen die Mittelwerte für alle drei Teilkompetenzen signifikant über dem OECD-Durchschnitt. Den höchsten Mittelwert erreichen die Fünfzehnjährigen in Deutschland in der Teilkompetenz *Phänomene naturwissenschaftlich erklären* (511 Punkte), den niedrigsten in der Teilkompetenz *naturwissenschaftliche Forschung bewerten und Untersuchungen planen*. Letzte stellt eine relative Schwäche der Schülerinnen und Schüler in Deutschland dar. Der Mittelwert von 506 Punkten ist signifikant niedriger als der Mittelwert für die Teilkompetenz *Phänomene naturwissenschaftlich erklären*. Auch die Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren* (507 Punkte) ist in Deutschland signifikant geringer ausgeprägt als die Teilkompetenz *Phänomene naturwissenschaftlich erklären*. Insgesamt haben auch die Staaten der Spitzengruppe – ähnlich wie in Deutschland – relative Stärken und Schwächen in den drei

© Waxmann Verlag GmbH

OECD-Staaten	Gesamt			Mädchen		Jungen	
	M	*(SE)*	*SD*	*M*	*(SE)*	*M*	*(SE)*
Japan	539	(3.3)	97	526	(3.3)	551	(4.5)
Finnland	534	(2.4)	95	540	(2.9)	527	(2.7)
Estland	533	(2.0)	90	526	(2.2)	540	(2.7)
Kanada	530	(2.1)	93	525	(2.4)	535	(2.6)
Slowenien	515	(1.5)	96	512	(2.4)	518	(1.9)
Neuseeland	511	(2.6)	106	503	(2.9)	519	(3.5)
Deutschland	511	(2.8)	100	502	(2.8)	520	(3.4)
Australien	510	(1.6)	105	505	(1.8)	516	(2.2)
Korea	510	(3.4)	98	509	(3.6)	511	(4.9)
Vereinigtes Königreich	509	(2.7)	103	505	(3.6)	513	(3.0)
Niederlande	509	(2.5)	99	503	(2.7)	516	(3.1)
Irland	505	(2.5)	91	497	(2.9)	514	(3.1)
Schweiz	505	(3.1)	102	497	(3.7)	512	(3.3)
Norwegen	502	(2.3)	99	497	(2.9)	507	(2.9)
Dänemark	502	(2.7)	94	495	(3.4)	509	(3.0)
Polen	501	(2.8)	95	492	(3.1)	509	(3.3)
Österreich	499	(2.7)	101	484	(3.3)	514	(3.8)
Belgien	499	(2.4)	102	488	(2.8)	509	(3.3)
Portugal	498	(2.5)	94	489	(2.7)	507	(3.0)
Schweden	498	(3.7)	104	497	(3.9)	499	(4.2)
Tschechische Republik	496	(2.5)	97	487	(2.8)	505	(3.3)
Spanien	494	(2.2)	92	486	(2.5)	502	(2.6)
OECD-Durchschnitt	493	(0.5)	97	487	(0.5)	499	(0.6)
Vereinigte Staaten	492	(3.4)	102	484	(3.6)	499	(3.8)
Frankreich	488	(2.2)	102	482	(2.8)	494	(2.9)
Lettland	488	(1.8)	84	490	(2.7)	486	(2.0)
Luxemburg	482	(1.1)	105	474	(1.4)	490	(1.7)
Italien	481	(2.7)	96	468	(3.7)	493	(3.6)
Ungarn	478	(2.5)	98	471	(3.0)	485	(3.4)
Island	468	(2.0)	93	464	(2.4)	472	(2.7)
Slowakische Republik	464	(2.7)	102	460	(3.4)	468	(3.2)
Israel	463	(3.5)	109	457	(4.3)	468	(4.7)
Griechenland	454	(3.9)	94	457	(3.9)	451	(4.7)
Chile	446	(2.6)	90	436	(3.0)	457	(3.3)
Türkei	426	(4.2)	84	425	(4.7)	427	(4.7)
Mexiko	414	(2.3)	74	405	(2.4)	422	(2.7)

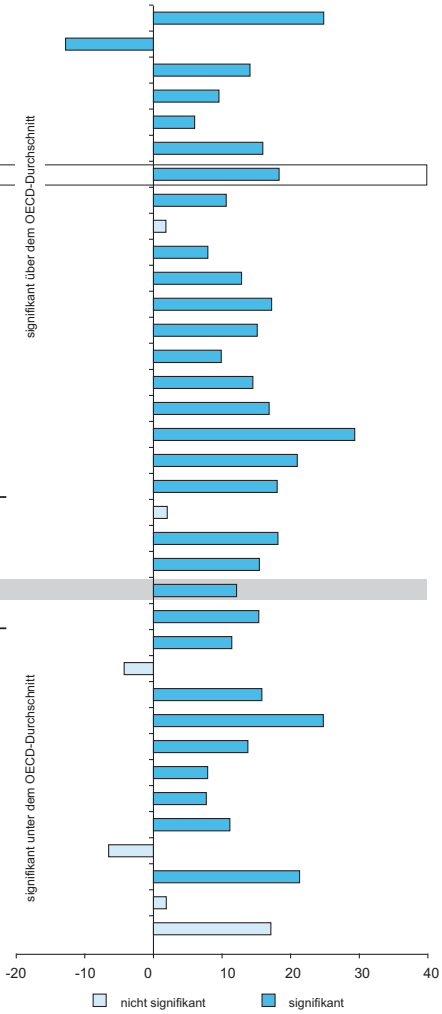


Abbildung 2.13: Mittelwerte und Streuungen der Teilskala *Phänomene naturwissenschaftlich erklären* in den OECD-Staaten

Teilkompetenzen. Relative Stärken in der Teilkompetenz *Phänomene naturwissenschaftlich erklären* haben beispielsweise die Schülerinnen und Schüler in Finnland, Slowenien und Irland. In der Teilkompetenz *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* hat Neuseeland eine relative Stärke und in der Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren* Korea. In letzterem Kompetenzbereich weisen die Schülerinnen und Schüler in Kanada und den Niederlanden hingegen eine relative Schwäche auf. Gleiches gilt für die Teilkompetenz *Phänomene naturwissenschaftlich erklären* für Korea sowie für die Teilkompetenz *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* für Japan. Besonders große Unterschiede zwischen

© Waxmann Verlag GmbH

OECD-Staaten	Gesamt			Mädchen		Jungen		Differenz
	M	*(SE)*	*SD*	*M*	*(SE)*	*M*	*(SE)*	Jungen – Mädchen
Japan	536	(3.3)	98	535	(3.5)	538	(4.6)	
Estland	535	(2.6)	93	539	(2.5)	530	(3.8)	
Kanada	530	(2.7)	102	535	(3.1)	525	(3.0)	
Finnland	529	(2.9)	102	543	(2.9)	517	(3.7)	
Neuseeland	517	(3.1)	114	518	(3.4)	516	(4.1)	
Korea	515	(3.3)	100	525	(3.6)	507	(4.9)	
Australien	512	(2.0)	111	517	(2.5)	508	(2.6)	
Slowenien	511	(2.0)	99	521	(2.8)	503	(2.3)	
Niederlande	511	(2.5)	109	513	(2.8)	509	(3.4)	
Vereinigtes Königreich	508	(2.8)	107	511	(3.6)	504	(3.5)	
Belgien	507	(2.5)	105	506	(3.2)	508	(3.5)	
Schweiz	507	(3.5)	108	509	(4.4)	505	(3.7)	
Deutschland	506	(2.9)	103	505	(3.0)	507	(3.6)	
Dänemark	504	(2.6)	95	506	(3.4)	503	(3.0)	
Vereinigte Staaten	503	(3.6)	105	504	(4.0)	501	(4.2)	
Portugal	502	(2.7)	98	502	(2.9)	503	(3.3)	
Polen	502	(3.0)	94	503	(3.3)	500	(3.5)	
Irland	500	(2.6)	95	499	(3.1)	500	(3.5)	
Frankreich	498	(2.5)	109	502	(3.3)	493	(3.4)	
Norwegen	493	(2.6)	99	495	(3.5)	491	(3.1)	
OECD-Durchschnitt	493	(0.5)	100	495	(0.6)	490	(0.6)	
Schweden	491	(4.0)	108	497	(4.5)	485	(4.7)	
Lettland	489	(2.0)	86	499	(2.8)	480	(2.5)	
Spanien	489	(2.7)	91	490	(3.1)	488	(3.2)	
Österreich	488	(2.6)	102	485	(3.5)	492	(4.0)	
Tschechische Republik	486	(2.8)	96	486	(3.4)	487	(3.5)	
Luxemburg	479	(1.7)	105	479	(2.2)	479	(2.5)	
Italien	477	(2.7)	98	472	(4.0)	482	(3.6)	
Island	476	(2.5)	94	483	(3.3)	469	(3.0)	
Ungarn	474	(2.8)	101	476	(3.3)	472	(4.1)	
Israel	471	(3.8)	113	473	(4.5)	470	(5.4)	
Slowakische Republik	457	(3.2)	105	463	(4.1)	451	(3.4)	
Griechenland	453	(4.2)	99	460	(4.6)	448	(4.9)	
Chile	443	(2.9)	93	441	(3.2)	446	(4.0)	
Türkei	428	(4.0)	83	436	(4.5)	419	(4.6)	
Mexiko	415	(2.9)	83	415	(3.1)	414	(3.8)	

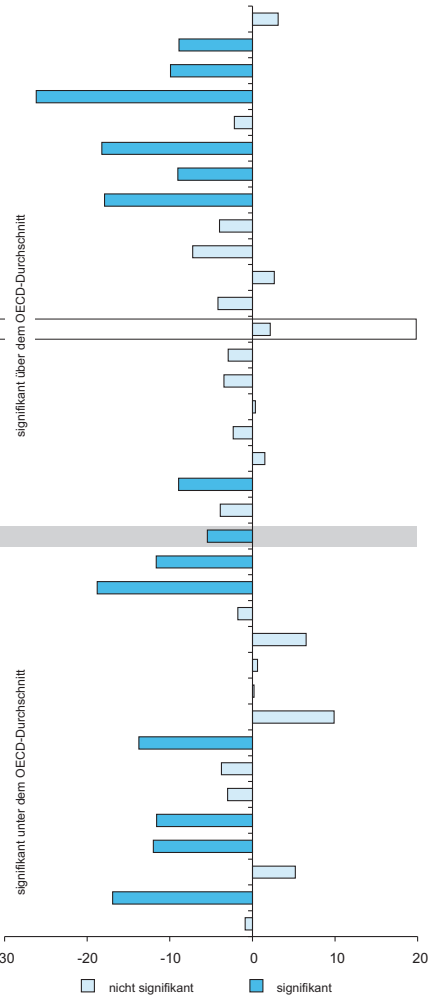


Abbildung 2.14: Mittelwerte und Streuungen der Teilskala *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* in den OECD-Staaten

den Teilkompetenzen sind in der Spitzengruppe für Korea zu beobachten, wohingegen in Estland keine signifikanten Unterschiede zwischen den drei Teilkompetenzen zu beobachten sind. Auch in Australien, dem Vereinigten Königreich und der Schweiz gibt es keine signifikanten Unterschiede.

Zusammenfassend zeigt sich, dass auch in der Spitzengruppe und der Gruppe mit ähnlicher naturwissenschaftlicher Kompetenz wie Deutschland eine gleichmäßige Förderung und Entwicklung aller drei Teilkompetenzen nur in einigen Staaten vergleichsweise gut gelingt. Insbesondere gilt dies für Estland, Australien, das Vereinigte Königreich und die Schweiz. In den anderen Staaten sind, ähnlich wie in Deutschland, relative Stärken

© Waxmann Verlag GmbH

OECD-Staaten	Gesamt			Mädchen		Jungen	
	M	*(SE)*	*SD*	*M*	*(SE)*	*M*	*(SE)*
Japan	541	(3.1)	97	536	(3.1)	545	(4.4)
Estland	537	(2.7)	95	538	(2.7)	535	(3.6)
Finnland	529	(2.8)	104	541	(3.3)	518	(3.1)
Kanada	525	(2.7)	97	525	(2.8)	525	(3.2)
Korea	523	(3.2)	100	533	(3.4)	514	(4.7)
Neuseeland	512	(2.5)	104	514	(2.8)	511	(3.3)
Slowenien	512	(2.0)	102	518	(2.6)	507	(2.5)
Vereinigtes Königreich	509	(2.9)	103	510	(3.7)	508	(3.4)
Deutschland	509	(3.0)	105	506	(3.2)	511	(3.5)
Australien	508	(1.8)	104	509	(2.2)	508	(2.4)
Niederlande	506	(2.5)	109	507	(2.8)	505	(3.2)
Schweiz	506	(3.0)	102	504	(3.6)	507	(3.3)
Belgien	503	(2.5)	103	499	(3.1)	507	(3.4)
Portugal	503	(2.6)	96	500	(3.0)	505	(3.2)
Polen	501	(2.6)	91	503	(3.0)	500	(3.2)
Frankreich	501	(2.5)	107	503	(3.1)	499	(3.1)
Irland	500	(2.7)	91	497	(3.0)	504	(3.6)
Dänemark	500	(2.6)	93	499	(3.3)	500	(3.2)
Norwegen	498	(2.8)	101	499	(3.1)	497	(3.3)
Vereinigte Staaten	497	(3.5)	99	495	(3.7)	498	(4.1)
Lettland	494	(1.7)	88	501	(2.4)	486	(2.1)
OECD-Durchschnitt	493	(0.5)	98	494	(0.5)	493	(0.6)
Tschechische Republik	493	(2.8)	102	491	(3.1)	495	(3.8)
Spanien	493	(2.4)	93	492	(2.7)	494	(2.9)
Österreich	493	(2.6)	101	486	(3.4)	500	(4.0)
Schweden	490	(3.7)	107	493	(3.8)	487	(4.3)
Luxemburg	486	(1.8)	100	485	(1.9)	487	(2.7)
Italien	482	(2.9)	97	476	(3.9)	487	(3.7)
Island	478	(2.1)	97	481	(2.5)	474	(3.1)
Ungarn	476	(2.7)	102	479	(3.4)	473	(3.7)
Israel	467	(3.7)	109	467	(4.4)	467	(5.2)
Slowakische Republik	459	(2.9)	103	462	(3.6)	456	(3.3)
Griechenland	454	(4.1)	95	460	(3.9)	449	(4.9)
Chile	447	(2.7)	89	441	(3.0)	453	(3.6)
Türkei	423	(4.2)	83	428	(4.6)	418	(5.1)
Mexiko	415	(2.3)	74	413	(2.6)	416	(2.6)

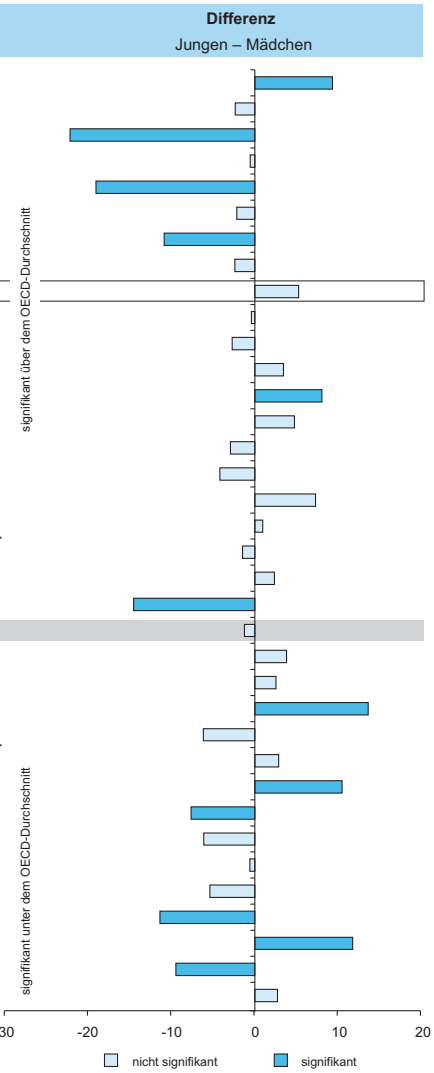


Abbildung 2.15: Mittelwerte und Streuungen der Teilskala *Daten und Evidenz naturwissenschaftlich interpretieren* in den OECD-Staaten

und Schwächen zu beobachten. Dies liefert Hinweise auf unterschiedliche Unterrichtstraditionen (vgl. Kapitel 4) und weist auf Verbesserungspotenziale im naturwissenschaftlichen Unterricht hin. Für ein besseres Verständnis der Unterschiede wären weiterführende Untersuchungen zu Unterrichtstraditionen und ggf. besonderen Schwerpunkten in der Erarbeitung einzelner Teilkomponenten erforderlich. Um zur Spitzengruppe der OECD-Staaten im Gesamtmittelwert der naturwissenschaftlichen Kompetenz Japan, Estland, Finnland und Kanada aufzuschließen, sollten in Deutschland alle Teilkompetenzen weiter gefördert werden. Ein besonderes Augenmerk sollte dabei auf der Förderung der Fähigkeit der Schülerinnen und Schüler *naturwissenschaftliche Forschung zu bewerten*

© Waxmann Verlag GmbH

und Untersuchungen zu planen sowie *Daten und Evidenz naturwissenschaftlich zu interpretieren* liegen – Teilkompetenzen, die sich auf den Prozess der naturwissenschaftlichen Erkenntnisgewinnung und den Umgang mit Evidenz beziehen.

Geschlechterdifferenzen in den naturwissenschaftlichen Teilkompetenzen

Die Unterschiede, die sich in der Gesamtskala naturwissenschaftliche Kompetenz zugunsten der Jungen zeigen, können mit Blick auf Geschlechterunterschiede in den Teilkompetenzen genauer betrachtet werden. Dabei zeigt sich, dass im OECD-Durchschnitt die Jungen vor allem besser *Phänomene naturwissenschaftlich erklären* können (zwölf Punkte Differenz zwischen Jungen und Mädchen). Dies ist in fast allen OECD-Staaten der Fall. Die größten Differenzen zugunsten der Jungen finden sich in Österreich (29 Punkte, gerundet) und Italien (25 Punkte), aber auch im Spitzenstaat Japan erreichen die Jungen in der Teilkompetenz *Phänomene naturwissenschaftlich erklären* im Mittel 25 Punkte mehr als die Mädchen.

Auch in Deutschland haben die Jungen in dieser Teilkompetenz mit einem Mittelwert von 520 Punkten einen signifikanten Kompetenzvorsprung von 18 Punkten gegenüber den Mädchen. Einzig in Finnland erreichen Mädchen (540 Punkte) in der Teilkompetenz *Phänomene naturwissenschaftlich erklären* signifikant höhere Mittelwerte als Jungen (527 Punkte). Damit hat Finnland auch das einzige Bildungssystem, in dem die Mädchen in allen drei Teilkompetenzen bessere Ergebnisse erzielen als die Jungen. In der Teilkompetenz *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* erreichen die Mädchen in zwölf weiteren OECD-Staaten bessere Ergebnisse. Dieser Vorteil der Mädchen ist bei den Staaten oberhalb des OECD-Durchschnitts in Finnland (mit einem Vorsprung der Mädchen von 26 Punkten), Slowenien und Korea besonders stark ausgeprägt, aber auch in Estland, Kanada und Australien zu finden. In Deutschland gibt es in dieser Teilkompetenz sowie auch in der Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren* keine signifikanten Unterschiede zwischen Mädchen und Jungen.

Für die Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren* zeigt sich in den OECD-Staaten in Bezug auf Geschlechterunterschiede ein vergleichsweise heterogenes Bild. In vielen Bildungssystemen sind hier ebenfalls keine Geschlechterunterschiede zu finden. Unter den Staaten, deren Kompetenzmittelwert in der Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren* signifikant über dem OECD-Durchschnitt liegt, erzielen beispielsweise in Japan und Belgien die Jungen bessere Ergebnisse, wohingegen in Finnland, Korea und Slowenien die Mädchen signifikant besser abschneiden. Diese Unterschiede zwischen den Staaten weisen darauf hin, dass Geschlechterdifferenzen nicht „naturgegeben" und daher nicht unvermeidbar sind. Eine gezielte Förderung kann helfen, sie zu verringern. In Deutschland, wie in den meisten anderen OECD-Staaten auch, gibt es vor allem in der Teilkompetenz *Phänomene naturwissenschaftlich erklären* differenziellen Förderbedarf für die Mädchen.

© Waxmann Verlag GmbH

Ergebnisse für die naturwissenschaftlichen Wissensbereiche und Wissenssysteme

Betrachtet man die Ergebnisse für die zwei Teilskalen der naturwissenschaftlichen Wissensbereiche *konzeptuelles Wissen* sowie *prozedural-epistemisches Wissen,* so zeigt sich im OECD-Durchschnitt kein signifikanter Unterschied (vgl. OECD, 2016b). Betrachtet man die Staaten, die sich durch eine vergleichbare naturwissenschaftliche Kompetenz wie Deutschland auszeichnen, erreichen Jugendliche in Irland und Slowenien höhere Kompetenzwerte im Wissensbereich *konzeptuelles Wissen* (jeweils vier Punkte Differenz). Unter den Spitzenstaaten ist in Finnland diese Differenz mit zehn Punkten besonders groß. Auch in Deutschland ist das *konzeptuelle Wissen* mit 512 Punkten stärker ausgeprägt als das *prozedural-epistemische Wissen* (507 Punkte). Beide Kompetenzmittelwerte liegen statistisch signifikant über dem OECD-Durchschnitt. Betrachtet man zusätzlich Geschlechterunterschiede zwischen Mädchen und Jungen, so haben Jungen in Deutschland einen deutlichen Vorsprung im *konzeptuellen Wissen* von 20 Punkten gegenüber den Mädchen (Mädchen: 502 Punkte, Jungen: 522 Punkte). Ähnlich große Differenzen zeigen sich in Japan (23 Punkte Differenz), Irland und der Schweiz (jeweils 17 Punkte Differenz). Für den Wissensbereich *prozedural-epistemisches Wissen* finden sich in Deutschland hingegen keine signifikanten Geschlechterunterschiede (Mädchen: 505 Punkte, Jungen: 508 Punkte).

Betrachtet man die drei Wissenssysteme *physikalische Systeme*, *lebende Systeme* sowie *Erd- und Weltraumsysteme,* zeigt sich, dass es gerade in den Spitzenstaaten Japan und Kanada gelingt, die naturwissenschaftliche Kompetenz in allen drei Wissenssystemen gleich gut zu fördern (vgl. OECD, 2016b). In Deutschland erzielen die Jugendlichen leicht höhere Kompetenzwerte in den *Erd- und Weltraumsystemen* (512 Punkte) im Vergleich zu den *physikalischen Systemen* (505 Punkte). Der Unterschied zu den *lebenden Systemen* (509 Punkte) ist nicht signifikant. Auch hier sind Geschlechterunterschiede zu finden. Bei den *physikalischen Systemen* erzielen Jungen im OECD-Durchschnitt einen signifikanten Vorsprung von neun Punkten. Im Spitzenland Estland haben die Jungen zehn Punkte Vorsprung vor den Mädchen. In Japan sind es sogar 18 Punkte. Auch in Deutschland erreichen die Jungen (514 Punkte) im Mittel 17 Punkte mehr als die Mädchen (497 Punkte). Keine signifikanten Unterschiede in der Gruppe von Staaten mit einer vergleichbaren naturwissenschaftlichen Kompetenz wie Deutschland lassen sich in Korea, Slowenien und dem Vereinigten Königreich finden. Für die *lebenden Systeme* zeigen sich in den meisten OECD-Staaten keine signifikanten Geschlechterunterschiede. Dies trifft auch auf Deutschland zu (Mädchen: 506 Punkte, Jungen: 512 Punkte). In sechs Staaten (Belgien, Chile, Irland, Italien, Japan und Portugal) erzielen die Jungen im Schnitt zehn Punkte mehr. In sieben Staaten (Finnland, Griechenland, Island, Korea, Lettland, Slowenien und Türkei) haben die Mädchen einen durchschnittlichen Vorsprung von vierzehn Punkten. Bei den *Erd- und Weltraumsystemen* erreichen in vierzehn Staaten die Jungen höhere Kompetenzwerte als die Mädchen. Auch in Deutschland erzielen die Jungen (518 Punkte) im Mittel zwölf Punkte mehr als die Mädchen (506 Punkte). Unter den Spitzenstaaten und Staaten mit einer vergleichbaren

© Waxmann Verlag GmbH

naturwissenschaftlichen Kompetenz wie Deutschland lässt sich ein solcher Unterschied, wenn auch auf einem deutlich höheren Niveau, nur noch in Japan finden (Mädchen: 534 Punkte, Jungen: 548 Punkte). Schaut man sich für Deutschland die drei Wissenssysteme im Vergleich an, so zeigt sich, dass Mädchen eine Schwäche im Bereich *physikalische Systeme* haben, während Jungen eine Stärke im Bereich *Erd- und Weltraumsysteme* haben.

Zusammenfassend lässt sich für Deutschland feststellen, dass die Fünfzehnjährigen in allen Teilkompetenzen, Wissensbereichen und Wissenssystemen Kompetenzwerte über dem OECD-Durchschnitt erreichen. Um Anschluss an die Spitzenstaaten zu erreichen, sollten *alle* Teilkompetenzen noch stärker im naturwissenschaftlichen Unterricht gefördert werden. Ein besonderes Augenmerk sollte auf die Teilkompetenzen, die sich auf den Prozess der Erkenntnisgewinnung und den Umgang mit Evidenz beziehen (*naturwissenschaftliche Forschung bewerten und Untersuchungen planen* und *Daten und Evidenz naturwissenschaftlich interpretieren*), gelegt werden. Differenzieller Förderbedarf zeichnet sich für die Mädchen in der Teilkompetenz *Phänomene naturwissenschaftlich erklären* ab. Eine besondere Herausforderung bleibt außerdem die gezielte Förderung von Mädchen im Bereich *physikalische Systeme.* Hier könnten Vorschläge für Förderansätze (u. a. Häußler & Hoffman, 1998; Häußler, Hoffman, Langeheine, Rost & Sievers, 1998; Hannover & Kessels, 2004; Kessels & Hannover, 2007) genutzt werden, um das Ziel einer Verringerung von Geschlechterunterschieden in den naturwissenschaftlichen Kompetenzen zu erreichen.

2.3.3 Schulartspezifische Analysen der naturwissenschaftlichen Kompetenz in Deutschland

Neben den bisher vorgestellten Ergebnissen zur naturwissenschaftlichen Kompetenz fünfzehnjähriger Schülerinnen und Schüler in Deutschland im internationalen Vergleich erlaubt die PISA-Stichprobe auch vertiefende Analysen aus nationaler Perspektive. Im Folgenden werden Befunde zu Unterschieden zwischen den Schularten in Deutschland vorgestellt. Im Gegensatz zu früheren Erhebungsrunden (Prenzel et al., 2007; Schiepe-Tiska et al., 2013) wird dabei nur zwischen dem Gymnasium und nicht gymnasialen Schularten unterschieden. In den letzten Jahren wurden zahlreiche Schulen mit einer Sekundarstufe I – mit Ausnahme der Gymnasien – durch die Kombination von Bildungsgängen in neue oder neu bezeichnete Schularten überführt (Pant, Stanat, Pöhlmann et al., 2013). Eine differenzierte Betrachtung der nicht gymnasialen, allgemeinbildenden Schularten ist daher nicht mehr tragfähig und sinnvoll. Bei PISA 2015 betrug die Bildungsbeteiligung am Gymnasium 33 Prozent, an den nicht gymnasialen Schularten 62 Prozent. Die verbleibenden 5 Prozent verteilen sich auf Schülerinnen und Schüler an Berufs- und Förderschulen, die aufgrund der geringen Stichprobengröße nicht gesondert betrachtet werden.

© Waxmann Verlag GmbH

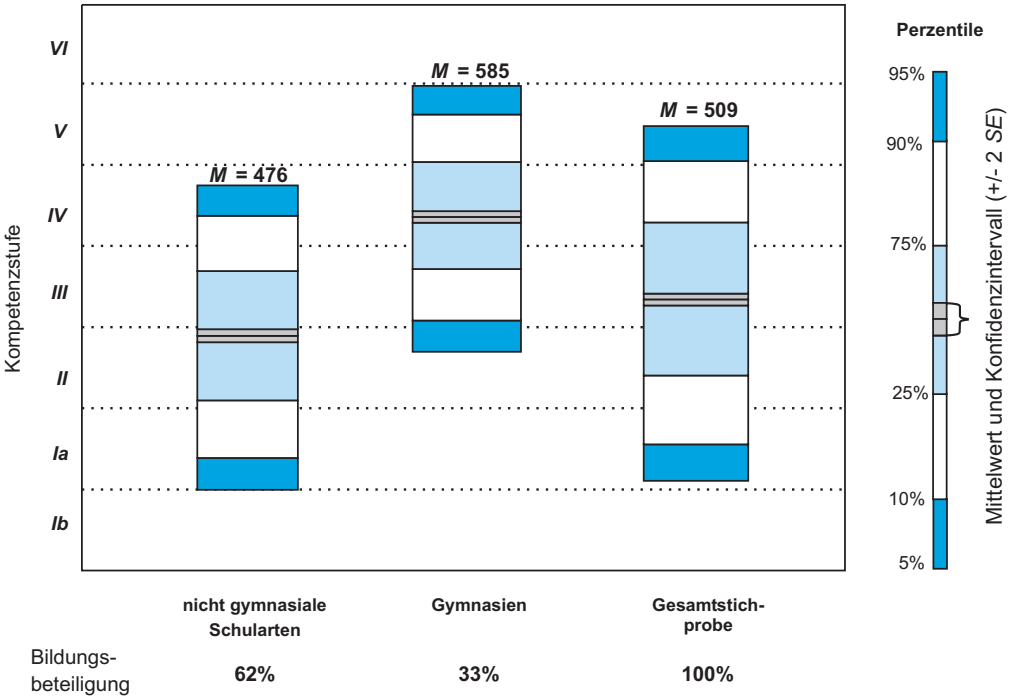


Abbildung 2.16: Perzentilbänder naturwissenschaftlicher Kompetenz in Deutschland nach Schulart

Mittelwerte und Streuung der naturwissenschaftlichen Kompetenz

Abbildung 2.16 zeigt die Mittelwerte und Streuungen der naturwissenschaftlichen Kompetenz für die Gesamtskala Naturwissenschaften. Erwartungsgemäß unterscheiden sich die erreichten Mittelwerte zwischen den Schularten beträchtlich. Während die Schülerinnen und Schüler am Gymnasium einen Mittelwert von 585 Punkten erreichen, liegt der Mittelwert der nicht gymnasialen Schularten mit 476 Punkten mehr als eine Standardabweichung bzw. nahezu eineinhalb Kompetenzstufen niedriger.

Neben Unterschieden in den Mittelwerten zeigen sich auch Unterschiede in der Streuung. Für die nicht gymnasialen Schularten (*SD* = 86 Punkte), aber auch innerhalb der Gymnasien (*SD* = 75 Punkte) ist eine deutliche Streuung der naturwissenschaftlichen Kompetenz zu beobachten. Dieser Befund erscheint insbesondere für die Gruppe der nicht gymnasialen Schularten (aufgrund der Mischung unterschiedlicher Bildungsgänge) nachvollziehbar. Die Streuung am Gymnasium lässt sich nicht zwangsläufig – wie man vermuten könnte – auf eine zunehmende Heterogenität dieser Schulform zurückführen. Wie in Kapitel 5 gezeigt wird, lassen sich lediglich zwei Prozent der Varianz in der naturwissenschaftlichen Kompetenz auf Unterschiede zwischen Schulen zurückführen. In Bezug auf die Domäne Mathematik hatte sich drei Jahre zuvor, in PISA 2012, gezeigt, dass 6 Prozent der Varianz zwischen den Gymnasien zu finden sind (Sälzer, Reiss, Schiepe-Tiska, Prenzel & Heinze, 2013, S. 162). Ein möglicher Grund für die Unterschiede

© Waxmann Verlag GmbH

könnten Leistungsunterschiede in den Naturwissenschaften zwischen den Bundesländern in Deutschland sein, wie sie etwa im letzten Ländervergleich festgestellt wurden (vgl. Pant et al., 2013).

Aus Abbildung 2.16 lassen sich auch Überlappungen der Kompetenzverteilungen an den Gymnasien und den nicht gymnasialen Schularten ablesen. Der mittlere Leistungsbereich für die nicht gymnasialen Schularten liegt auf den Kompetenzstufen II und III. Jugendliche am Gymnasium erreichen vorwiegend die Kompetenzstufe IV. Wie in früheren Erhebungen erreichen jedoch circa 10 Prozent der Schülerinnen und Schüler nicht gymnasialer Schularten ein Kompetenzniveau, das über dem Mittelwert der Gymnasien liegt. Umgekehrt erreichen am Gymnasium aber auch circa 10 Prozent der Fünfzehnjährigen ein Kompetenzniveau unterhalb des Mittelwertes der nicht gymnasialen Schularten.

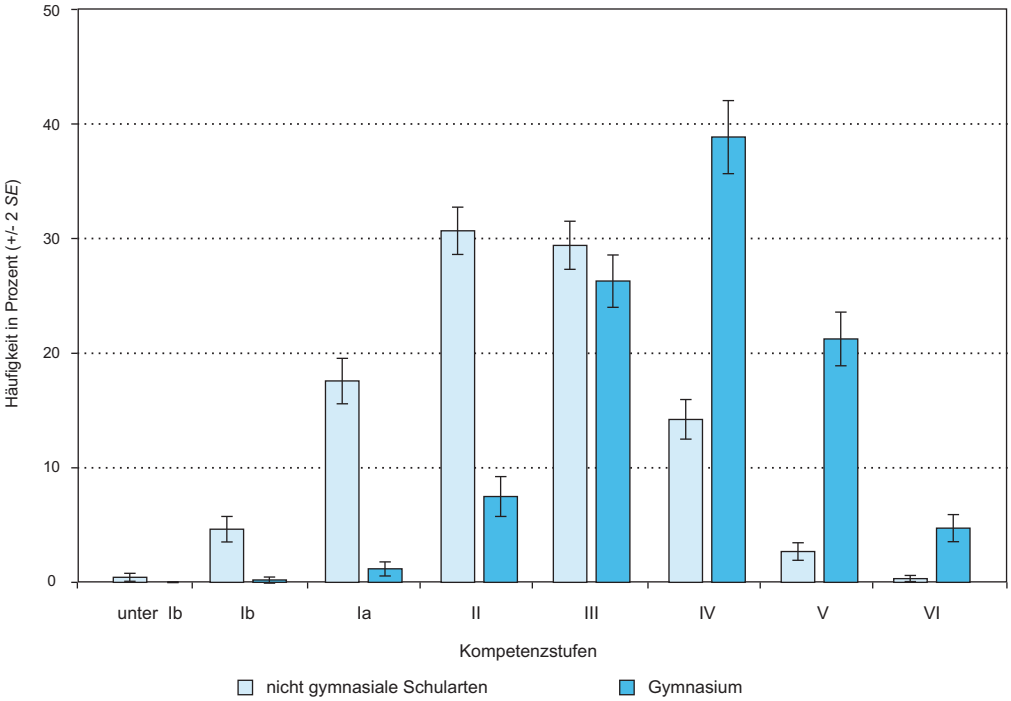


Abbildung 2.17: Prozentuale Anteile Fünfzehnjähriger auf den Stufen der naturwissenschaftlichen Kompetenz in Deutschland nach Schulart

Verteilung auf die Stufen naturwissenschaftlicher Kompetenz

Diese Befunde spiegeln sich in der Verteilung der Jugendlichen auf die Stufen der naturwissenschaftlichen Kompetenz wider, die in Abbildung 2.17 dargestellt ist.

Die Häufigkeitsverteilungen sind deutlich gegeneinander verschoben. An den nicht gymnasialen Schularten beträgt der Anteil der Jugendlichen mit sehr niedrigen Kompetenzen (Kompetenzstufe Ia und darunter) 22.7 Prozent, während nur 3 Prozent dieser

© Waxmann Verlag GmbH

Fünfzehnjährigen die Kompetenzstufen V und IV erreichen. Für die Gymnasien findet sich der größte Anteil der Schülerinnen und Schüler (38.9 Prozent) auf Kompetenzstufe IV. Jugendliche mit sehr niedrigen Kompetenzen, d.h. auf Kompetenzstufe Ia und darunter, sind am Gymnasium praktisch nicht zu finden (1.4 Prozent), während sich gut ein Viertel der Fünfzehnjährigen auf den Stufen V und VI befindet und damit ausgezeichnete Voraussetzungen etwa für eine berufliche Karriere im naturwissenschaftlich-technischen Bereich mitbringt. Eine weitere Förderung dieser Spitzengruppe am Gymnasium ist vor dem Hintergrund der nach wie vor besseren Spitzenleistungen in anderen Staaten wünschenswert. Darüber hinaus müssen insbesondere in den nicht gymnasialen Schularten weitere Anstrengungen unternommen werden, um eine stärkere Breitenförderung im Sinne einer Verringerung der Jugendlichen auf den untersten Kompetenzstufen zu erreichen. Die Förderung der hochkompetenten Schülerinnen und Schüler ist selbstredend auch hier eine Aufgabe von gezielten Fördermaßnahmen.

Ergebnisse für die naturwissenschaftlichen Teilkompetenzen, Wissensbereiche und Wissenssysteme

Die im internationalen Vergleich berichteten Ergebnisse für die verschiedenen Teilkomponenten der naturwissenschaftlichen Kompetenz getrennt nach Gymnasium und nicht gymnasialen Schularten zeigt Tabelle 2.12. Bei der Interpretation ist zu beachten, dass nur Teilkompetenzen, Wissensbereiche und Wissenssysteme untereinander verglichen werden können. Ein Vergleich zum Beispiel des Wissensbereichs *konzeptuelles Wissen* und des Wissenssystems *physikalische Systeme* ist aufgrund der vorgenommenen Skalierungen nicht möglich.

Erwartungsgemäß erreichen die Fünfzehnjährigen am Gymnasium in allen Teilbereichen höhere Kompetenzmittelwerte als diejenigen in nicht gymnasialen Schularten. Interessante Unterschiede ergeben sich jedoch bei der Betrachtung von relativen Stärken und Schwächen innerhalb der einzelnen Bereiche. So zeigt sich am Gymnasium im Vergleich zu den anderen Teilkompetenzen eine relative Stärke bei *Daten und Evidenz naturwissenschaftlich interpretieren*. Die erreichten 590 Punkte liegen signifikant über den Leistungen in den Teilkompetenzen *Phänomene naturwissenschaftlich erklären* und *naturwissenschaftliche Forschung bewerten und planen*. Bei den nicht gymnasialen Schularten zeigt sich eine signifikante Stärke in der Teilkompetenz *Phänomene naturwissenschaftlich erklären* im Vergleich zu den anderen Teilkompetenzen.

In den Wissensbereichen sind Jugendliche in den nicht gymnasialen Schularten im Bereich *konzeptuelles Wissen* signifikant besser als im *prozedural-epistemischen Wissen*. Am Gymnasium deutet sich dieser Unterschied auch an, hier ist er aber nicht signifikant. Dies legt die Vermutung nahe, dass der naturwissenschaftliche Unterricht in Deutschland eher auf die Vermittlung konzeptuellen Wissens fokussiert, wenngleich die Bildungsstandards für den mittleren Schulabschluss explizit eine gleichmäßige Förderung aller Wissens- bzw. Kompetenzbereiche fordern (KMK, 2005a, 2005b, 2005c).

© Waxmann Verlag GmbH

Tabelle 2.12: Mittelwerte der Teilskalen in Deutschland nach Schulart

	nicht gymnasiale Schularten		Gymnasium	
	M	*(SE)*	*M*	*(SE)*
Phänomene naturwissenschaftlich erklären	478	(3.3)	586	(2.9)
Naturwissenschaftliche Forschung bewerten und Untersuchungen planen	473	(3.5)	582	(3.6)
Daten und Evidenz naturwissenschaftlich interpretieren	473	(3.2)	590	(3.2)
Konzeptuelles Wissen	479	(3.4)	588	(3.2)
Prozedural-epistemisches Wissen	473	(3.1)	585	(2.7)
Physikalische Systeme	473	(3.2)	579	(3.4)
Lebende Systeme	473	(3.2)	589	(3.0)
Erd- und Weltraumsysteme	479	(3.4)	588	(3.1)
Gesamtskala	476	(3.1)	585	(2.6)

2.3.4 Die Veränderung der naturwissenschaftlichen Kompetenz zwischen PISA 2006 und PISA 2015

PISA 2015 schließt mit den Naturwissenschaften als Hauptdomäne den zweiten vollständigen PISA-Erhebungszyklus ab. Da bei einem Status als Nebendomäne immer eine kleinere Anzahl an Aufgaben eingesetzt und die Kompetenz weniger differenziert erfasst wird, konzentrieren sich die folgenden Betrachtungen auf den Vergleich von PISA 2006 und PISA 2015. Wie in den Abschnitten 2.1 und 2.2 beschrieben, wurden die Rahmenkonzeption, die Testaufgaben und die Testprozeduren aus dem Jahr 2006 (als die Naturwissenschaften erstmals Hauptdomäne waren) in verschiedener Hinsicht modifiziert. Durch diese Anpassungen werden ohne Zweifel neue Entwicklungen angemessen berücksichtigt. Allerdings limitieren entsprechende Modifikationen eine Vergleichbarkeit der Befunde und damit Aussagen über die Veränderungen und Entwicklungen in der naturwissenschaftlichen Kompetenz in der letzten Dekade (vgl. auch Kapitel 12).

Mit entsprechenden Vorbehalten müssen Veränderungen auf der Gesamtskala naturwissenschaftlicher Kompetenz betrachtet werden. An sich sind Trendaussagen bei einem längerfristig angelegten Bildungsmonitoring aber von besonderem Interesse, da sie Hinweise auf die Wirksamkeit von (möglicherweise zu bestimmten Zeitpunkten ergriffenen) Maßnahmen zur Förderung naturwissenschaftlicher Kompetenz liefern.

Abbildung 2.18 zeigt die Veränderung der naturwissenschaftlichen Kompetenz von PISA 2006 zu PISA 2015 für die OECD-Staaten.

© Waxmann Verlag GmbH

OECD-Staaten	2015		2006		Differenz
	M	*(SE)*	*M*	*(SE)*	2015 – 2006
Japan	538	(3.0)	531	(3.4)	
Estland	534	(2.1)	531	(2.5)	
Finnland	531	(2.4)	563	(2.0)	
Kanada	528	(2.1)	534	(2.0)	
Korea	516	(3.1)	522	(3.4)	
Neuseeland	513	(2.4)	530	(2.7)	
Slowenien	513	(1.3)	519	(1.1)	
Australien	510	(1.5)	527	(2.3)	
Vereinigtes Königreich	509	(2.6)	515	(2.3)	
Deutschland	509	(2.7)	516	(3.8)	
Niederlande	509	(2.3)	525	(2.7)	
Schweiz	506	(2.9)	512	(3.2)	
Irland	503	(2.4)	508	(3.2)	
Belgien	502	(2.3)	510	(2.5)	
Dänemark	502	(2.4)	496	(3.1)	
Polen	501	(2.5)	498	(2.3)	
Portugal	501	(2.4)	474	(3.0)	
Norwegen	498	(2.3)	487	(3.1)	
Vereinigte Staaten	496	(3.2)	489	(4.2)	
Österreich	495	(2.4)	511	(3.9)	
Frankreich	495	(2.1)	495	(3.4)	
Schweden	493	(3.6)	503	(2.4)	
OECD-Durchschnitt	493	(0.4)	498	(0.5)	
Tschechische Republik	493	(2.3)	513	(3.5)	
Spanien	493	(2.1)	488	(2.6)	
Lettland	490	(1.6)	490	(3.0)	
Luxemburg	483	(1.1)	486	(1.1)	
Italien	481	(2.5)	475	(2.0)	
Ungarn	477	(2.4)	504	(2.7)	
Island	473	(1.7)	491	(1.6)	
Israel	467	(3.4)	454	(3.7)	
Slowakische Republik	461	(2.6)	488	(2.6)	
Griechenland	455	(3.9)	473	(3.2)	
Chile	447	(2.4)	438	(4.3)	
Türkei	425	(3.9)	424	(3.8)	
Mexiko	416	(2.1)	410	(2.7)	

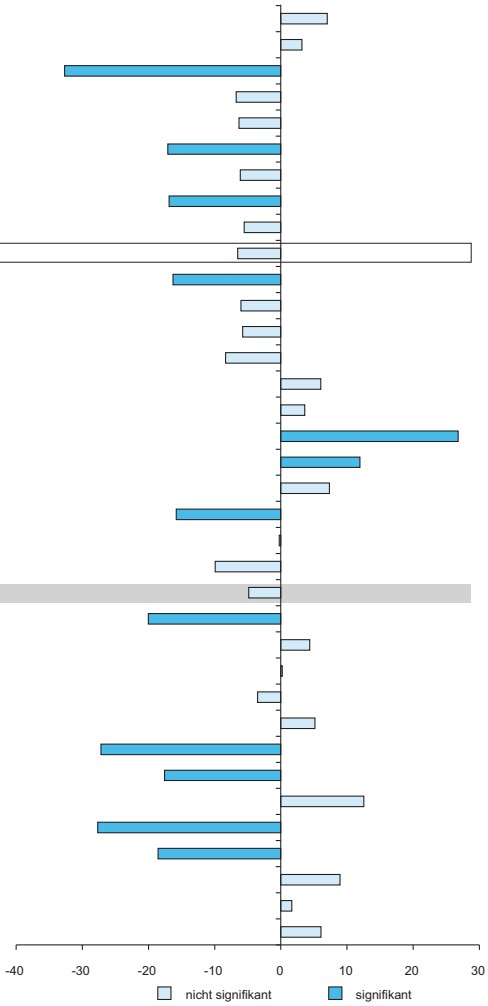


Abbildung 2.18: Mittelwerte naturwissenschaftlicher Kompetenz in den OECD-Staaten in PISA 2006 und PISA 2015

Im Mittel zeigt sich im Vergleich zu PISA 2006 eine leichte Abnahme der naturwissenschaftlichen Kompetenz für die aktuellen OECD-Staaten von 498 Punkte auf 493 Punkte, die sich jedoch unter Berücksichtigung des Link-Fehlers (Maß für die Messunsicherheit, vgl. Kapitel 12) nicht als statistisch signifikant absichern lässt. Der OECD-Mittelwert darf folglich im Vergleich von PISA 2006 und 2015 als unverändert gelten. In 23 Staaten findet sich entsprechend keine statistisch bedeutsame Veränderung der naturwissenschaftlichen Kompetenz zwischen PISA 2006 und PISA 2015. Zu dieser Gruppe zählen im oberen Kompetenzbereich neben Deutschland (mit einer Differenz von sieben Punkten) auch Japan, Estland, Kanada, Korea, Slowenien, das Vereinigte Königreich, die

© Waxmann Verlag GmbH

Schweiz und Irland. Zehn Staaten erreichen bei PISA 2015 signifikant niedrigere Mittelwerte als bei PISA 2006, allen voraus Finnland mit einem Verlust von 32 Punkten. Ebenfalls deutliche Rückgänge mit mehr als 20 Punkten sind im unteren Leistungsbereich für Ungarn und die Slowakische Republik zu verzeichnen. Zu dieser Gruppe zählen außerdem auch Neuseeland (–17 Punkte), Australien (–17 Punkte) und die Niederlande (–16 Punkte) sowie Österreich (–16 Punkte). Nur für zwei Staaten – Portugal und Norwegen – ist eine signifikante Verbesserung zu beobachten. Der Zuwachs fällt insbesondere für Portugal mit 27 Punkten sehr deutlich aus. Auch wenn sich in Deutschland tendenziell eine Abnahme der Kompetenz zeigt, liegt diese noch im Bereich möglicher Zufallsschwankungen. Die Ergebnisse für Deutschland weisen somit darauf hin, dass sich die naturwissenschaftliche Kompetenz der Schülerinnen und Schüler auf einem hohen, signifikant über dem OECD-Durchschnitt liegenden Niveau etabliert hat. Dass dies nicht selbstverständlich ist, zeigen Staaten wie Finnland, Australien und Kanada, in denen ein statistisch bedeutsamer Rückgang der naturwissenschaftlichen Kompetenz zu verzeichnen ist. Dass deutlich mehr Staaten in den Leistungen verloren als gewonnen haben, mag auch etwas mit dem Wechsel des Testmediums zu tun haben. Gesicherte Aussagen sind aufgrund der Daten aus PISA 2015 allerdings nicht möglich.

Die Verteilung der Schülerinnen und Schüler in Deutschland auf die Kompetenzstufen im Vergleich zwischen PISA 2006 und PISA 2015 gibt Hinweise auf potenzielle Bereiche eines besonderen Förderbedarfs. Der Anteil der Jugendlichen, deren Kompetenz unterhalb der Stufe II liegt, stieg leicht von 15.4 Prozent auf 17.0 Prozent an, der Anteil der Jugendlichen auf den Kompetenzstufen V und VI sank leicht von 11.8 Prozent auf 10.6 Prozent ab. Beide Veränderungen sind jedoch wiederum nicht statistisch signifikant und deuten eher darauf hin, dass sich die Befunde stabilisiert haben.

Für Deutschland lassen sich außer zu den Veränderungen auf der Gesamtskala der naturwissenschaftlichen Kompetenz auch Überlegungen zu Trends einer Veränderung der Mittelwerte von Jungen und Mädchen sowie der Gymnasien und der nicht gymnasialen Schularten anstellen. Für die Betrachtung der nicht gymnasialen Schularten wurden die Daten von PISA 2006 neu ausgewertet, indem die nicht gymnasialen Schularten zusammengefasst wurden. Die Mittelwerte und Streuungen sind in Tabelle 2.13 dargestellt.

Tabelle 2.13: Naturwissenschaftliche Kompetenz in Deutschland bei PISA 2006 und PISA 2015 getrennt nach Geschlecht und Schulart

	2006				2015			
	M	*(SE)*	*SD*	*(SE)*	*M*	*(SE)*	*SD*	*(SE)*
Mädchen	512	(3.8)	97	(2.0)	504	(3.2)	97	(1.6)
Jungen	519	(4.6)	103	(2.5)	514	(2.8)	93	(1.6)
Nicht gymnasiale Schularten	486	(4.3)	87	(2.3)	476	(3.0)	86	(1.5)
Gymnasium	598	(2.9)	68	(1.7)	585	(2.6)	75	(2.0)
Gesamtstichprobe	516	(3.8)	100	(2.0)	509	(2.7)	99	(1.5)

© Waxmann Verlag GmbH

Bei PISA 2006 lag die Differenz zwischen Mädchen und Jungen bei sieben Punkten, in PISA 2015 sind es zehn Punkte. Statistisch gesehen haben Jungen im Vergleich zu Mädchen nun einen signifikanten Vorsprung in der naturwissenschaftlichen Kompetenz. Dieser Befund lässt sich jedoch nicht auf eine signifikante Verschlechterung der Mittelwerte der Mädchen zurückführen. Wie bei den Jungen ist innerhalb der Gruppen die Veränderung in der Kompetenz bei den Mädchen statistisch nicht bedeutsam, was auf die Berücksichtigung des Link-Fehlers zurückzuführen ist (vgl. Kapitel 12).

Ein Blick auf die Veränderung der Kompetenzstufenverteilung zeigt, dass bei PISA 2006 bereits signifikant mehr Jungen (13.7 Prozent) auf den Kompetenzstufen V und VI zu finden waren als Mädchen (9.8 Prozent). Dies ist auch in PISA 2015 der Fall (Jungen: 12.4 Prozent, Mädchen: 8.7 Prozent). Der Anteil von Mädchen mit sehr niedrigen Kompetenzwerten (Kompetenzstufe Ia und darunter) ist allerdings 2015 mit 18.1 Prozent signifikant höher als der der Jungen (15.9 Prozent). Dies war 2006 nicht der Fall. Jedoch ist auch hier die Differenz zwischen 2006 und 2015 nicht statistisch signifikant. Auffällig ist bei der Betrachtung der Geschlechter auch die Verringerung der Streuung bei den Jungen auf 93 Punkte (im Vergleich zu 103 Punkten in PISA 2006). Auch wenn die meisten Veränderungen der Mittelwerte und Kompetenzstufenverteilungen bei Mädchen und Jungen nicht als statistisch bedeutsam abgesichert werden können, sollte gerade die Entwicklung der Mädchen sorgfältig im Blick behalten werden, damit sich die Schere zwischen Jungen und Mädchen nicht vergrößert.

Betrachtet man die Veränderung der mittleren Kompetenzwerte differenziert nach Schularten, lassen sich Fragen dazu aufwerfen, inwieweit es dem differenzierten Schulsystem in Deutschland gelingt, eine breite naturwissenschaftliche Grundbildung anzulegen, Talente zu fördern und die Leistungsspitze auszubauen. Für die nicht gymnasialen Schularten zeigen sich keine signifikanten Veränderungen der naturwissenschaftlichen Kompetenz. Am Gymnasium ist jedoch ein statistisch signifikanter Rückgang der naturwissenschaftlichen Kompetenz um 13 Punkte zu finden. Ebenso ist am Gymnasium eine Zunahme der Streuung um sieben Punkte zu beobachten, die zwar nicht statistisch signifikant ist, aber auf eine leichte Zunahme der Heterogenität der Schülerschaft hindeutet. Eine mögliche Erklärung könnte die leichte Erhöhung der Bildungsbeteiligung am Gymnasium von 31 Prozent in 2006 auf 33 Prozent in 2015 sein. Wiederum kann nicht ausgeschlossen werden, dass die Rückgänge auch dem Wechsel des Testmediums geschuldet sein können.

2.4 Zusammenfassung und Diskussion

Die naturwissenschaftliche Kompetenz der Fünfzehnjährigen in Deutschland in PISA 2015 ist mit 509 Punkten gegenüber dem OECD-Mittelwert von 493 Punkten erneut überdurchschnittlich ausgeprägt. Deutschland positioniert sich damit innerhalb der Gruppe von Staaten (Korea, Neuseeland, Slowenien, Australien, Vereinigtes Königreich, Niederlande, Schweiz und Irland), die den Spitzenstaaten Japan, Estland, Finnland und

© Waxmann Verlag GmbH

Kanada folgt. Auch der Anteil an hochkompetenten Jugendlichen ist signifikant größer als im OECD-Durchschnitt, jedoch niedriger als in anderen Ländern. Betrachtet man die Veränderung der naturwissenschaftlichen Kompetenz im Vergleich zu PISA 2006 – unter den Vorbehalten einer weiterentwickelten Rahmenkonzeption und eines veränderten Tests –, dann lässt sich keine signifikante Veränderung der naturwissenschaftlichen Kompetenz in Deutschland feststellen. In Bezug auf die Veränderung des Darbietungsmodus von Papier auf den Computer weisen die Ergebnisse zusätzlicher Analysen der Feldtestdaten durch das Zentrum für Internationale Bildungsvergleichsstudien (ZIB) darauf hin, dass die Naturwissenschaftsaufgaben am Computer etwas schwieriger waren als auf Papier (Robitzsch et al., im Druck). Nicht auszuschließen ist deshalb, dass die Kompetenzveränderung in Deutschland von 2006 zu 2015 eher überschätzt wird. Insgesamt scheint es in Deutschland gelungen zu sein, die naturwissenschaftliche Kompetenz auf einem im internationalen Vergleich hohen Niveau zu stabilisieren, auch wenn der Abstand zur internationalen Spitze noch beträchtlich ist. Immerhin hat sich Deutschland im Kreise der Staaten, die naturwissenschaftlich-technische Nachwuchsförderung auf hohem Niveau betreiben, etabliert. Dies ist nicht selbstverständlich. In einer Reihe von Staaten – darunter auch solche, die wie Finnland, Neuseeland und Australien in 2006 Spitzenleistungen erbrachten – ist ein deutlicher Punktverlust bezüglich der mittleren Kompetenz zu beobachten.

Bei der differenzierten Betrachtung der Teilkompetenzen zeigte sich in Deutschland ein positives Bild: Die Fünfzehnjährigen erzielen über alle drei Kompetenzbereiche hinweg eine vergleichsweise ausgeglichene Leistung, die über dem OECD-Durchschnitt liegt. Damit gelingt die gleichmäßige Ausbildung der Teilkompetenzen in Deutschland vergleichsweise gut. Allerdings zeigen erneut Staaten wie Japan oder Estland, deren Jugendliche in allen drei Teilkompetenzen höhere Werte erzielen, dass es hier Entwicklungspotenzial gibt und im naturwissenschaftlichen Unterricht in Deutschland alle Teilkompetenzen stärker gefördert werden sollten. Des Weiteren sollte ein besonderes Augenmerk auf die Vermittlung von *prozeduralem und epistemischem Wissen* gelegt werden. Zudem zeigen sich Unterschiede in den Wissenssystemen zugunsten der *lebenden Systeme*, was mit einem höheren Anteil an Biologiestunden im deutschen Schulsystem gegenüber der Physik und der Chemie zusammenhängen könnte.

Im Vergleich zu früheren PISA-Erhebungsrunden überrascht der Befund, dass nun und erstmals sowohl im OECD-Durchschnitt als auch in Deutschland die Jungen eine signifikant bessere naturwissenschaftliche Kompetenz erzielen als die Mädchen. Deutschland gehört zu einem der wenigen Staaten, in denen der Anteil von Mädchen auf den unteren Kompetenzstufen (Stufe Ia und darunter) signifikant größer ist als der Anteil der Jungen. Es gibt also offensichtlich im Vergleich zu 2006 inzwischen mehr Mädchen, die nicht einmal über Basiskompetenzen in den Naturwissenschaften verfügen. Auch wenn der Anteil hochkompetenter Mädchen in Deutschland größer ist als im OECD-Durchschnitt, bleibt auch dieser hinter dem Anteil der Jungen zurück. Unter den hochkompetenten Jugendlichen befinden sich demnach deutlich weniger Mädchen als Jungen. Die Vermutung, dass möglicherweise das neue, computerbasierte Testfor-

© Waxmann Verlag GmbH

mat die Jungen bei der Testung bevorteilt, konnte durch bisherige detaillierte Analysen nicht bestätigt werden. Ein Blick auf die Ausprägung der Teilkompetenzen liefert Hinweise auf einen differenziellen geschlechterspezifischen Förderbedarf: Jungen erzielen vor allem einen deutlichen Vorsprung in der Teilkompetenz *Phänomene naturwissenschaftlich erklären*. Das gerade für diese Teilkompetenz benötigte *konzeptuelle Wissen* ist bei ihnen ebenfalls stärker ausgeprägt. Die Schwäche der Mädchen liegt im Bereich *physikalische Systeme*, während Jungen in allen Wissenssystemen homogene Leistungen zeigen. Es sei angemerkt, dass die Befundlage zu Geschlechterunterschieden in den Naturwissenschaften in Deutschland nicht einheitlich ist. So untersuchte der Ländervergleich 2012, dessen Schwerpunkt auf der Ausprägung der mathematisch-naturwissenschaftlichen Kompetenz am Ende der neunten Jahrgangsstufe lag, ebenfalls Geschlechterunterschiede in den Bildungsstandarddimensionen Fachwissen und Erkenntnisgewinnung für Biologie, Chemie und Physik. Hier fanden sich – im Gegensatz zu PISA – in allen Bereichen außer im Fachwissen Physik Vorsprünge zugunsten der Mädchen (Schroeders, Penk, Jansen & Pant, 2013). Es zeigte sich entsprechend, dass von den Naturwissenschaften nur die Physik eher als die Domäne der Jungen gesehen werden kann. Eine an PISA 2012 angeknüpfte Längsschnittstudie, welche die Entwicklung der naturwissenschaftlichen Kompetenz von der neunten zur zehnten Klassenstufe untersuchte, ergab, dass die Mädchen am Ende der Klasse 10 ein vergleichbares Niveau naturwissenschaftlicher Kompetenz erreichten wie die Jungen (Schiepe-Tiska et al., im Druck). In der Konsequenz kann man feststellen, dass die Kompetenzunterschiede zwischen Jungen und Mädchen in den Naturwissenschaften nicht unbedingt aufgeregt gesehen werden müssen, aber weiterhin aufmerksam beobachtet werden sollten. Mit Blick auf das Erreichen mehrdimensionaler Bildungsziele sollten neben den kognitiven Lernergebnissen auch motivationale Orientierungen und Einstellungen gefördert werden, die für die Entwicklung naturwissenschaftlicher Kompetenz ebenfalls von zentraler Bedeutung sind. Jedoch zeigen sich auch für diese motivational-affektiven Lernergebnisse deutliche Unterschiede zwischen den Geschlechtern zugunsten der Jungen (vgl. Kapitel 3).

Eine differenzierte Betrachtung der naturwissenschaftlichen Kompetenz nach Schulart zeigt, dass gerade Jugendliche am Gymnasium eine relative Stärke in der Teilkompetenz *Daten und Evidenz naturwissenschaftlich interpretieren* haben. Schülerinnen und Schüler in nicht gymnasialen Schularten haben eine relative Stärke in der Teilkompetenz *Phänomene naturwissenschaftlich erklären*. In allen Schularten scheint der Fokus darüber hinaus eher auf dem Aufbau *konzeptuellen Wissens* und weniger auf dem Entwickeln *prozeduralen und epistemischen Wissens* zu liegen. Hier zeigen sich Entwicklungsmöglichkeiten für den naturwissenschaftlichen Unterricht in Deutschland.

Im Vergleich zu PISA 2006 ist am Gymnasium außerdem ein statistisch bedeutsamer Rückgang der naturwissenschaftlichen Kompetenz zu finden. Auch wenn dieser Trend aufgrund der veränderten Testanlage nur mit Vorsicht interpretiert werden darf, zeigt sich auch im internationalen Vergleich, dass die Förderung der Spitzengruppe nicht so gut gelingt wie in anderen Staaten. Damit schöpft Deutschland die Möglichkeiten eines differenzierten Schulsystems zur Entwicklung einer naturwissenschaftlichen Grundbil-

© Waxmann Verlag GmbH

dung in der Breite, bei gleichzeitigem Anspruch einer Förderung von Talenten in den Naturwissenschaften, weiterhin nicht hinreichend aus. Auch bei einer leicht gestiegenen Bildungsbeteiligung am Gymnasium sollte mit einer Auswahl der leistungsstärksten Schülerinnen und Schüler die Talentförderung besser gelingen als in nicht differenzierenden Schulsystemen. Für den Wohlstand und das Wohlergehen eines Staates ist gerade die Ausbildung hochkompetenter Jugendlicher, die in der Lage sind, komplexe naturwissenschaftliche Sachverhalte zu durchdringen und zu verstehen, im Sinne einer Nachwuchsförderung zentral. Ein wichtiger Schritt zur Förderung dieser Jugendlichen ist die im letzten Jahr verabschiedete bundesweite Strategie zur Förderung leistungsstarker Schülerinnen und Schüler (KMK, 2015). Diese Strategie greift insbesondere die Bedeutung einer systematischen Beobachtung und Diagnose von Lernvoraussetzungen durch Lehrkräfte, ergänzt durch schulexterne Diagnostik (z. B. Schulberatungsstellen, schulpsychologische Dienste), auf. Sie fasst gezielte Maßnahmen auf der Ebene von Schule und Unterricht zusammen (Enrichment, Akzeleration, Gruppierung sowie integrierte Förderung) und macht Vorschläge, wie diese umgesetzt werden können. Darüber hinaus regt sie zum Aufbau von Netzwerken und zur Zusammenarbeit mit externen Partnern an. Diese Strategie ergänzt damit den Beschluss zur Förderstrategie für leistungsschwächere Schülerinnen und Schüler (KMK, 2010), dessen Ziel eine Verringerung des Anteils der Jugendlichen ist, die nicht das Mindestniveau der Kompetenzentwicklung erreichen. Damit wurde auf der politischen Ebene der Fokus auf beide Zielgruppen gerichtet. Beide Strategien unterstreichen die Bedeutung einer qualitativ hochwertigen Aus- und Weiterbildung der Lehrkräfte, um Jugendliche individuell zu fördern und den eigenen Unterricht kompetenzorientiert zu gestalten.

Literatur

Aktionsrat Bildung. (2015). *Bildung. Mehr als Fachlichkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Baumert, J., Lehmann, R. H., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller, O. & Neubrand, J. (1997). *TIMSS. Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich: Deskriptive Befunde.* Opladen: Leske + Budrich. http://dx.doi.org/10.1007/978-3-322-95096-3

Bayrhuber, H., Bögeholz, S., Elster, D., Hammann, M., Hößle, C., Lücken, M., Mayer, J., Nerdel, C., Neuhaus, B., Prechtl, H. & Sandmann, A. (2007). Biologie im Kontext – Ein Programm zur Kompetenzförderung durch Kontextorientierung im Biologieunterricht und zur Unterstützung von Lehrerprofessionalisierung. *Der mathematische und naturwissenschaftliche Unterricht, 60* (5), 282–286.

Beaton, A. E., Martin, M. O., Mullis, I. V., Gonzalez, E. J., Smith, T. A. & Kelly, D. L. (1996). *Science achievement in the middle school years: IEA's third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

© Waxmann Verlag GmbH

Bybee, R. W. & Fuchs, B. (2006). Preparing the 21st century workforce: A new reform in science and technology education. *Journal of Research in Science Teaching, 43*, 349–352. https://doi.org/10.1002/tea.20147

Bybee, R. & McCrae, B. (2011). Scientific literacy and student attitudes: Perspectives from PISA 2006 science. *International Journal of Science Education, 33*, 7–26. http://dx.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

DeBoer, G. E. (2011). The globalization of science education. *Journal of Research in Science Teaching, 48*, 567–591. https://doi.org/10.1002/tea.20421

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

Fensham, P. J. (1985). Science for all: A reflective essay. *Journal of Curriculum Studies, 17*, 415–435. http://dx.doi.org/10.1080/0022027850170407

Hannover, B. & Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science. *Learning and Instruction, 14*, 51–67. https://doi.org/10.1016/j.learninstruc.2003.10.002

Häußler, P. & Hoffman, L. (1998). *BLK-Programmförderung „Steigerung der Effizienz des mathematisch naturwissenschaftlichen Unterrichts". Erläuterungen zu Modul 7 mit Unterrichtsbeispielen für den Physikunterricht* (Förderung von Mädchen und Jungen) [Stand 08/1998]. Kiel: Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel (IPN).

Häußler, P., Hoffman, L., Langeheine, R., Rost, J. & Sievers, K. (1998). A typology of students' interest in physics and the distribution of gender and age within each type. *International Journal of Science Education, 20*, 223–238. http://dx.doi.org/10.1080/0950069980200207

Kessels, U. & Hannover, B. (2007). How the image of math and science affects the development of academic interests. In M. Prenzel (Hrsg.), *Studies on the educational quality of schools. The final report on the DFG priority programme* (S. 283–297). Münster: Waxmann.

KMK (2005a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005a). *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. (2005b). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards 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. (2005c). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Physik für den mittleren Bildungsabschluss (Beschluss vom 16. Dezember 2004)*. München: Wolters Kluwer.

© Waxmann Verlag GmbH

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)*. Zugriff am 27.10.2016. Verfügbar unter http://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)*. Zugriff am 27.10.2016. Verfügbar unter http://www.kmk.org/fileadmin/Dateien/pdf/350-KMK-TOP-011-Fu-Leistungsstarke_-_neu.pdf

Labudde, P. & Möller, K. (2012). Stichwort: Naturwissenschaftlicher Unterricht. *Zeitschrift für Erziehungswissenschaften, 15,* 11–36. http://dx.doi.org/ 10.1007/s11618-012-0257-0

Mikelskis-Seifert, S. & Duit, R. (2010). *Physik im Kontext – Konzepte, Ideen, Materialien für effizienten Physikunterricht*. Seelze: Friedrich Verlag.

Millar, R. (2006). Twenty first century science: Insights from the design and implementation of a scientific literacy approach in school science. *International Journal of Science Education, 28*, 1499–1521. http://dx.doi.org/10.1080/09500690600718344

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. (2006). *Assessing scientific, reading and mathematical literacy. A framework for PISA 2006*. Paris: OECD. Zugriff am 27.10.2016. Verfügbar unter http://www.oecd.org/pisa/pisaproducts/pisa2006/37464175.pdf

OECD. (2016a). *PISA 2015 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. Paris: OECD.

OECD. (2016b). *PISA 2015 results (Volume I): Excellence and equity in education*. Paris: OECD.

OECD. (in Vorbereitung). *PISA 2015 technical report*. Paris: OECD.

Osterlind, S. J. (1990). Toward a uniform definition of a test item. *Educational Research Quarterly, 14* (4), 2–5.

Pant, H. A., Stanat, P., Pöhlmann, C., Hecht, M., Jansen, M., Kampa, N., Lenski, A. E., Penk, C., Radmann, S., Roppelt, A., Schroeders, U., Siegle, T. & Ziemke, A. (2013). Der Blick in die Länder. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 159–247). Münster: Waxmann.

Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T. & Pöhlmann, C. (Hrsg.). (2013). *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Münster: Waxmann.

Prenzel, M., Friedrich, A. & Stadler, M. A. (2009). *Von SINUS lernen: Wie Unterrichtsentwicklung gelingt*. Seelze-Velber: Klett/Kallmeyer.

Prenzel, M., Reiss, K. & Hasselhorn, M. (2009). Förderung der Kompetenzen von Kindern und Jugendlichen. In J. Milberg (Hrsg.), *Förderung des Nachwuchses in Technik und Naturwissenschaft* (S. 15–60). Berlin, Heidelberg: Springer.

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, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA*

© Waxmann Verlag GmbH

2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich (S. 192–250). Opladen: Leske + Budrich.

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.

Roberts, D. A. (2007). Scientific literacy/science literacy. In S. K. Abell & N. G. Lederman (Hrsg.), *Handbook of research on science education* (S. 729–780). Mahwah, N.J: Lawrence Erlbaum Associates.

Robitzsch, A., Lüdtke, O., Köller, O., Kröhne, U., Goldhammer, F. & Heine, J.-H. (im Druck). Herausforderungen bei der Schätzung von Trends in Schulleistungsstudien: Eine Skalierung der deutschen PISA Daten. *Diagnostica*.

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, W. Schneider, P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 177–198). 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., Rönnebeck, S., Heitmann, P., Schöps, K., Prenzel, M. & Nagy, G. (im Druck). 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*.

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.

Schroeders, U., Penk, C., Jansen, M. & Pant, H. A. (2013). Geschlechtsbezogene Disparitäten. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 249–274). Münster: Waxmann.

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

© Waxmann Verlag GmbH

3 Motivationale Orientierungen, Selbstbilder und Berufserwartungen in den Naturwissenschaften in PISA 2015

Anja Schiepe-Tiska, Inga Simm & Stefanie Schmidtner

Die Entwicklung motivationaler Orientierungen und Selbstbilder in den Naturwissenschaften ist neben dem Wissens- bzw. Kompetenzzuwachs wichtiges Ziel naturwissenschaftlichen Unterrichts und beeinflusst zukünftige Ausbildungs- und Berufswahlentscheidungen. In Bezug auf diese Bildungsziele zeichnet sich in Deutschland im internationalen Vergleich ein aufschlussreiches Bild ab, das bei differenzierter Betrachtung Handlungsbedarf anzeigt. Die Fünfzehnjährigen berichten allgemein über weniger Freude und Interesse an den Naturwissenschaften, interessieren sich jedoch im Vergleich zum OECD-Durchschnitt überdurchschnittlich stark für verschiedene naturwissenschaftliche Themen – vor allem aus den Bereichen *lebende Systeme* sowie *Erd- und Weltraumsysteme*. Die instrumentelle Motivation liegt ebenfalls unterhalb des OECD-Durchschnitts. Nur der Hälfte der Jugendlichen ist bewusst, dass Naturwissenschaften für ihr späteres Leben wichtig sein werden. Auch in dem, was sie sich in Bezug auf die Naturwissenschaften zutrauen, liegt deutliches Verbesserungspotenzial. Die Ausprägungen dieser Merkmale sind außerdem im Vergleich zu PISA 2006 gesunken. Die Tendenz, später einen naturwissenschaftlichen Beruf ergreifen zu wollen, ist in Deutschland im Vergleich zum OECD-Durchschnitt zwar etwas geringer ausgeprägt, im Vergleich zu PISA 2006 jedoch angestiegen. Im internationalen Vergleich besonders auffällig sind die in Deutschland auftretenden Geschlechterdifferenzen in allen Schülermerkmalen zugunsten der Jungen. Auch können sich mehr Jungen vorstellen, zukünftig einen naturwissenschaftlichen Beruf auszuüben. Weiterführende Analysen zeigen, dass in Deutschland die Bereitschaft, einen MINT-Beruf zu ergreifen, damit zusammenhängt, inwieweit Jugendliche Freude und Interesse an Naturwissenschaften allgemein erleben und die Bedeutung und die Relevanz von Naturwissenschaften für ihr zukünftiges Leben erkennen.

Naturwissenschaften und Technik sind für die Zukunft und Innovationsfähigkeit einer Gesellschaft von zentraler Bedeutung, um den Herausforderungen der heutigen Zeit wie beispielsweise der zunehmenden Digitalisierung, der älter werdenden Bevölkerung oder dem Klimawandel angemessen begegnen zu können. Aber auch für jeden Einzelnen sind Naturwissenschaften und Technik Gegenstandsbereiche, die helfen, die Welt und sich

© Waxmann Verlag GmbH

selbst zu verstehen und verantwortungsbewusst zu handeln. Naturwissenschaftliches Wissen bildet dafür eine wichtige Grundlage – es reicht allein jedoch nicht aus, um sich auch über den Unterricht und die Schule hinaus im Alltag mit Naturwissenschaften auseinandersetzen zu wollen und diese für mögliche Ausbildungs- und Berufswahlentscheidungen in Betracht zu ziehen (vgl. Schiepe-Tiska, Rozcen, Müller, Prenzel & Osborne, im Druck). Dabei bieten Berufe im Bereich Mathematik, Informatik, Naturwissenschaften und Technik (MINT) nach wie vor gute Berufschancen: Die Beschäftigung ist im Jahr 2015 im Vergleich zum Vorjahr gestiegen, insgesamt war die Arbeitslosigkeit tendenziell rückläufig und eine weiter steigende Nachfrage nach MINT-Arbeitskräften war festzustellen – jede vierte neu gemeldete Stelle bei der Bundesagentur für Arbeit richtete sich an MINT-Fachkräfte (Bundesagentur für Arbeit, 2016). Aktuell ist zwar kein flächendeckender Fachkräftemangel zu verzeichnen, allerdings zeichnet sich ab, dass der technologische Fortschritt und die rasante Digitalisierung die Zahl der benötigten Fachkräfte zukünftig ansteigen lässt. Außerdem ist ein Drittel der MINT-Erwerbstätigen 50 Jahre oder älter und geht demnach in absehbarer Zeit in den Ruhestand. Engpässe sowohl in den akademischen als auch in den nicht akademischen MINT-Berufen finden sich vor allem in den Bereichen Informatik, Ingenieurwissenschaften und Technik. Während die Anzahl der Studierenden in den MINT-Fächern deutlich gestiegen ist, gerade in den ingenieurwissenschaftlichen Fächern, sank der Anteil der Fachkräfte mit einer beruflichen Ausbildung und die Anzahl neu abgeschlossener MINT-Ausbildungsverträge nahm ab (acatech, 2015; Bundesagentur für Arbeit, 2016). Gerade in den nicht akademischen Berufen wird sich demnach der Fachkräftemangel in Zukunft verstärken.

Damit Jugendliche auch über die Schulzeit hinaus den Naturwissenschaften gegenüber aufgeschlossen bleiben und diese für ihr zukünftiges Berufsleben in Erwägung ziehen, müssen sie die Bedeutung und Relevanz von Naturwissenschaften für ihren Alltag *und* die Gesellschaft erfahren und erkennen. Wichtige Ziele naturwissenschaftlicher Grundbildung sind deshalb neben der Vermittlung von Wissen die Entwicklung motivationaler Orientierungen wie das Erleben von Freude und Interesse an Naturwissenschaften oder das Herausbilden positiver naturwissenschaftsbezogener Selbstbilder (Schiepe-Tiska et al., im Druck). Diese mehrdimensionalen Bildungsziele – Wissenszuwachs einerseits und die Förderung motivationaler Orientierungen und Selbstbilder andererseits – beeinflussen sich wechselseitig. Wenn sich Jugendliche für Naturwissenschaften interessieren und Zutrauen in ihre eigenen Fähigkeiten haben, können sie selbstbewusster mit naturwissenschaftlichen Themen umgehen und nutzen effektivere Lernstrategien (Schneider et al., 2015), was wiederum den Wissens- bzw. Kompetenzzuwachs fördert. Andererseits beeinflusst die naturwissenschaftliche Kompetenz, was sich Schülerinnen und Schüler in Bezug auf ihre Fähigkeiten zutrauen und inwieweit sie sich dafür interessieren (z. B. Köller, Trautwein, Lüdtke & Baumert, 2006). Wenn eine Jugendliche oder ein Jugendlicher zum Beispiel regelmäßig Misserfolge in den Naturwissenschaften erlebt, wird sie oder er vermutlich keine Freude daran haben, die Relevanz nicht erkennen sowie sich zukünftig noch weniger zutrauen. Motivationale Orien-

© Waxmann Verlag GmbH

tierungen und Selbstbilder sind demnach nicht nur wichtige Bildungsziele, sondern auch Bedingungs- und Prozessvariablen für die Leistungsentwicklung.

Die Schule und der naturwissenschaftliche Unterricht spielen für eine positive Entwicklung dieser Orientierungen und Einstellungen eine wichtige Rolle. Sie bieten Schülerinnen und Schülern Lerngelegenheiten für die Auseinandersetzung mit Naturwissenschaften und Technik. Hier kann und soll das Interesse an Naturwissenschaft geweckt sowie die Bedeutung für das eigene Leben und die Zukunft verdeutlicht werden. Darüber hinaus finden Jugendliche heraus, was sie sich selbst in Bezug auf ihre naturwissenschaftliche Kompetenz zutrauen können. Diese mehrdimensionalen Bildungsziele sind in Deutschland explizit in Schulgesetzen und Lehrplänen verankert – sie geraten allerdings im schulischen Alltag häufig aus dem Blick (vgl. Aktionsrat Bildung, 2015).

Auch in PISA 2015 wird die Bedeutung motivationaler Orientierungen und Einstellungen zu den Naturwissenschaften explizit in der Rahmenkonzeption hervorgehoben (OECD, 2016a; vgl. Kapitel 2). Diese beeinflussen – neben dem *konzeptuellen*, *prozeduralen* und *epistemischen Wissen* –, ob und wie Fünfzehnjährige die naturwissenschaftlichen Teilkompetenzen in konkreten Problemsituationen nutzen. Diese Orientierungen und Einstellungen wurden mithilfe des Schülerfragebogens erfasst, den die Fünfzehnjährigen im Anschluss an den Test bearbeiteten.

Im Folgenden wird zunächst der theoretische Hintergrund zu Orientierungen, Selbstbildern und Berufserwartungen beschrieben, und es werden die Herausforderungen, die es bei deren Erhebung zu berücksichtigen gilt, dargestellt. Im ersten Teil der Ergebnisse werden dann die Schülermerkmale in Deutschland mit denen ausgewählter OECD-Staaten verglichen, und es wird auf Unterschiede in den Merkmalsausprägungen zwischen Jungen und Mädchen eingegangen. Es schließt sich eine differenzierte Betrachtung für Deutschland getrennt nach Schularten sowie die Darstellung der Veränderung dieser Merkmale im Vergleich zu PISA 2006 an. Im zweiten Teil werden die Berufserwartungen der Fünfzehnjährigen genauer betrachtet. Neben einem Vergleich mit ausgewählten OECD-Staaten werden ebenfalls differenzielle Effekte für das Geschlecht und die Schulart untersucht. Abschließend wird der Zusammenhang zwischen kognitiven sowie motivational-affektiven Schülermerkmalen mit der Absicht, im Alter von 30 Jahren einen naturwissenschaftlichen Beruf auszuüben, betrachtet.

3.1 Theoretischer Hintergrund

In PISA 2015 wird die Bedeutung mehrdimensionaler Bildungsziele für die naturwissenschaftliche Grundbildung berücksichtigt: Motivationale Orientierungen und Einstellungen werden in der Rahmenkonzeption explizit genannt und bilden neben den Wissensbereichen (*konzeptuelles*, *prozedurales*, *epistemisches Wissen*) die Grundlage für die Anwendung der Teilkompetenzen in verschiedenen Kontexten (vgl. Kapitel 2; OECD, 2016a). Motivationale Orientierungen umfassen dabei zum Beispiel Freude und Interesse an den Naturwissenschaften, aber auch die instrumentelle Motivation. Die naturwissen-

© Waxmann Verlag GmbH

schaftsbezogene Selbstwirksamkeitserwartung wird ebenfalls als wichtiges Bildungsziel berücksichtigt. Des Weiteren bezieht die Rahmenkonzeption Einstellungen über die Naturwissenschaften als Wissenschaft mit ein (Osborne, Simon & Collins, 2003). Damit sind wissenschaftliche Denk- und Erkenntnishaltungen gemeint, die sich auf Dispositionen beziehen, nach naturwissenschaftlichen Erklärungen für Phänomene zu suchen und bestehende Erklärungen nicht als unveränderbar und für alle Zeiten gültig anzusehen. Ein zentraler Aspekt einer solchen Haltung ist es, empirische Beweise als Grundlage für Vorstellungen und Erklärungen anzusehen und wertzuschätzen. Zusätzlich wird die besondere Bedeutung des Verantwortungsbewusstseins gegenüber der Umwelt aufgegriffen. In diesem Kapitel soll der Schwerpunkt auf motivationalen Orientierungen und naturwissenschaftsbezogenen Selbstwirksamkeitserwartungen liegen, für die Betrachtung der anderen beiden Bereiche sei auf den internationalen Ergebnisbericht verwiesen (OECD, 2016b). Dazu werden im Folgenden die in PISA 2015 betrachteten Merkmale Freude und Interesse an Naturwissenschaften, Interesse an naturwissenschaftlichen Themen, instrumentelle Motivation sowie naturwissenschaftsbezogene Selbstwirksamkeitserwartungen vorgestellt und deren theoretische Hintergründe erläutert (für eine detailliertere Übersicht vgl. auch Schiepe-Tiska et al., im Druck).

3.1.1 Freude und Interesse an den Naturwissenschaften

Eine Reihe von Untersuchungen belegt, dass sich Kinder bis zum Ende der Grundschulzeit sehr für naturwissenschaftliche Phänomene und Fragestellungen interessieren (z. B. Prenzel, Geiser, Langeheine & Lobemeier, 2003; Prenzel, Lankes & Minsel, 2000). Im Laufe der Sekundarstufe I nimmt das Interesse jedoch über alle Schularten hinweg ab. Dies betrifft im mathematisch-naturwissenschaftlichen Bereich vor allem die Mädchen, Verluste treten ganz allgemein aber auch in anderen Fächern auf (vgl. Daniels, 2008).

Aus einer theoretischen Perspektive wird Interesse als Beziehung zwischen einer Person und einem Objekt bzw. einem Interessengegenstand verstanden. Interessen entstehen, wenn sich die Person mit dem Interessengegenstand beschäftigt und eine Beziehung dazu aufbaut (Krapp & Prenzel, 2011). Diese Objekte können Gegenstände oder Themen, spezifische Anwendungskontexte oder bestimmte Aktivitäten sein (Häußler & Hoffmann, 2000). In den Naturwissenschaften kann sich eine Schülerin oder ein Schüler beispielsweise allgemein für Naturwissenschaften, für das Experimentieren, für ein Schulfach wie Biologie oder Chemie (*Fachinteresse*) oder aber auch für ein konkretes naturwissenschaftliches Thema wie den menschlichen Körper (*Sachinteresse*) interessieren. In der Forschung werden häufig drei Komponenten von Interesse unterschieden, nämlich eine kognitive, eine wertbezogene und eine affektive Komponente (Krapp, 2002). Die kognitive Komponente bezieht sich auf den epistemischen Aspekt von Interessen. Sie beschreibt die Bereitschaft, sich neues Wissen über den Interessengegenstand anzueignen und mehr darüber zu erfahren, um ein tieferes Verständnis dafür zu entwickeln. Die wertbezogene Komponente bezeichnet die hohe persönliche Bedeutung, die

© Waxmann Verlag GmbH

einem Interessengegenstand beigemessen wird – die Jugendlichen identifizieren sich mit dem Interessengegenstand. Die affektive Komponente beschreibt, dass die Beschäftigung mit dem Interessengegenstand insgesamt vor allem mit positiven Emotionen wie Freude einhergeht. Diese Emotionen wiederum beeinflussen das Lern- und Leistungsverhalten von Jugendlichen und sind wichtiger Bestandteil ihres Wohlbefindens in der Schule (Pekrun, 2006).

Schule und Unterricht tragen entscheidend dazu bei, wie Jugendliche mit naturwissenschaftsbezogenen Anforderungssituationen umgehen. Fünfzehnjährige mit einem ausgeprägten Interesse und Freude an Naturwissenschaften nutzen elaboriertere metakognitive Strategien, lernen mehr dazu und entscheiden sich mit größerer Wahrscheinlichkeit für Schulen oder Schulzweige mit einem naturwissenschaftlichen Schwerpunkt oder naturwissenschaftsbezogene Berufe (vgl. Krapp & Prenzel, 2011; Pekrun & Linnenbrink-Garcia, 2014). Bei PISA 2006 wurden Freude und Interesse an Naturwissenschaften ebenfalls erhoben. In Deutschland waren beide im Vergleich zu den anderen OECD-Staaten unterdurchschnittlich ausgeprägt. Darüber hinaus berichteten Jungen über mehr Freude und Interesse an den Naturwissenschaften als Mädchen (Schütte, Frenzel, Asseburg & Pekrun, 2007). Auch die Fächer, für die sich Mädchen und Jungen interessieren, unterscheiden sich: Während Mädchen sich vor allem für Biologie interessieren, zeigen Jungen größeres Interesse für Physik und Chemie (vgl. Krapp & Prenzel, 2011).

3.1.2 Instrumentelle Motivation

Die naturwissenschaftsbezogene instrumentelle Motivation beschreibt, welche Bedeutung und Relevanz Jugendliche den Naturwissenschaften für ihre Zukunft zuschreiben. Die Selbstbestimmungstheorie (Deci & Ryan, 2002) geht davon aus, dass Schülerinnen und Schüler manchmal die externen Anforderungen, welche die Gesellschaft, die Schule, Lehrkräfte oder auch Eltern an sie stellen, in ihr eigenes Wertesystem integrieren müssen. Ihre Lerntätigkeiten sind demnach teilweise instrumentell auf den schulischen Unterricht allgemein oder speziell auf das spätere Ausbildungs- und Berufsleben ausgerichtet. Diese Art der Motivation muss daher nicht zwangsläufig mit dem Erleben von Freude und Interesse verbunden sein. Während interessengeleitete Handlungen intrinsisch sind, d. h. der Anreiz für die Handlung liegt in der Tätigkeit selbst, liegen bei instrumenteller Motivation die Anreize eher in den Folgen einer Handlung. Sie werden dadurch ausgelöst und aufrechterhalten, dass Schülerinnen und Schüler glauben, Naturwissenschaften seien für ihre zukünftigen Ausbildungs- und Berufschancen wichtig. Instrumentelle Motivation beeinflusst daher die Anstrengungsbereitschaft von Jugendlichen und sagt ihre zukünftige Ausbildungs- und Berufswahl vorher (z. B. Taskinen, Asseburg & Walter, 2008). Auch bei PISA 2006 wurde die naturwissenschaftsbezogene instrumentelle Motivation erfasst. In Deutschland schätzten Fünfzehnjährige die Relevanz von

© Waxmann Verlag GmbH

Naturwissenschaften für ihr zukünftiges Ausbildungs- und Berufsleben geringer ein als im OECD-Durchschnitt. Dies trifft vor allem auf die Mädchen zu. (Schütte et al., 2007).

3.1.3 Selbstwirksamkeitserwartungen

Neben dem Entwickeln motivationaler Orientierungen ist das Herausbilden positiver, realistischer Selbstbilder in Bezug auf die Naturwissenschaften ein zentrales Anliegen von Schule und Unterricht. Dazu gehört auch die Förderung naturwissenschaftsbezogener Selbstwirksamkeitserwartungen. Diese beziehen sich auf die subjektive Überzeugung, eine konkrete naturwissenschaftliche Problemstellung oder Handlung erfolgreich bewältigen zu können (Bandura, 1977). Wichtig ist dabei das Barriere-Konzept: Es geht nicht darum, dass sich Schülerinnen und Schüler generell erfolgreiche Lern- und Leistungshandlungen zutrauen, sondern dies vor allem in Situationen tun, in denen besondere Schwierigkeiten und Probleme auftreten, beispielsweise wenn Aufgaben besonders schwer sind. Selbstwirksamkeitserwartungen entwickeln sich aus vier Quellen: (a) der erfolgreichen Bewältigung ähnlicher Problemstellungen oder Handlungen in der Vergangenheit („Wenn ich es schon einmal geschafft habe, dann schaffe ich es wieder."), (b) Erfahrungen, die andere Personen stellvertretend für einen gemacht haben („Wenn diese Person es schafft, dann schaffe ich es auch."), (c) dem Vertrauen anderer Personen in einen selbst („Wenn meine Freunde/Eltern/Lehrkräfte denken, ich schaffe das, dann schaffe ich es auch.") und (d), einem gewissen körperlichen Erregungsniveau wie beispielsweise ein schnellerer Herzschlag („Ich bin zwar aufgeregt, aber ich schaffe das".).

Selbstwirksamkeitserwartungen hängen ebenfalls mit erhöhter Anstrengungsbereitschaft, Durchhaltevermögen und Zielorientierung zusammen (Bandura, 1993; Klassen & Usher, 2010; Pajares, Britner & Valiante, 2000). Sind Jugendliche der Meinung, dass sie nicht über die notwendigen Fähigkeiten und Fertigkeiten verfügen, um eine Aufgabe zu lösen, strengen sie sich beim Lösen der Aufgabe auch weniger an. Das führt dazu, dass im Sinne einer sich selbst erfüllenden Prophezeiung die Wahrscheinlichkeit sinkt, die Aufgabe tatsächlich richtig zu lösen. Dies beeinflusst wiederum ihre zukünftige Leistung – wenn sie die Aufgabe erneut nicht lösen, nimmt auch ihre Selbstwirksamkeitserwartung ab (Parker, Marsh, Ciarrochi, Marshall & Abduljabbar, 2014). Im naturwissenschaftlichen Unterricht hängen Selbstwirksamkeitserwartungen vor allem positiv mit prozeduralen Lernaktivitäten wie dem Experimentieren zusammen (Jansen, Scherer & Schroeders, 2015). Häufiges Experimentieren scheint die Selbstwirksamkeitserwartung von Jugendlichen zu fördern, wohingegen der umfassendere Prozess der naturwissenschaftlichen Erkenntnisgewinnung, der beispielsweise das eigenständige Planen von Experimenten umfasst, negativ mit der Selbstwirksamkeitserwartung zusammenhängt. Auch naturwissenschaftsbezogene Selbstwirksamkeitserwartungen hängen damit zusammen, dass sich Jugendliche in der Zukunft eher vorstellen können, in einem Beruf zu arbeiten, der etwas mit Naturwissenschaften zu tun hat (Jansen et al., 2015).

© Waxmann Verlag GmbH

Bei PISA 2006 war die Selbstwirksamkeitserwartung in Deutschland leicht überdurchschnittlich ausgeprägt. Jungen trauten sich dabei eher als Mädchen zu, konkrete naturwissenschaftliche Aufgaben zu lösen (Schütte et al., 2007).

3.2 Naturwissenschaftsbezogene Berufserwartungen

Im Alter von fünfzehn Jahren stellt die Frage nach dem zukünftigen Beruf bereits eine wichtige Entwicklungsaufgabe für Jugendliche in Deutschland dar. Sie besuchen häufig die neunte oder zehnte Klasse (vgl. Kapitel 1 und 5) und in den meisten nicht gymnasialen Schularten wird am Ende dieser Klassenstufen der erste allgemeinbildende (Hauptschulabschluss) oder mittlere Schulabschluss (Realschulabschluss) erworben. Die Schülerinnen und Schüler bewerben sich bereits lange vor dem Abschluss für eine Berufsausbildung oder entscheiden sich für weitere Ausbildungswege. Auch am Gymnasium stehen zukünftige Entscheidungen für die Wahl von Profilen oder Kursen in der gymnasialen Oberstufe an. Damit werden zu diesem Zeitpunkt bereits die ersten Weichen für oder gegen einen möglichen naturwissenschaftlichen Beruf gestellt.

Dass naturwissenschaftliche Karriereerwartungen die tatsächlichen Ausbildungs- und Berufswahlentscheidungen Jugendlicher beeinflussen, ist durch Längsschnittstudien gut belegt (z. B. Croll, 2008; Tai et al., 2006). Studien, die sich aus unterschiedlichen theoretischen Perspektiven mit der Entstehung von Berufserwartungen beschäftigen, finden Zusammenhänge vor allem mit Einstellungen zu Naturwissenschaften. Eine zentrale Rolle spielen dabei Freude und Interesse an den Naturwissenschaften (z. B. Taskinen, Schütte & Prenzel, 2013). Jugendliche, die Freude und Interesse an Naturwissenschaften erleben, beschäftigen sich gern mit naturwissenschaftlichen Themen und setzen sich eher mit möglichen naturwissenschaftlichen Berufsfeldern auseinander. Wenn sie eine Ausbildung oder einen Beruf wählen, der zu ihren Interessen passt, sind sie zufriedener und erfolgreicher in ihrer beruflichen Tätigkeit (vgl. Taskinen et al., 2013). Auch positive Einstellungen der Eltern gegenüber Naturwissenschaften können Erwägungen für einen solchen Beruf beeinflussen (z. B. DeWitt & Archer, 2015).

Betrachtet man, inwieweit die naturwissenschaftliche Kompetenz eine Rolle für die Berufserwartungen spielt, zeigen Studien, dass die tatsächliche Kompetenz – wie zum Beispiel bei PISA erfasst – nur einen geringen direkten Effekt hat (z. B. Kjaernsli & Li, 2011; Nagengast & Marsh, 2012; Taskinen, et al., 2008). Wichtiger scheint zu sein, ob die Jugendlichen glauben, dass sie den Anforderungen des Berufs auch gewachsen sind (Eccles, 2011). Dies spiegelt sich in ihren naturwissenschaftsbezogenen Selbstbildern wider – Vorstellungen, die Jugendliche von ihren Fähigkeiten haben und was sie sich selbst in Bezug auf die Naturwissenschaften zutrauen. Naturwissenschaften werden jedoch häufig eher als schwierig angesehen und deshalb unterschätzen Schülerinnen und Schüler oftmals ihre tatsächlichen Kompetenzen (Osborne et al., 2003). Das führt dazu, dass sie naturwissenschaftliche Berufe seltener in Betracht ziehen, obwohl sie die Fähigkeiten dazu mitbrächten.

© Waxmann Verlag GmbH

Auch die instrumentelle Motivation spielt eine wichtige Rolle dafür, ob Jugendliche sich vorstellen können, später einen naturwissenschaftlichen Beruf auszuüben (z. B. Kjaernsli & Li, 2011). Wenn Fünfzehnjährige zukünftig einem MINT-Beruf nachgehen wollen, ist ihnen eher bewusst, dass das Lernen in den naturwissenschaftlichen Fächern wichtig für ihr Berufsziel ist. Dies scheint insbesondere auf Berufe im Gesundheitsbereich zuzutreffen (Taskinen et al., 2008). Auch pragmatische Überlegungen wie Beschäftigungsaussichten, Sicherheit des Berufs und Verdienstmöglichkeiten tragen zur instrumentellen Motivation bei. Viele MINT-Berufe erfüllen diese Kriterien und stellen deshalb attraktive Berufsziele dar.

3.3 Die Erfassung naturwissenschaftsbezogener motivationaler Orientierungen, Selbstwirksamkeitserwartungen und Berufserwartungen in PISA 2015

3.3.1 Methodische Einschränkungen

Die PISA-Studie zielt auf den internationalen Vergleich motivationaler Orientierungen und Selbstwirksamkeitserwartungen ab. Dabei gilt es, einige methodische Einschränkungen zu berücksichtigen, die sich aus dem Design der Studie ergeben (vgl. Schiepe-Tiska et al., im Druck). So werden zum Beispiel relativ stabile Orientierungen und Selbstbilder erhoben. Außerdem zeigt die Forschung, dass Schülerinnen und Schüler Selbsteinschätzungen vor ihrem jeweiligen kulturellen Hintergrund treffen. Dies kann dazu führen, dass sie Aussagen grundsätzlich eher zustimmen, als sie abzulehnen, oder extreme Antwortkategorien grundsätzlich vermeiden (Johnson, Shavitt & Holbrook, 2011). Möglicherweise beantworten sie die Aussagen auch so, wie sie glauben, dass es von ihnen erwartet wird (soziale Erwünschtheit). Kulturspezifische Werte und Normen können demnach zu kulturbedingten Unterschieden im Antwortverhalten führen, die eine internationale Vergleichbarkeit beeinträchtigen (van de Gaer, Grisay, Schulz & Gebhardt, 2012). Davon abgesehen ist eine Selbsteinschätzung auch davon abhängig, welche Bezugsgruppe für den Vergleich ausgewählt wird (Heine, Lehmann, Peng & Greenholtz, 2002). Aus diesen Gründen werden in diesem Kapitel die Antworten der Fünfzehnjährigen im Vergleich zum OECD-Durchschnitt und zu einigen ausgewählten OECD-Staaten berichtet, deren Kulturräume gewisse Ähnlichkeiten mit Deutschland aufweisen (Schweiz, Niederlande, Finnland, Vereinigtes Königreich und Kanada).

Eine weitere Herausforderung besteht darin, dass die Naturwissenschaften – anders als beispielsweise die Mathematik – sehr unterschiedliche Bereiche umfassen wie etwa *lebende* oder *physikalische Systeme*. Damit spielt zusammen, dass in manchen Staaten Naturwissenschaften als integriertes Unterrichtsfach und in anderen Staaten – wie zum Großteil in Deutschland – getrennt nach Fächern wie Biologie, Physik und Chemie

© Waxmann Verlag GmbH

unterrichtet wird. Die Schülerinnen und Schüler geben demnach eine Wahrnehmung an, die unterschiedliche Themenbereiche und zum Teil auch Schulfächer abbildet.

Für die Erfassung der Berufserwartungen gilt es darüber hinaus zu berücksichtigen, dass ein offenes Antwortformat gewählt und die Antworten anschließend mithilfe der internationalen Berufsklassifikation ISCO-08 (International Standard Classification of Occupations; International Labour Office, 2012) kodiert wurden. Dies mag für die internationale Vergleichbarkeit die Validität der Erfassung erhöhen, allerdings muss beachtet werden, dass Jugendliche nur Berufe nennen können, die ihnen auch bekannt sind. Die Belastbarkeit der Aussagen kann außerdem dadurch eingeschränkt sein, dass Jugendliche eventuell keine realistische Vorstellung von den auszuführenden Tätigkeiten des Berufs haben. Deshalb müssen ihre Vorstellungen nicht unbedingt mit den realistischen Anforderungsprofilen übereinstimmen.

3.3.2 Eingesetzte Skalen

Freude und Interesse an Naturwissenschaften

Freude und Interesse an Naturwissenschaften wurde mittels fünf Aussagen erhoben (Cronbachs α = .95[1]): (a) „Im Allgemeinen macht es mir Spaß, mich mit naturwissenschaftlichen Themen zu befassen.“, (b) „Ich lese gerne etwas über Naturwissenschaften.“, (c) „Ich beschäftige mich gerne mit naturwissenschaftlichen Problemen.“, (d) „Ich eigne mir gerne neues Wissen in den Naturwissenschaften an.“ und (e) „Ich bin interessiert, Neues in den Naturwissenschaften zu lernen.“ (Frey et al., 2009). Die Einschätzungen erfolgten auf einer vierstufigen Antwortskala („stimme überhaupt nicht zu“, „stimme eher nicht zu“, „stimme eher zu“ und „stimme völlig zu“).

Interesse an naturwissenschaftlichen Themen

Mit insgesamt fünf Themengebieten wurden die Schülerinnen und Schüler aufgefordert, ihr Interesse an naturwissenschaftlichen Themen anzugeben (Cronbachs α = .76). Die Themengebiete waren (a) „Lebensräume (z. B. Ökosysteme, Nachhaltigkeit)“, (b) „Bewegung und Kräfte (z. B. Geschwindigkeit, Reibung, Magnetismus, Schwerkraft)“, (c) „Energie und ihre Umwandlung (z. B. Konservierung, chemische Reaktionen)“, (d) „Das Universum und seine Geschichte“ und (e) „Wie Naturwissenschaften uns helfen können, Krankheiten zu verhindern“. Sie spiegeln Beispiele der verschiedenen Wissenssysteme – *physikalische Systeme*, *lebende Systeme*, *Erd- und Weltraumsysteme* – naturwissenschaftlicher Kompetenz wider (vgl. Kapitel 2). Die Antwortkategorien umfassten eine fünfstufige Skala von „überhaupt nicht interessiert“, „kaum interessiert“, „eher interessiert“, „sehr interessiert“ bis „ich weiß nicht, was das ist“.

1 Cronbachs Alpha wurde jeweils auf der Basis der Stichprobe in Deutschland ermittelt.

© Waxmann Verlag GmbH

Instrumentelle Motivation

Wie Schülerinnen und Schüler ihre naturwissenschaftsbezogene instrumentelle Motivation einschätzen, wurde anhand von vier Aussagen erhoben (Cronbachs α = .92), welche die Jugendlichen auf einer vierstufigen Antwortskala („stimme völlig zu", „stimme eher zu", „stimme eher nicht zu" „stimme überhaupt nicht zu") beurteilten: (a) „Sich im Unterricht in naturwissenschaftlichen Fächern anzustrengen, zahlt sich aus, weil mir das bei der Arbeit, die ich später machen möchte, helfen wird.", (b) „Was ich in den naturwissenschaftlichen Fächern lerne, ist wichtig für mich, weil ich es für das brauche, was ich später machen möchte.", (c) „Für naturwissenschaftliche Unterrichtsfächer zu lernen, lohnt sich für mich, weil das Gelernte meine beruflichen Aussichten verbessern wird." und (d) „Viele Dinge, die ich in den naturwissenschaftlichen Unterrichtsfächern lerne, werden mir dabei helfen, einen Job zu bekommen." (Frey et al., 2009).

Selbstwirksamkeitserwartung

Um zu erfassen, ob sich Jugendliche hinsichtlich naturwissenschaftlicher Aufgabenstellungen als selbstwirksam erleben, wurden sie gebeten einzuschätzen, wie sehr sie sich die Bearbeitung verschiedener naturwissenschaftlicher Frage- und Problemstellungen zutrauen würden (Cronbachs α = .88). Diese sollten auf einer vierstufigen Antwortskala eingeschätzt werden („das wäre einfach für mich", „ich könnte das mit ein bisschen Mühe schaffen" „es würde mir schwerfallen, das allein zu schaffen", „das könnte ich nicht"). Die Frage- und Problemstellungen waren (a) „Die naturwissenschaftliche Fragestellung erkennen, die einem Zeitungsbericht über ein Gesundheitsthema zugrunde liegt.", (b) „Erklären, warum Erdbeben in manchen Gegenden häufiger vorkommen als in anderen.", (c) „Die Rolle der Antibiotika bei der Behandlung von Krankheiten beschreiben.", (d) „Wissenschaftliche Fragestellungen herausfinden, die mit der Müllentsorgung zusammenhängen.", (e) „Vorhersagen, wie Änderungen in der Natur das Überleben bestimmter Tierarten beeinflussen können.", (f) „Die wissenschaftlichen Informationen auf einem Lebensmitteletikett interpretieren.", (g) „Zeigen, wie neue Erkenntnisse zu einem neuen Verständnis über die Möglichkeit von Leben auf dem Mars führen können." sowie (h) „Die bessere von zwei Erklärungen über die Bildung von saurem Regen erkennen." (Frey et al., 2009).

Naturwissenschaftliche Berufserwartungen

Anhand der offenen Frage „Was meinst du, welchen Beruf du mit ca. 30 Jahren haben wirst?" wurde die Berufswahlneigung der Schülerinnen und Schüler erfasst. Die offenen Antworten wurden nach der internationalen Berufsklassifikation ISCO-08 (International Standard Classification of Occupations; International Labour Office, 2012) kodiert. Anschließend wurden die naturwissenschaftlichen Berufserwartungen in vier Kategorien eingeteilt: (a) „Naturwissenschaftler, Mathematiker, Ingenieure" (z. B. Architekten, Phy-

© Waxmann Verlag GmbH

siker, Ingenieure, Geologen), (b) „Gesundheitsberufe“ (z. B. Ärzte, Tierärzte, Krankenpflegekräfte, Hebammen, Physiotherapeuten, Apotheker), (c) „Fachkräfte in der Informations- und Kommunikationstechnologie“ (z. B. Softwareentwickler, Systemdesigner, Softwareingenieur, Webarchitekt), (d) „naturwissenschaftsbezogene Technikberufe“ (z. B. Biotechniker, Elektrotechniker, Technische Zeichner und Chemielaboranten; Frey et al., 2009).

3.4 Ergebnisse

Im Folgenden werden zunächst die motivationalen Orientierungen *Freude und Interesse an den Naturwissenschaften*, *Interesse an naturwissenschaftlichen Themen*, *instrumentelle Motivation* sowie *naturwissenschaftsbezogene Selbstwirksamkeitserwartungen* im Vergleich zu ausgewählten OECD-Staaten – Kanada, Vereinigtes Königreich, Finnland, Niederlande und Schweiz – dargestellt und Geschlechterdifferenzen betrachtet. Diese Staaten wurden ausgewählt, weil deren Kulturräume gewisse Ähnlichkeiten mit Deutschland aufweisen und demnach differenzielle Antwortverzerrungen aufgrund des kulturellen Hintergrunds weniger zu erwarten sind (vgl. Abschnitt 3.3). Außerdem zeichnen sich die Fünfzehnjährigen in diesen Staaten durch eine mit Deutschland vergleichbare oder höher ausgeprägte naturwissenschaftliche Kompetenz aus (vgl. Kapitel 2). Darüber hinaus gibt es in diesen Staaten unterschiedliche Unterrichtstraditionen, welche die Entwicklung motivationaler Orientierungen beeinflussen (vgl. Kapitel 4). Insgesamt wird der Frage nachgegangen, ob es im Vergleich zu anderen Staaten Besonderheiten in den motivational-affektiven Merkmalen der Schülerinnen und Schüler in Deutschland gibt. Vertiefende Analysen untersuchen Unterschiede zwischen diesen Merkmalen am Gymnasium und an nicht gymnasialen Schularten. Es schließt sich eine Betrachtung der Veränderung der Merkmale im Vergleich zu PISA 2006 an.

Im zweiten Teil werden die Berufserwartungen der Fünfzehnjährigen näher in den Blick genommen. Es wird untersucht, inwieweit Jugendliche in Erwägung ziehen, später einen naturwissenschaftsbezogenen Beruf auszuüben. Dabei wird auf Unterschiede zwischen Mädchen und Jungen sowie die Veränderung der Berufserwartungen im Vergleich zu PISA 2006 eingegangen. Weiterführende Analysen betrachten das Zusammenspiel kognitiver sowie motivational-affektiver Schülermerkmale mit den naturwissenschaftsbezogenen Berufserwartungen.

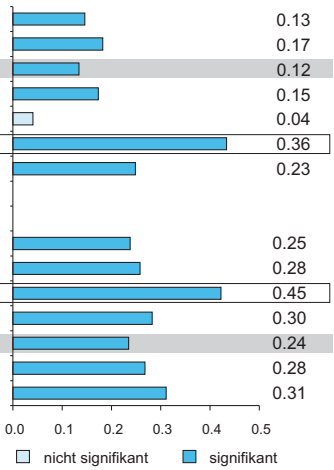
© Waxmann Verlag GmbH

3.4.1 Naturwissenschaftsbezogene motivationale Orientierungen und Selbstwirksamkeitserwartungen im internationalen Vergleich

Um die Ausprägungen der Schülermerkmale darzustellen, zeigen die folgenden Abbildungen für die ausgewählten Staaten einen am OECD-Mittelwert standardisierten Skalenwert, welcher auch einen Vergleich mit den Merkmalsausprägungen von PISA 2006 ermöglicht. Anhand dieser Skalenwerte ist ein Vergleich der Staaten untereinander möglich, es können jedoch keine Aussagen über die absolute Höhe der motivationalen Orientierungen und Selbstwirksamkeitserwartung getroffen werden. Deshalb wird zusätzlich für jedes Schülermerkmal ein charakteristisches Beispielitem und die dazugehörige prozentuale Zustimmung zu den entsprechenden Antwortkategorien angegeben, um das Antwortverhalten der Fünfzehnjährigen genauer in den Blick nehmen zu können. Die Items wurden so ausgewählt, dass sie eine möglichst gute Repräsentation der Verteilung der relativen Häufigkeitsangaben für die jeweilige Skala abbilden.

Freude und Interesse an den Naturwissenschaften

Schülermerkmale	Zustimmung in %[1]	OECD-Staaten	*M*	*(SE)*	*SD*	Effektstärke
Freude und Interesse Im Allgemeinen macht es mir Spaß, mich mit naturwissenschaftlichen Themen zu befassen.						
	74.9	Kanada	**0.40**	(0.01)	1.14	0.13
	66.9	Vereinigtes Königreich	**0.15**	(0.02)	1.06	0.17
	62.8	OECD-Durchschnitt	0.02	(0.00)	1.11	0.12
	66.1	Schweiz	-0.02	(0.02)	1.15	0.15
	64.3	Finnland	**-0.07**	(0.02)	1.01	0.04
	58.6	Deutschland	**-0.18**	(0.02)	1.21	0.36
	39.6	Niederlande	**-0.52**	(0.02)	1.07	0.23
Interesse an naturwissenschaftlichen Themen Wie Naturwissenschaften uns helfen können, Krankheiten zu verhindern.						
	76.0	Kanada	**0.26**	(0.01)	0.96	0.25
	73.5	Schweiz	**0.15**	(0.02)	0.92	0.28
	72.7	Deutschland	**0.04**	(0.02)	0.96	0.45
	74.3	Vereinigtes Königreich	0.01	(0.02)	0.96	0.30
	68.4	OECD-Durchschnitt	0.00	(0.00)	0.97	0.24
	64.6	Finnland	**-0.09**	(0.02)	0.97	0.28
	61.0	Niederlande	**-0.27**	(0.02)	1.00	0.31



Anmerkung: Statistisch signifikante Unterschiede zum OECD-Durchschnitt sind **fett** hervorgehoben.

[1] Die Antwortkategorien „stimme völlig zu“ und „stimme eher zu“ wurden zusammengefasst.

Abbildung 3.1: Freude und Interesse an Naturwissenschaften im internationalen Vergleich

Abbildung 3.1 zeigt zunächst die Skalenwerte für *Freude und Interesse an den Naturwissenschaften*. Im Vergleich zum OECD-Durchschnitt berichten Fünfzehnjährige in Deutschland signifikant weniger Freude und Interesse an den Naturwissenschaften. Dies trifft ebenfalls auf die Schülerinnen und Schüler in Finnland und den Niederlanden zu. Im Vereinigten Königreich und gerade in Kanada berichten die Jugendlichen

© Waxmann Verlag GmbH

hingegen über signifikant mehr Freude und Interesse an den Naturwissenschaften als Jugendliche im OECD-Durchschnitt.

Eine Betrachtung der Zustimmung des charakteristischen Items zeigt, dass in Deutschland 59 Prozent der Schülerinnen und Schüler angeben, sich mit naturwissenschaftlichen Themen zu beschäftigen, weil es ihnen Spaß macht. Dies bedeutet jedoch auch, dass rund 40 Prozent der Schülerinnen und Schüler im Allgemeinen wenig bis keine Freude an der Beschäftigung mit naturwissenschaftlichen Themen erleben. Im Gegensatz dazu stimmen in Kanada sogar drei Viertel der Fünfzehnjährigen zu, dass ihnen die Beschäftigung mit naturwissenschaftlichen Themen Spaß macht. In den Niederlanden sind es hingegen nur 40 Prozent.

Für Freude und Interesse an Naturwissenschaften lassen sich im OECD-Durchschnitt Geschlechterunterschiede zugunsten der Jungen finden. Auch in Deutschland und dem Großteil der ausgewählten Vergleichsstaaten geben Jungen an, mehr Freude und Interesse an Naturwissenschaften zu erleben als Mädchen. Dieser Unterschied ist in Deutschland im Vergleich zum OECD-Durchschnitt und den ausgewählten Staaten am stärksten ausgeprägt (d = 0.36). Lediglich in Finnland berichten Mädchen und Jungen über ähnlich viel Freude und Interesse an Naturwissenschaften.

Das *Interesse an naturwissenschaftlichen Themen* liegt in Deutschland leicht, aber dennoch signifikant über dem OECD-Durchschnitt (Abbildung 3.1). In der Schweiz und in Kanada ist das Interesse an naturwissenschaftlichen Themen noch einmal deutlich stärker ausgeprägt als in Deutschland. Ein unterdurchschnittliches Interesse berichten – wie auch bei Freude und Interesse an Naturwissenschaften – die Schülerinnen und Schüler in Finnland und besonders in den Niederlanden. Geschlechterdifferenzen zugunsten der Jungen finden sich sowohl im OECD-Durchschnitt als auch in allen ausgewählten Vergleichsstaaten. Auch in Deutschland interessieren sich die Jungen deutlich mehr für naturwissenschaftliche Themen als Mädchen. Wie bei der Freude an den Naturwissenschaften ist dieser Geschlechterunterschied in Deutschland am stärksten ausgeprägt (d = 0.45).

Inwieweit sich die Fünfzehnjährigen für die verschiedenen naturwissenschaftlichen Themen interessieren, zeigt Tabelle 3.1. Im OECD-Durchschnitt finden die Jugendlichen das Universum und seine Geschichte sowie wie Naturwissenschaften helfen können, Krankheiten zu verhindern, am interessantesten (beide 68 Prozent). Auch in Deutschland ist das Interesse der Schülerinnen und Schüler an diesen beiden Themen am größten. Wie Naturwissenschaften helfen können, Krankheiten zu verhindern, interessiert sogar fast drei Viertel der Fünfzehnjährigen in Deutschland. Ein vergleichsweise hohes Interesse zeigen auch die Jugendlichen in Kanada, dem Vereinigten Königreich und der Schweiz. Am wenigsten interessieren sich die Fünfzehnjährigen im OECD-Durchschnitt für Lebensräume – auch in Finnland, den Niederlanden, dem Vereinigten Königreich und der Schweiz geben die Jugendlichen bei diesem Thema das geringste Interesse an. In Deutschland ist das nicht der Fall. Hier gibt über die Hälfte der Jugendlichen an, sich für Ökosysteme und Nachhaltigkeit zu interessieren. In Deutschland interessieren sich die Fünfzehnjährigen hingegen am wenigsten für die Themen Bewegung und

© Waxmann Verlag GmbH

Tabelle 3.1: *Interesse an naturwissenschaftlichen Themen* im internationalen Vergleich

OECD-Staaten	Lebensräume (z. B. Ökosysteme, Nachhaltigkeit)		Bewegungen und Kräfte (z. B. Geschwindigkeit, Reibung, Magnetismus, Schwerkraft)		Energie und ihre Umwandlung (z. B. Konservierung, chemische Reaktionen)		Das Universum und seine Geschichte		Wie Naturwissenschaften uns helfen können, Krankheiten zu verhindern	
	Zustimmung in %[1]	*(SE)*	Zustimmung in %[1]	*(SE)*	Zustimmung in %[1]	*(SE)*	Zustimmung in %[1]	*(SE)*	Zustimmung in %[1]	*(SE)*
Deutschland	55.6	(0.8)	44.5	(0.8)	42.4	(0.9)	66.1	(0.9)	72.7	(0.9)
Finnland	28.8	(0.7)	45.7	(0.9)	44.9	(0.8)	66.0	(1.0)	64.5	(0.9)
Kanada	54.6	(0.6)	55.9	(0.5)	64.2	(0.6)	71.2	(0.6)	76.0	(0.5)
Niederlande	32.8	(0.8)	40.7	(0.7)	38.9	(0.6)	54.6	(0.9)	61.0	(0.7)
Schweiz	50.9	(0.9)	50.4	(0.8)	55.1	(0.9)	72.8	(0.8)	73.5	(0.8)
Vereinigtes Königreich	41.1	(0.8)	45.5	(0.8)	51.1	(0.8)	72.7	(0.7)	74.3	(0.8)
OECD-Durchschnitt	42.3	(0.1)	46.8	(0.1)	49.4	(0.1)	67.9	(0.1)	68.4	(0.1)

[1] Die Antwortkategorien „sehr interessiert“ und „eher interessiert“ wurden zusammengefasst.

Kräfte sowie Energie und ihre Umwandlung. Insgesamt scheint in Deutschland das Interesse für *lebende Systeme* (vgl. Kapitel 2) und auch für *Erd- und Weltraumsysteme* stärker ausgeprägt zu sein als das Interesse für *physikalische Systeme*, welche auch chemische Themen berücksichtigen. Auch in den Vergleichsstaaten – mit Ausnahme von Finnland – scheint das Interesse an *lebenden Systemen* größer zu sein als das an *physikalischen Systemen*. Der Unterschied ist in den anderen Vergleichsstaaten jedoch nicht so stark ausgeprägt wie in Deutschland.

Betrachtet man Unterschiede zwischen Mädchen und Jungen, zeigt sich, dass im OECD-Durchschnitt, in Deutschland sowie in allen Vergleichsstaaten Mädchen stärker daran interessiert sind, wie Naturwissenschaften helfen können, Krankheiten zu verhindern. In Deutschland stimmen dieser Aussage 77 Prozent der Mädchen und 68 Prozent der Jungen zu (siehe OECD, 2016b, für die prozentualen Angaben für Jungen und Mädchen in den Vergleichsstaaten). Jungen interessieren sich in all diesen Staaten signifikant stärker für die beiden Themen aus dem Bereich *physikalische Systeme*. In Deutschland beträgt die Geschlechterdifferenz für das Thema Energie und ihre Umwandlung 28 Prozent zugunsten der Jungen, für das Thema Bewegung und Kräfte sind es sogar 31 Prozent. In keinem anderen der ausgewählten Vergleichsstaaten ist dieser Unterschied so stark ausgeprägt. Diese Geschlechterdifferenz spiegelt sich auch in der naturwissenschaftlichen Kompetenz in den Wissenssystemen wider. Sind bei PISA 2015 in Deutschland keine Kompetenzunterschiede für die *lebenden Systeme* zu finden, ist die naturwissenschaftliche Kompetenz in den *physikalischen Systemen* bei Jungen deutlich höher ausgeprägt als bei Mädchen (vgl. Kapitel 2).

© Waxmann Verlag GmbH

Instrumentelle Motivation und Selbstwirksamkeitserwartungen

Abbildung 3.2 zeigt die Skalenwerte der instrumentellen Motivation und der naturwissenschaftlichen Selbstwirksamkeitserwartung im internationalen Vergleich. Die *instrumentelle Motivation* ist in Deutschland signifikant niedriger ausgeprägt als im OECD-Durchschnitt. Etwas mehr als die Hälfte der Schülerinnen und Schüler in Deutschland stimmt der Aussage zu, dass es sich auszahlt, sich in naturwissenschaftlichen Fächern anzustrengen, weil es ihnen bei der späteren Arbeit helfen wird. Damit ist der Hälfte der Jugendlichen nicht bewusst, dass die Naturwissenschaften eine wichtige Grundlage für ihr weiteres Ausbildungs- und Berufsleben bilden können. Der Vergleich mit den ausgewählten Staaten zeigt ein ähnliches Bild wie bei Freude und Interesse an den Naturwissenschaften: Die Jugendlichen in der Schweiz und den Niederlanden berichten über eine ähnlich ausgeprägte instrumentelle Motivation wie Deutschland und liegen damit ebenfalls unter dem OECD-Durchschnitt. Das Vereinigte Königreich und Kanada erzielen Werte deutlich über dem OECD-Durchschnitt. Sowohl in Kanada als auch im Vereinigten Königreich sind circa 80 Prozent der Schülerinnen und Schüler der Meinung, dass es sich für die spätere Arbeit lohnt, sich in naturwissenschaftlichen Fächern anzustrengen.

Die instrumentelle Motivation ist im OECD-Durchschnitt bei Jungen stärker ausgeprägt als bei Mädchen. Auch in Deutschland und den Niederlanden schätzen die Jungen Naturwissenschaften für ihr späteres Ausbildungs- und Berufsleben nützlicher ein als die Mädchen. Im Unterschied dazu sind in Kanada Mädchen höher instrumentell motiviert als Jungen. Erneut ist der Geschlechterunterschied am stärksten in Deutschland ausgeprägt ($d = 0.27$). Dass Geschlechterunterschiede in der Einschätzung der Nützlichkeit von Naturwissenschaften nicht zwingend sind, zeigen das Vereinigte Königreich, Finnland und die Schweiz. Hier finden sich keine Unterschiede zwischen Jungen und Mädchen.

Die *naturwissenschaftliche Selbstwirksamkeitserwartung* erfasst, wie sehr sich Schülerinnen und Schüler zutrauen, bestimmte Aufgabenstellungen in den Naturwissenschaften erfolgreich zu lösen. Im Vergleich zum OECD-Durchschnitt ist die Selbstwirksamkeitserwartung in Deutschland leicht unterdurchschnittlich ausgeprägt. Für die Aufgabe „die wissenschaftliche Information auf einem Lebensmitteletikett interpretieren“ gab nur ein knappes Fünftel der Schülerinnen und Schüler an, diese lösen zu können. Während in der Schweiz die Selbstwirksamkeitserwartung noch signifikant niedriger ausgeprägt ist als in Deutschland, berichten Schülerinnen und Schüler sowohl im Vereinigten Königreich als auch in Kanada eine deutlich höhere Selbstwirksamkeit, die jeweils über dem OECD-Durchschnitt liegt. Dementsprechend trauen sich in Kanada auch rund ein Viertel der Schülerinnen und Schüler die Interpretation der Informationen auf einem Lebensmitteletikett zu, während sich in der Schweiz und den Niederlanden ebenso wie in Deutschland mehr als 80 % der Schülerinnen und Schüler dies nicht zutrauen.

Ein klares Bild zeigt sich beim Blick auf die Geschlechterunterschiede: Jungen haben in allen betrachteten Staaten eine signifikant höhere Selbstwirksamkeitserwartung als

© Waxmann Verlag GmbH

Schülermerkmale	Zustimmung in %[1]	OECD-Staaten	*M*	*(SE)*	*SD*	Differenz Jungen – Mädchen	Effektstärke
Instrumentelle Motivation Sich im Unterricht in naturwissenschaftlichen Fächern anzustrengen, zahlt sich aus, weil mir das bei der Arbeit, die ich später machen möchte, helfen wird.	81.0	Kanada	**0.46**	(0.01)	0.99		-0.05
	79.7	Vereinigtes Königreich	**0.38**	(0.02)	0.97		0.04
	65.0	Finnland	0.16	(0.02)	0.92		-0.05
	68.8	OCED-Durchschnitt	0.14	(0.00)	0.98		0.04
	54.9	Niederlande	**-0.21**	(0.02)	1.06		0.14
	54.4	Deutschland	**-0.24**	(0.02)	1.04		0.27
	54.5	Schweiz	**-0.25**	(0.02)	1.04		-0.03
Selbstwirksamkeitserwartung Die wissenschaftliche Information auf einem Lebensmitteletikett interpretieren.	25.1	Kanada	**0.35**	(0.02)	1.29		0.21
	19.5	Vereinigtes Königreich	**0.27**	(0.02)	1.22		0.21
	20.0	OCED-Durchschnitt	0.04	(0.00)	1.23		0.16
	16.6	Deutschland	**-0.01**	(0.02)	1.19		0.31
	19.9	Finnland	**-0.04**	(0.02)	1.18		0.22
	15.5	Niederlande	**-0.08**	(0.02)	1.27		0.21
	14.2	Schweiz	**-0.17**	(0.02)	1.22		0.20

Anmerkung: Statistisch signifikante Unterschiede zum OECD-Durchschnitt sind **fett** hervorgehoben.

[1] Die Antwortkategorien „stimme völlig zu" und „stimme eher zu" wurden zusammengefasst.

Abbildung 3.2: Naturwissenschaftsbezogene instrumentelle Motivation und Selbstwirksamkeitserwartungen im internationalen Vergleich

Mädchen. Auch hier ist der Effekt in Deutschland vergleichsweise stark ausgeprägt (d = 0.31), im Falle der Selbstwirksamkeitserwartung jedoch ähnlicher zu den Vergleichsstaaten als bei Freude und Interesse und instrumenteller Motivation.

In der Zusammenschau ergibt sich für Deutschland im internationalen Vergleich ein aufschlussreiches Bild. Erleben die Jugendlichen wenig Freude und Interesse an den Naturwissenschaften allgemein, interessieren sie sich doch überdurchschnittlich für die verschiedenen naturwissenschaftlichen Themen. Dies legt die Vermutung nahe, dass die Schülerinnen und Schüler Naturwissenschaften, wie sie sie aus der Schule kennen, anders wahrnehmen als naturwissenschaftliche Themen allgemein. Sobald es einen thematischen Bezug gibt, scheinen Naturwissenschaften für die Fünfzehnjährigen in Deutschland – und auch in der Schweiz – interessanter zu werden. Die instrumentelle Motivation liegt ebenfalls unterhalb des OECD-Durchschnitts. Nur der Hälfte der Jugendlichen ist bewusst, dass Naturwissenschaften für ihr späteres Leben wichtig sein werden. Auch in dem, was sie sich in Bezug auf die Naturwissenschaften zutrauen, liegt deutliches Verbesserungspotenzial.

International zeigt sich ein ähnliches Bild nur in den Niederlanden. In Kanada hingegen – einem Bildungssystem, das sich durch besonders stark ausgeprägte naturwissenschaftliche Kompetenz auszeichnet – sind die motivationalen Orientierungen durchweg positiv ausgeprägt. Dies ist zum Großteil auch im Vereinigten Königreich der Fall. In diesen Staaten wird im naturwissenschaftlichen Unterricht sowohl Wert auf die kognitive Anregung („*Minds-on*") der Schülerinnen und Schüler gelegt als auch häufig expe-

© Waxmann Verlag GmbH

rimentiert („*Hands-on*“, vgl. Kapitel 4). Ein solcher Unterricht scheint gerade die positive Ausprägung motivational-affektiver Lernergebnisse zu fördern, ohne zwangsläufig mit geringerer Leistung einhergehen zu müssen. In Finnland hingegen – einem ebenfalls kompetenzstarken Bildungssystem – zeichnet sich der Unterricht eher durch lehrerzentrierte Klassengespräche aus. Es werden zwar durchschnittlich oft strukturierte Laborexperimente durchgeführt, allerdings deutlich seltener Ideen ausgetestet und Experimente selbst entwickelt. Dies scheint vor allem zulasten der Freude und des Interesses an Naturwissenschaften und des Interesses an naturwissenschaftlichen Themen zu gehen.

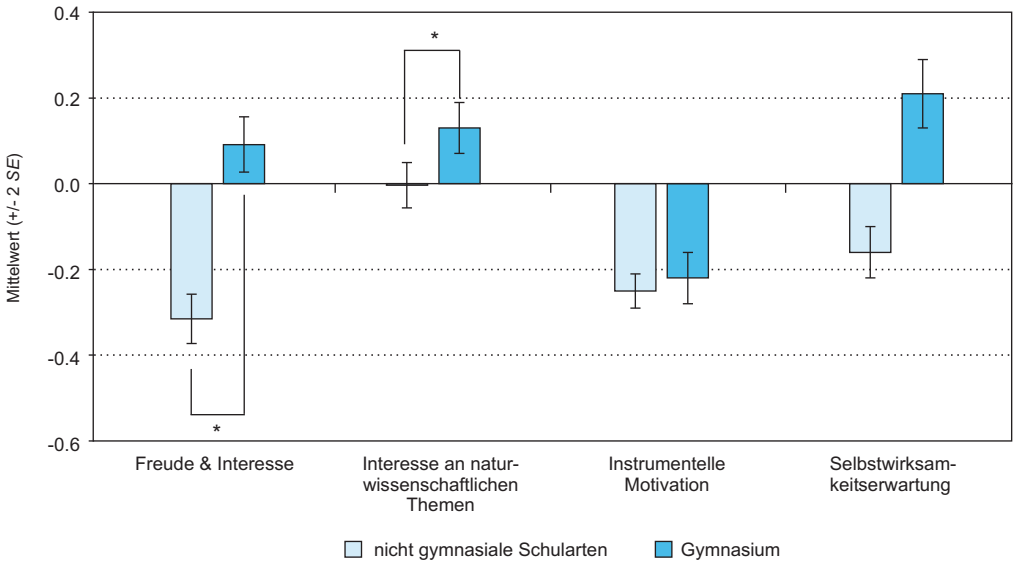
Die Geschlechterunterschiede in den betrachteten Merkmalen sind in Deutschland besonders auffällig und weisen die vergleichsweise höchsten Effektstärken auf. Jungen berichten im Vergleich zu Mädchen deutlich höher ausgeprägte motivationale Orientierungen und eine höhere Selbstwirksamkeitserwartung. Auch in den ausgewählten Vergleichsstaaten geben die Jungen mehr Freude und Interesse an den Naturwissenschaften sowie eine höher ausgeprägte Selbstwirksamkeitserwartung an als die Mädchen. Hier zeichnet sich also nicht nur in Deutschland Förderungspotenzial ab. Auch Mädchen für Naturwissenschaften zu begeistern, ihnen die Bedeutsamkeit für das Leben zu verdeutlichen und ihr naturwissenschaftliches Selbstbild zu stärken – darauf sollte im naturwissenschaftlichen Unterricht ein besonderes Augenmerk gelegt werden (vgl. Labudde & Möller, 2012).

3.4.2 Schulartspezifische Unterschiede in den naturwissenschaftsbezogenen motivationalen Orientierungen und der Selbstwirksamkeitserwartung in Deutschland

Einen differenzierteren Blick auf die motivationalen Orientierungen und die Selbstwirksamkeitserwartung Fünfzehnjähriger in Deutschland ermöglicht auch die Betrachtung getrennt nach Schularten. Dabei wird nach Gymnasien und nicht gymnasialen Schularten unterschieden. Zahlreiche Schulen mit einer Sekundarstufe I wurden in den letzten Jahren durch die Kombination von Bildungsgängen in neue oder neu bezeichnete Schularten überführt (Pant et al., 2013). Daher ist eine differenzierte Betrachtung der nicht gymnasialen, allgemeinbildenden Schularten nicht mehr tragfähig und sinnvoll (vgl. auch Kapitel 1).

Abbildung 3.3 zeigt, dass Jugendliche am Gymnasium sowohl mehr Freude als auch mehr Interesse an naturwissenschaftlichen Themen berichten als Fünfzehnjährige an nicht gymnasialen Schularten. Gerade die Freude an Naturwissenschaften ist bei Schülerinnen und Schülern nicht gymnasialer Schularten besonders gering ausgeprägt. Auch für die Selbstwirksamkeitserwartung findet sich dieser Unterschied, der ebenfalls zugunsten der Jugendlichen am Gymnasium ausfällt. Einzig für die instrumentelle Motivation – wie nützlich Fünfzehnjährige Naturwissenschaften für ihren späteren Beruf einschätzen – zeigen sich keine signifikanten Unterschiede zwischen den Schularten.

© Waxmann Verlag GmbH



* Schularten unterscheiden sich signifikant ($p < .05$).

Abbildung 3.3: Motivationale Orientierungen und Selbstwirksamkeitserwartung differenziert nach Schulart

3.4.3 Veränderungen der Schülermerkmale zwischen PISA 2006 und PISA 2015

Bei PISA 2015 stehen nach PISA 2006 zum zweiten Mal die Naturwissenschaften als Hauptdomäne im Mittelpunkt der Betrachtung. Demnach können für einige Merkmale Veränderungen im Vergleich zu 2006 untersucht werden. Abbildung 3.4 zeigt die Entwicklung von Freude und Interesse an Naturwissenschaften, instrumenteller Motivation und Selbstwirksamkeitserwartung zwischen 2006 und 2015.

Die Ergebnisse zeichnen für Deutschland ein problematisches Bild: Sowohl Freude und Interesse an Naturwissenschaften als auch die instrumentelle Motivation sind im Vergleich zu PISA 2006 signifikant gesunken. Die Freude hat vor allem bei den Mädchen abgenommen (von M_{2006} = -0.17 auf M_{2015} = -0.39); für die Jungen zeigt sich keine Veränderung (M_{2006} = -0.01, M_{2015} = 0.05). Auch für die instrumentelle Motivation ist insgesamt eine Abnahme zu verzeichnen. Sowohl Mädchen (M_{2006} = -0.16, M_{2015} = -0.38) als auch Jungen (M_{2006} = 0.01, M_{2015} = -0.10) geben bei PISA 2015 im Vergleich zu PISA 2006 an, dass sie Naturwissenschaften weniger nützlich für ihr späteres Leben finden. Auch die Selbstwirksamkeitserwartung der Fünfzehnjährigen ist bei PISA 2015 signifikant geringer als sie es noch bei PISA 2006 war. Dieser Rückgang wird nur bei den Mädchen sichtbar (M_{2006} = -0.01, M_{2015} = -0.19). Sie trauen sich in Bezug auf das Erkennen, Beschreiben und Erklären naturwissenschaftlicher Phänomene weniger zu als 2006. Bei den Jungen zeigt sich keine signifikante Veränderung (M_{2006} = 0.13, M_{2015} = 0.18).

© Waxmann Verlag GmbH

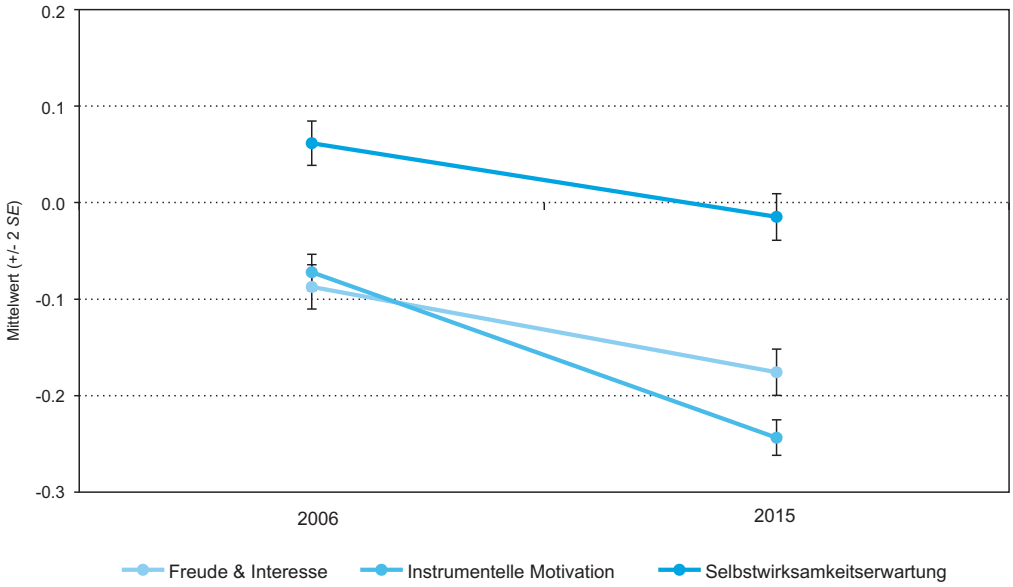


Abbildung 3.4: Veränderungen der motivationalen Orientierungen und Selbstwirksamkeitserwartungen in Deutschland zwischen PISA 2006 und PISA 2015

Ein Blick auf die Schularten offenbart, dass Freude und Interesse an den Naturwissenschaften am Gymnasium und an den nicht gymnasialen Schularten (beides mit einer Differenz von -0.11) im Vergleich zu PISA 2006 abgenommen hat. Auch die instrumentelle Motivation ist in beiden Schularten bei PISA 2015 gesunken (Differenz am Gymnasium von -0.23, an nicht gymnasialen Schularten von -0.16). Für die Selbstwirksamkeitserwartung zeigt sich eine Abnahme nur am Gymnasium (Differenz von -0.22). Hier fühlen sich die Fünfzehnjährigen weniger als bei PISA 2006 in der Lage, naturwissenschaftliche Phänomene zu erkennen, zu beschreiben und zu erklären. In Anbetracht dessen, dass ein enger Zusammenhang gerade zwischen Selbstwirksamkeitserwartung und naturwissenschaftlicher Kompetenz besteht (Jansen et al, 2015; Parker et al., 2014), ist dieser Befund besorgniserregend, spiegelt aber auch wider, dass sich zwischen PISA 2006 und PISA 2015 die naturwissenschaftliche Kompetenz der Jugendlichen am Gymnasium verringert hat (vgl. Kapitel 2).

Mit diesem Rückgang steht Deutschland nicht allein da. Auch in den Niederlanden berichten die Fünfzehnjährigen bei PISA 2015 weniger Freude und Interesse (Differenz von -0.20) und geringere Selbstwirksamkeitserwartung (Differenz von -0.10). In der Schweiz lassen sich insgesamt keine Veränderungen in den motivationalen Orientierungen und der Selbstwirksamkeitserwartung ausmachen. In Finnland, einem der Spitzenstaaten in der naturwissenschaftlichen Kompetenz, ergibt sich ein gemischtes Bild: Die instrumentelle Motivation ist angestiegen. Sowohl Mädchen als auch Jungen berichten in PISA 2015 häufiger, dass Naturwissenschaften für ihr späteres Leben und den Beruf nützlich sind (Differenz von +0.37). Jedoch ist die Freude an den Naturwissenschaf-

© Waxmann Verlag GmbH

ten sowohl bei den Mädchen als auch bei den Jungen signifikant gesunken (Differenz von -0.19). Auch die Selbstwirksamkeitserwartungen haben etwas abgenommen (Differenz von -0.09). Gerade die Mädchen trauen sich in Bezug auf das Erkennen, Beschreiben und Erklären naturwissenschaftlicher Phänomene weniger zu als bei PISA 2006. In Kanada hingegen – ebenfalls einer der kompetenzstärksten OECD-Staaten in den Naturwissenschaften (vgl. Kapitel 2) – lässt sich eine signifikante Zunahme für alle drei motivational-affektiven Lernergebnisse im Vergleich zu PISA 2006 feststellen (Freude und Interesse: Differenz von +0.23, instrumentelle Motivation: Differenz von +0.14, Selbstwirksamkeitserwartung: Differenz von +0.14). Dies ist sowohl für die Mädchen als auch für die Jungen der Fall. Auch im Vereinigten Königreich berichten die Jugendlichen – Mädchen und Jungen – einen signifikanten Anstieg in allen drei Merkmalen (Freude und Interesse: Differenz von +0.23, instrumentelle Motivation: Differenz von +0.21, Selbstwirksamkeitserwartung: Differenz von +0.09).

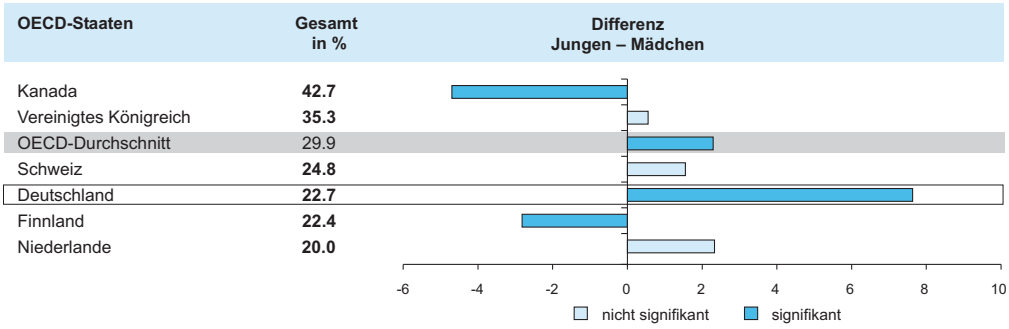
3.4.4 Naturwissenschaftsbezogene Berufserwartungen im internationalen Vergleich

Gerade mit Blick auf die hohe wirtschaftliche Bedeutung naturwissenschaftlich-technischer Berufe kommt der Betrachtung der naturwissenschaftsbezogenen Berufserwartungen eine bedeutende Rolle zu. Im Folgenden wird der Frage nachgegangen, inwieweit sich Schülerinnen und Schüler vorstellen können, einen naturwissenschaftlichen Beruf mit ca. 30 Jahren in Betracht zu ziehen.

In Deutschland geben 23 Prozent der Fünfzehnjährigen an, dass sie einen naturwissenschaftsbezogenen Beruf in Erwägung ziehen (Abbildung 3.5). Dieser Wert liegt signifikant unter dem OECD-Durchschnitt – im Mittel können sich 30 Prozent der Jugendlichen vorstellen, mit 30 Jahren einen solchen Beruf auszuüben. Auch in der Schweiz, in Finnland und den Niederlanden können sich im Vergleich zum OECD-Durchschnitt weniger Fünfzehnjährige vorstellen, einen naturwissenschaftsbezogenen Beruf zu ergreifen: Ähnlich wie in Deutschland gibt dies ein Fünftel bis ein Viertel der Schülerinnen und Schüler an. Im Vereinigten Königreich und Kanada hingegen geben überdurchschnittlich viele Fünfzehnjährige an, dass sie sich die Ausübung eines solchen Berufs vorstellen können. In Kanada können sich dies sogar über 40 Prozent der Jugendlichen vorstellen. In diesem Bildungssystem scheint es besonders gut zu gelingen, Schülerinnen und Schüler für naturwissenschaftsbezogene Berufe zu begeistern. Da Fünfzehnjährige in Kanada sich auch durch hohe naturwissenschaftliche Kompetenzen und motivationale Orientierungen auszeichnen, scheinen optimale Voraussetzungen für hochqualifizierte Nachwuchskräfte in den Naturwissenschaften gegeben zu sein. In Deutschland zeichnet sich hier Nachholbedarf ab.

In Deutschland ziehen wie auch im OECD-Durchschnitt signifikant mehr Jungen einen naturwissenschaftsbezogenen Beruf in Erwägung als Mädchen. In Deutschland können sich 27 Prozent der Jungen, aber nur 18 Prozent der Mädchen vorstellen, mit

© Waxmann Verlag GmbH



Anmerkung: Statistisch signifikante Unterschiede zum OECD-Durchschnitt sind **fett** hervorgehoben.

Abbildung 3.5: Naturwissenschaftsbezogene Berufserwartungen im internationalen Vergleich

30 Jahren einen solchen Beruf auszuüben. In allen anderen Vergleichsstaaten wollen mehr Mädchen einen naturwissenschaftsbezogenen Beruf ausüben als in Deutschland (zwischen 24 Prozent in Finnland und 45 Prozent in Kanada). Eine Geschlechterdifferenz zugunsten der Jungen für naturwissenschaftsbezogene Berufe muss jedoch nicht selbstverständlich sein. Dies verdeutlichen das Vereinigte Königreich, die Schweiz und die Niederlande. In diesen Staaten sind keine Geschlechterunterschiede zu finden. In Kanada und Finnland sieht das Bild umgekehrt aus. In diesen Bildungssystemen ziehen mehr Mädchen als Jungen einen naturwissenschaftsbezogenen Beruf in Betracht.

Berufserwartungen nach Arten naturwissenschaftsbezogener Berufe

Ein genauerer Blick auf die verschiedenen Arten naturwissenschaftsbezogener Berufe in Abbildung 3.6 verdeutlicht, dass die Fünfzehnjährigen in Deutschland sich vor allem vorstellen können, als Naturwissenschaftler, Mathematiker und Ingenieure zu arbeiten. Diese Kategorie wird von fast der Hälfte der Jugendlichen in Betracht gezogen. Ein knappes Drittel gibt Berufe aus dem Gesundheitsbereich an. Einen Beruf im Bereich der Informations- und Kommunikationstechnik wie z. B. Softwareentwickler oder Webarchitekt geben 18 Prozent der Schülerinnen und Schüler an, 7 Prozent können sich einen naturwissenschaftsbezogenen Technikberuf wie z. B. Elektrotechniker oder Chemielaborant vorstellen.

Im Vergleich zum OECD-Durchschnitt und den ausgewählten Vergleichsstaaten fällt die Verteilung in Deutschland dadurch auf, dass in den anderen Staaten am häufigsten Gesundheitsberufe als potenzielle Berufserwartungen angegeben werden. Die Werte liegen zwischen 41 Prozent in der Schweiz und 63 Prozent in Finnland (OECD-Durchschnitt: 50 Prozent). Naturwissenschaftler, Mathematiker und Ingenieure ziehen ein knappes Viertel (Finnland) bis etwas mehr als ein Drittel (Kanada) der Fünfzehnjährigen in Betracht (OECD-Durchschnitt: 35 Prozent) – außer im Vereinigten Königreich, da sind es ähnlich wie in Deutschland 44 Prozent. Diese beiden Berufskategorien zusammengenommen werden von 73 Prozent in der Schweiz bis zu 90 Prozent der

© Waxmann Verlag GmbH

Schülerinnen und Schüler im Vereinigten Königreich angegeben und machen damit den Großteil der Berufsvorstellungen aus. Berufe aus dem Bereich der Informations- und Kommunikationstechnik – ein Bereich, in dem in der zunehmend digitalisierten Welt Nachwuchskräfte benötigt werden – scheint für Fünfzehnjährige wenig vorstellbar zu sein. Im OECD-Durchschnitt ziehen 11 Prozent einen solchen Beruf in Erwägung. In den ausgewählten Vergleichsstaaten liegen die Werte alle unter dem von Deutschland und schwanken zwischen 6 Prozent in Kanada und 12 Prozent in der Schweiz. Hier scheinen die Jugendlichen in Deutschland bereits besser über solche Berufe informiert zu sein und interessieren sich für eine Karriere in diesem Bereich. Naturwissenschaftsbezogene Technikberufe werden in fast allen Vergleichsstaaten am seltensten angegeben. Die Werte liegen zwischen 1 Prozent im Vereinigten Königreich und vergleichbar mit Deutschland – 8 Prozent in den Niederlanden (OECD-Durchschnitt 4 Prozent). Hier fällt besonders die Schweiz auf, wo 15 Prozent der Jugendlichen angeben, sich vorstellen zu können, mit 30 Jahren einen solchen Beruf auszuüben.

Betrachtet man die Häufigkeiten der verschiedenen Kategorien naturwissenschaftlichen Berufe getrennt für Mädchen und Jungen, offenbaren sich in Deutschland interessante Unterschiede (Abbildung 3.6). Gibt die Hälfte der Jungen an, dass sie sich vorstellen können, Naturwissenschaftler, Mathematiker oder Ingenieure zu werden, zieht etwa ein Drittel der Mädchen diese Berufe in Betracht. Bei den Mädchen wiederum kann sich die Hälfte vorstellen, einen Beruf im Gesundheitsbereich auszuüben, wohingegen dieser Berufswunsch nur von 14 Prozent der Jungen angegeben wird. Dies steht im Einklang mit den Ergebnissen zum Interesse an naturwissenschaftlichen Themen, bei denen die Mädchen über ein höheres Interesse daran berichten, wie Naturwissenschaften hel-

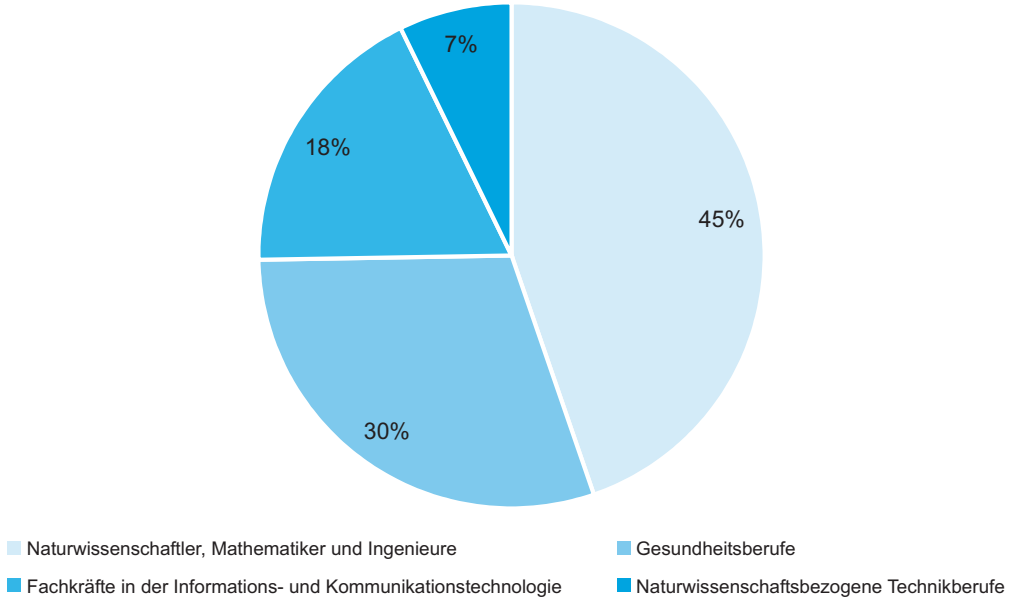


Abbildung 3.6: Verteilung der Arten naturwissenschaftsbezogener Berufe in Deutschland

© Waxmann Verlag GmbH

fen können, Krankheiten zu verhindern (vgl. Abschnitt 4.4.1). Bedeutende Unterschiede zeigen sich auch für den Bereich der Informations- und Kommunikationstechnik. Während 29 Prozent der Jungen diese zukunftsträchtigen Berufe in Erwägung ziehen, sind es nur 3 Prozent der Mädchen. Hier scheint es in Deutschland notwendig, gerade Mädchen stärker über solche Berufschancen zu informieren und ihr Interesse daran zu wecken. Bei den naturwissenschaftsbezogenen Technikberufen unterscheiden sich Mädchen und Jungen nicht signifikant voneinander. Diese sind jedoch mit 9 bzw. 6 Prozent eher gering ausgeprägt. Das mag darauf zurückzuführen sein, dass Technik als Inhalt im schulischen Fächerkanon wenig präsent ist und Jugendliche solche Berufe als wenig gesellschaftlich sinnvoll und mit monotonen Aufgaben assoziieren sowie davon ausgehen, dass man dabei kaum mit Menschen zu tun hat (vgl. acatech, 2015). Dies ist jedoch in vielen technischen Berufen nicht mehr zwangsläufig die Realität und auch die Verdienstmöglichkeiten und Karriereperspektiven gerade einer technischen Ausbildung werden vielfach unterschätzt.

Ein ähnliches Bild wie in Deutschland zeigt sich in den Vergleichsstaaten. Auch hier interessieren sich Mädchen vor allem für Berufe im Gesundheitsbereich, Jungen geben eher naturwissenschafts- und ingenieursbezogene Berufe an. Allerdings sind die Geschlechterpräferenzen deutlich stärker ausgeprägt. Zwischen 66 Prozent der Mädchen in der Schweiz und fast 90 Prozent der Mädchen in Finnland ziehen einen Beruf im Gesundheitsbereich in Betracht, bei den Jungen sind es zwischen 18 Prozent in der Schweiz und 31 Prozent in Finnland. Naturwissenschaftler, Mathematiker und Ingeni-

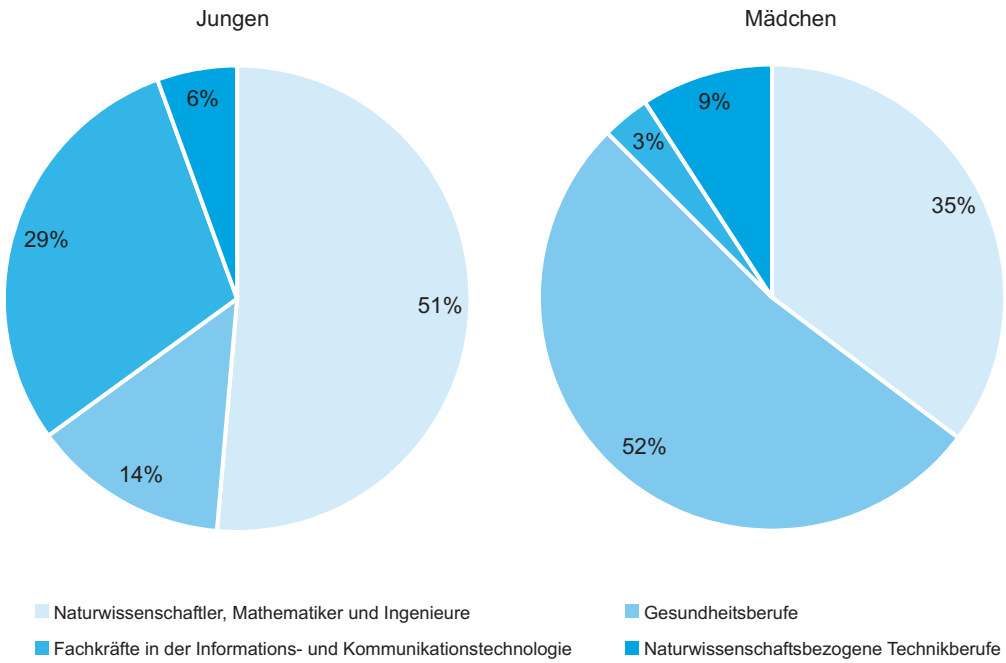


Abbildung 3.7: Verteilung der Arten naturwissenschaftsbezogener Berufe in Deutschland differenziert nach Geschlecht

© Waxmann Verlag GmbH

eure zu werden, erwägen zwischen 42 Prozent der Jungen in der Schweiz und 59 Prozent der Jungen im Vereinigten Königreich. Bei den Mädchen geben zwischen 7 Prozent in der Schweiz und 28 Prozent im Vereinigten Königreich einen solchen Beruf an. Berufe im Bereich der Informations- und Kommunikationstechnik werden von den Mädchen in den Vergleichsstaaten fast gar nicht angegeben (1 bis 2 Prozent). Ähnlich wie in Deutschland sind es vor allem die Jungen, die einen Beruf in diesem für die Zukunft gute Chancen bietenden Bereich in Erwägung ziehen. Zwischen 12 Prozent der Jungen in Kanada und 22 Prozent der Jungen in der Schweiz können sich einen solchen Beruf vorstellen. Für die naturwissenschaftsbezogenen Technikberufe zeigen sich wie in Deutschland keine Geschlechterunterschiede.

3.4.5 Naturwissenschaftsbezogene Berufserwartungen in Deutschland differenziert nach Schularten

Die 23 Prozent der Jugendlichen in Deutschland, die einen naturwissenschaftlichen Beruf für ihr späteres Leben in Erwägung ziehen, verteilen sich zu 12 Prozent auf das Gymnasium und 11 Prozent auf nicht gymnasiale Schularten. Betrachtet man diese Anteile genauer im Hinblick auf die in Erwägung gezogenen Berufskategorien (Abbildung 3.8), zeigt sich, dass knapp die Hälfte der Jugendlichen am Gymnasium in Erwägung zieht, Naturwissenschaftler, Mathematiker oder Ingenieure zu werden. In nicht gymnasialen Schularten sind es mit 39 Prozent etwas weniger. Schülerinnen und Schüler am Gymnasium geben mit 33 Prozent auch etwas häufiger an, dass sie sich einen

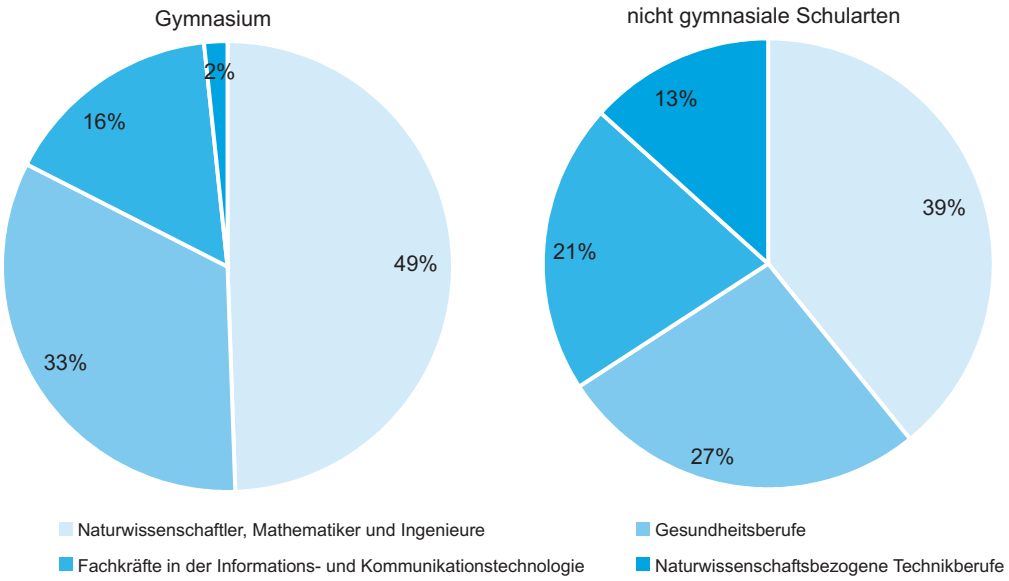


Abbildung 3.8: Verteilung der Arten naturwissenschaftsbezogener Berufe in Deutschland differenziert nach Schulart

© Waxmann Verlag GmbH

Beruf im Gesundheitsbereich vorstellen können, als Fünfzehnjährige an nicht gymnasialen Schularten (27 Prozent). Ein deutlicher Unterschied zugunsten der nicht gymnasialen Schularten resultiert für die naturwissenschaftsbezogenen Technikberufe. Diese werden von 13 Prozent der Jugendlichen an nicht gymnasialen Schularten, aber nur von 2 Prozent der Gymnasiastinnen und Gymnasiasten angegeben. Für den Bereich der Informations- und Kommunikationstechnik finden sich hingegen keine signifikanten Unterschiede. Diese scheinen für Fünfzehnjährige am Gymnasium und an nicht gymnasialen Schularten ähnlich attraktiv zu sein.

3.4.6 Veränderungen der naturwissenschaftsbezogenen Berufserwartungen zwischen PISA 2006 und PISA 2015

Da die Jugendlichen auch bereits bei PISA 2006, als die Naturwissenschaften das letzte Mal die Hauptdomäne waren, danach gefragt wurden, welchen Beruf sie sich mit ca. 30 Jahren vorstellen können, kann betrachtet werden, wie sich die naturwissenschaftsbezogene Berufswahlneigung verändert hat (Abbildung 3.9). In Deutschland geben bei PISA 2015 4 Prozent (gerundet) mehr Jugendliche als noch bei PISA 2006 an, dass sie einen naturwissenschaftsbezogenen Beruf in Betracht ziehen. Diese Zunahme von 18 Pro-

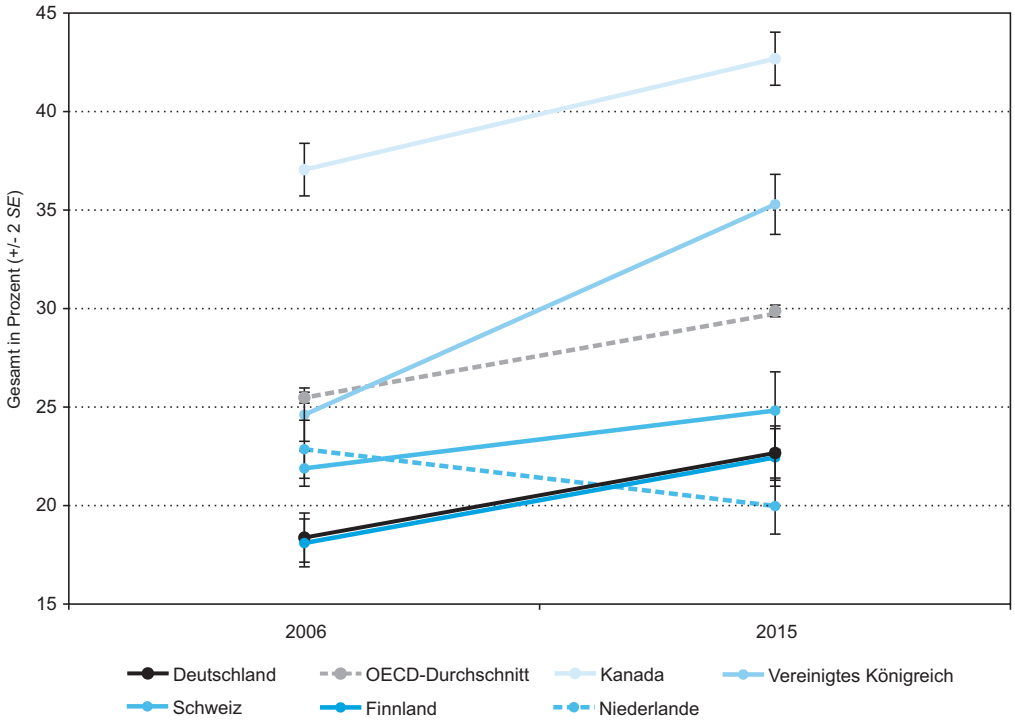


Abbildung 3.9: Die Veränderung der naturwissenschaftsbezogenen Berufserwartungen zwischen PISA 2006 und PISA 2015

© Waxmann Verlag GmbH

zent bei PISA 2006 auf 23 Prozent bei PISA 2015 ist statistisch signifikant. Ein solcher Anstieg zeigt sich auch im Durchschnitt aller OECD-Staaten. Die naturwissenschaftsbezogene Karriereerwartung ist seit PISA 2006 um 4 Prozent (gerundet) auf 30 Prozent angestiegen. Ein deutlich stärkerer Anstieg zeigt sich im Vereinigten Königreich. Hier ziehen 11 Prozent mehr Fünfzehnjährige einen naturwissenschaftsbezogenen Beruf in Betracht als bei PISA 2006. Auch in Kanada ziehen 6 Prozent mehr Jugendliche einen solchen Beruf in Erwägung – eine zusätzliche Verbesserung auf einem bereits hohen Niveau. Ein leichter, aber dennoch signifikanter Rückgang (von 3 Prozent) ist in den Niederlanden zu verzeichnen.

Betrachtet man in Deutschland die Veränderung für Mädchen und Jungen getrennt, zeigt sich die Zunahme naturwissenschaftsbezogener Karriereerwartungen nur für die Jungen (Differenz von +9 Prozent). Bei den Mädchen lassen sich keine signifikanten Unterschiede beobachten. In Bezug auf den prozentualen Anteil naturwissenschaftlicher Karriereerwartungen getrennt nach Schulart stieg der Anteil an Gymnasien um 6 und an anderen Schularten um 2 Prozent an.

3.4.7 Der Zusammenhang zwischen naturwissenschaftsbezogenen Berufserwartungen und Schülermerkmalen

Im Folgenden soll untersucht werden, wie die Neigung für oder gegen einen naturwissenschaftsbezogenen Beruf mit kognitiven sowie motivational-affektiven Schülermerkmalen zusammenhängt. Dazu wurden logistische Regressionen berechnet. Ein erstes Modell untersucht den Zusammenhang zwischen allgemeinen Schülervoraussetzungen wie Geschlecht (0 = Jungen, 1 = Mädchen), soziale Herkunft (HISEI – höchster sozioökonomischer Index der Eltern, vgl. Kapitel 8), Zuwanderungshintergrund (0 = kein Zuwanderungshintergrund, 1 = Zuwanderungshintergrund) sowie die besuchte Schulart (0 = nicht gymnasiale Schularten, 1 = Gymnasium) und der Neigung, sich für einen naturwissenschaftlichen Beruf zu entscheiden. In einem zweiten Modell wurde zusätzlich die naturwissenschaftliche Kompetenz berücksichtigt. Modell 3 zeigt schließlich den Zusammenhang zwischen motivational-affektiven Schülermerkmalen und der Berufserwartung.

Die Ergebnisse sind in Tabelle 3.1 dargestellt. Modell 1 zeigt, dass alle Schülervoraussetzungen mit der Berufserwartung zusammenhängen. Der Anteil der erklärten Varianz beträgt 6 Prozent. Berücksichtigt man zusätzlich die naturwissenschaftliche Kompetenz (Modell 2), hängt diese zwar positiv mit der Neigung, einen naturwissenschaftlichen Beruf in Erwägung zu ziehen, zusammen. Allerdings ist dieser Zusammenhang sehr klein. Die Varianzaufklärung steigt um zusätzliche 2 Prozent. Deutlich größere Effekte zeigen sich in Modell 3 für die motivational-affektiven Schülermerkmale. Sowohl Freude an den Naturwissenschaften als auch die instrumentelle Motivation hängen positiv mit der Berufserwartung zusammen. Nimmt die Freude an Naturwissenschaften zu, steigt auch die Wahrscheinlichkeit, mit 30 Jahren einen naturwissenschaftlichen Beruf auszu-

© Waxmann Verlag GmbH

üben. Noch deutlicher wird der Effekt für die instrumentelle Motivation. Interesse an naturwissenschaftlichen Themen und naturwissenschaftsbezogene Selbstwirksamkeitserwartung hingegen hängen nicht mit der Berufserwartung zusammen. Die Varianzaufklärung des Gesamtmodells beträgt 15 Prozent.

Betrachtet man diese Zusammenhänge getrennt für das Gymnasium und nicht gymnasiale Schularten, zeigen sich für die nicht gymnasialen Schularten ähnliche Ergebnisse wie für die Gesamtstichprobe. Die Wahrscheinlichkeit, einen naturwissenschaftlichen Beruf auszuüben (Cox-Snell R^2 = .09), steigt mit zunehmender Freude (b = 0.32, SE = .10, Eb = 1.37) und instrumenteller Motivation (b = 0.26, SE = .09, Eb = 1.30). Für das Gymnasium finden sich im Modell 3 zunächst keine Zusammenhänge mehr mit dem Geschlecht, dem Zuwanderungshintergrund, der sozialen Herkunft und der naturwissenschaftlichen Kompetenz. Auch Freude und Interesse an Naturwissenschaften hängen nicht mit der Erwartung, einen naturwissenschaftlichen Beruf auszuüben, zusammen. Einzig die instrumentelle Motivation – inwieweit Gymnasiastinnen und Gymnasiasten glauben, dass ihnen Naturwissenschaften für ihren späteren Beruf nützlich sein werden – hängt positiv mit der Berufserwartung zusammen (b = 0.9, SE = .11, Eb = 2.66). Am Gymnasium beträgt die Varianzaufklärung 21 Prozent. Insgesamt lässt sich daraus schlussfolgern, dass die naturwissenschaftliche Kompetenz weniger relevant für die Entscheidung, einen naturwissenschaftlichen Beruf auszuüben, ist. Die Förderung von Freude und Interesse sowie instrumenteller Motivation hingegen scheinen wichtige Ansatzpunkte zu sein, um Jugendliche für einen naturwissenschaftlichen Beruf zu begeistern.

Tabelle 3.2: Zusammenhang zwischen Schülermerkmalen und der naturwissenschaftsbezogenen Berufserwartung mit 30 Jahren

	Modell I			**Modell II**			**Modell III**		
	b	*(SE)*	*Eb*	*b*	*(SE)*	*Eb*	*b*	*(SE)*	*Eb*
Konstante	**-1.90**	0.14	0.15	**-4.21**	0.36	0.01	**-3.04**	0.56	0.05
Schulart	**0.76**	0.11	2.14	**0.31**	0.12	1.36	**0.53**	0.17	1.70
Geschlecht	**-0.44**	0.09	0.64	**-0.34**	0.09	0.71	-0.10	0.12	0.91
Zuwanderungshintergrund	**-0.27**	0.10	0.77	**-0.50**	0.10	0.61	**-0.55**	0.16	0.58
HISEI	**0.01**	0.00	1.01	**0.01**	0.00	1.01	**0.01**	0.00	1.01
Naturwissenschaftliche Kompetenz				**0.01**	0.00	1.01	**0.00**	0.00	1.00
Freude & Interesse							**0.21**	0.06	1.24
Interesse an naturwissenschaftlichen Themen							0.11	0.09	1.11
Instrumentelle Motivation							**0.58**	0.06	1.78
Selbstwirksamkeit							-0.10	0.06	0.91
Cox-Snell R^2	**0.06**			**0.08**			**0.15**		
Nagelkerke R^2	**0.09**			**0.12**			**0.22**		

fett: signifikante Regressionskoeffizienten (p < .05). b = unstandardisierte Regressionskoeffizienten, Eb = Odds Ratio.

© Waxmann Verlag GmbH

3.5 Zusammenfassung und Diskussion

Mehrdimensionale Bildungsziele, die neben einer fundierten Wissensbasis die Förderung naturwissenschaftsbezogener motivationaler Orientierungen und Selbstbilder berücksichtigen, sind wichtige Grundlagen für lebenslanges Lernen, prägen die Persönlichkeit von Jugendlichen und beeinflussen zukünftige Ausbildungs- und Berufswahlentscheidungen (vgl. Schiepe-Tiska et al., im Druck). Diese Orientierungen und Selbstbilder sind deshalb als Bildungsziele ebenso wichtig wie die in Leistungstests abgebildeten Kompetenzen. Die Ergebnisse für Deutschland bei PISA 2015 zeichnen in Bezug auf diese Merkmale ein Bild mit einigen wenigen Stärken und einer Reihe von Problemlagen, bei deren Interpretation allerdings Einschränkungen aufgrund des international gewählten Designs berücksichtigt werden müssen.

Für die motivationalen Orientierungen und die Selbstwirksamkeitserwartung geben die Befunde – im internationalen Vergleich – Anlass zur Sorge. Freude und Interesse an Naturwissenschaften sowie die instrumentelle Motivation sind unterdurchschnittlich ausgeprägt. Es bereitet 40 Prozent der Fünfzehnjährigen wenig oder keine Freude, sich mit naturwissenschaftlichen Themen zu befassen und fast der Hälfte ist nicht bewusst, dass Naturwissenschaften auch für das zukünftige Leben relevant sind. Was sich Jugendliche in Bezug auf ihre naturwissenschaftlichen Fähigkeiten zutrauen, liegt ebenfalls leicht unterhalb des OECD-Durchschnitts. Einzig das Interesse an naturwissenschaftlichen Themen ist leicht überdurchschnittlich ausgeprägt. In Deutschland interessieren sich Schülerinnen und Schüler dabei vor allem für Themen aus den Bereichen *lebende Systeme* sowie *Erd- und Weltraumsysteme*. Für Letzteres erzielten die Jugendlichen auch die vergleichsweise höchsten Kompetenzwerte (vgl. Kapitel 2). Der Unterschied zwischen Freude und Interesse an Naturwissenschaften allgemein und dem Interesse an naturwissenschaftlichen Themen könnte auf eine unterschiedliche Wahrnehmung von Naturwissenschaften in der Schule und naturwissenschaftlichen Themen allgemein zurückzuführen sein. Sobald es um konkrete Themen geht, scheint es für die Jugendlichen interessanter zu werden.

Betrachtet man kompetenzstarke Vergleichsstaaten wie Kanada oder Finnland, zeigen sich interessante Unterschiede in den motivational-affektiven Lernergebnissen: Während in Kanada Jugendliche Freude und Interesse an den Naturwissenschaften erleben, ihnen die Bedeutung für ihr späteres Leben bewusst ist und sie sich selbst auch einiges in Bezug auf die Naturwissenschaften zutrauen, ist in Finnland zwar die instrumentelle Motivation der Fünfzehnjährigen stark ausgeprägt, Freude und Interesse sowie die Selbstwirksamkeitserwartung sind jedoch leicht unterdurchschnittlich ausgeprägt. Hier scheinen Unterschiede in den Unterrichtstraditionen zum Tragen zu kommen. Während sich der Unterricht in beiden Staaten durch kognitiv aktivierende Lernaktivitäten wie Ideen erklären („*Minds-on*“) auszeichnet, ist der Unterricht in Kanada zusätzlich durch häufige Diskussionen sowie „*Hands-on*“-Aktivitäten wie im Labor strukturierte Experimente durchzuführen oder selbst Experimente zu entwickeln geprägt (vgl. Kapitel 4). Darüber hinaus werden häufig Anwendungsbezüge hergestellt und die Relevanz für das

© Waxmann Verlag GmbH

eigene Leben verdeutlicht. Dies scheint neben dem Wissenszuwachs auch die Entwicklung motivationaler Orientierungen und das Herausbilden positiver Selbstbilder zu fördern, was auch durch Metaanalysen gut belegt ist (z. B. Furtak, Seidel, Iverson & Briggs, 2012; Seidel & Shavelson, 2007). Der Unterricht in Deutschland, der größtenteils durchschnittlich kognitiv anregend ist, eher selten Möglichkeiten zum Experimentieren bietet und wenig Anwendungsbezüge herstellt (vgl. Kapitel 4), scheint demnach wenig dazu beizutragen, dass sich Jugendliche für Naturwissenschaften interessieren, deren Relevanz für ihr zukünftiges (Berufs-)Leben erfahren und sich als selbstwirksam erleben.

Dies wird noch einmal deutlicher, wenn man die Veränderung im Vergleich zu PISA 2006 betrachtet. Für Deutschland ist eine Abnahme an Freude, instrumenteller Motivation und Selbstwirksamkeitserwartung zu beobachten. Sowohl Schülerinnen und Schüler am Gymnasium als auch an den nicht gymnasialen Schularten erleben weniger Freude und Interesse und schätzen die Bedeutung von Naturwissenschaften für ihre Zukunft geringer ein als in PISA 2006. Zusätzlich hat bei den Gymnasiastinnen und Gymnasiasten die naturwissenschaftsbezogene Selbstwirksamkeitserwartung abgenommen. Sie sehen sich weniger in der Lage, naturwissenschaftliche Aufgabenstellungen zu lösen. Da sich im Vergleich zu PISA 2006 auch die naturwissenschaftliche Kompetenz am Gymnasium bei PISA 2015 verringert hat (vgl. Kapitel 2), mag dieser Abfall ein Ausdruck einer realistischen Selbsteinschätzung sein. Nichtsdestotrotz zeigt sich hier – ebenso wie für Freude und Interesse und instrumentelle Motivation – Handlungsbedarf, um gerade auch die kompetenzstarken Jugendlichen in Deutschland für Naturwissenschaften zu begeistern.

Besonders auffällig für Deutschland sind die ausgeprägten Geschlechterunterschiede. Die Jungen erleben bei PISA 2015 deutlich mehr Freude und Interesse an den Naturwissenschaften, sehen eher die Bedeutung für ihr zukünftiges Leben und trauen sich in Bezug auf ihre Fähigkeiten mehr zu als die Mädchen. Anlass zur Sorge gibt auch hier der Vergleich mit PISA 2006: Die Abnahme der motivationalen Orientierungen und Selbstwirksamkeitserwartung zeigt sich vor allem bei den Mädchen. Betrachtet man Unterschiede in den Interessen genauer, sind bei PISA 2015 die Jungen eher an physikalischen sowie Erd- und Weltraumsystemen interessiert, wohingegen die Mädchen eher die Tendenz haben, sich für lebende Systeme zu interessieren. Dieser Befund spiegelt auch den aktuellen Forschungsstand zu Interessensunterschieden bei Jungen und Mädchen wider (vgl. Krapp & Prenzel, 2011). Insgesamt zeichnet sich hier Handlungsbedarf ab, denn bei PISA 2015 sind zusätzlich zu den Differenzen in den motivational-affektiven Lernergebnissen erstmals sowohl im OECD-Durchschnitt als auch in Deutschland signifikante Geschlechterunterschiede in der naturwissenschaftlichen Kompetenz zugunsten der Jungen zu beobachten (vgl. Kapitel 2). Hier sind Unterrichtskonzepte gefragt, die gerade auch Mädchen für Naturwissenschaften begeistern und ihnen Erfolgserlebnisse ermöglichen. Ansätze dazu wie zum Beispiel das Anknüpfen an geschlechtsspezifische Interessen, das gezielte Fördern von Selbstbildern durch Zutrauen in die Fähigkeiten der Mädchen oder der Abbau einer eher männlich konnotierten Wahrnehmung von Physik und Technik sind in der Forschung bekannt, scheinen aber im Unterricht noch zu wenig

© Waxmann Verlag GmbH

umgesetzt zu werden (vgl. Labudde & Möller, 2012). Nicht zuletzt braucht es für diese Umsetzung gut ausgebildete, motivierte MINT-Lehrkräfte, die ihre Begeisterung für naturwissenschaftliche Themen an Mädchen *und* Jungen weitergeben. Hier sind Universitäten und Landesinstitute gefragt, Lehrkräfte in der Aus- und Weiterbildung stärker für den Umgang mit Heterogenität auch in Bezug auf Geschlechterdifferenzen in den Naturwissenschaften zu sensibilisieren.

Die Tendenz, später einen naturwissenschaftlichen Beruf ergreifen zu wollen, ist in Deutschland im Vergleich zum OECD-Durchschnitt ebenfalls geringer ausgeprägt. Nicht einmal ein Viertel der Jugendlichen gibt an, dass sie einen MINT-Beruf in Erwägung ziehen. Fast die Hälfte von diesen Fünfzehnjährigen zieht vor allem eine Tätigkeit als Naturwissenschaftler, Mathematiker und Ingenieure in Betracht. Diese Tendenz ist primär bei Gymnasiastinnen und Gymnasiasten zu beobachten. Im Vergleich zum OECD-Durchschnitt und den Vergleichsstaaten ist auffällig, dass in anderen Staaten vor allem Berufe aus dem Gesundheitsbereich in Erwägung gezogen werden. Ein möglicher Grund kann sein, dass solche Berufe in anderen Staaten im Vergleich zu Deutschland häufiger akademische Berufe sind. In Deutschland bieten viele dieser Berufe im Vergleich zu beispielsweise Ingenieursberufen oder auch Technikerberufen geringere Verdienst- und weniger soziale Aufstiegsmöglichkeiten.

Auch für die Berufserwartungen zeigen sich – ähnlich zu den motivationalen Orientierungen – Geschlechterdifferenzen: Insgesamt können sich die Jungen eher die Ausübung eines solchen Berufs vorstellen als die Mädchen. Dies deckt sich mit den aktuellen Zahlen zu den MINT-Studienfächern: Hier betrug der Frauenanteil im Jahr 2015 lediglich 28 Prozent; insgesamt lag der Frauenanteil bei den MINT-Berufen bei 15 Prozent (Bundesagentur für Arbeit, 2016). Andererseits sank die Anzahl arbeitsloser Frauen im MINT-Bereich überproportional und bietet ihnen daher sehr gute Zukunftschancen. Betrachtet man die Berufserwartungen mit Blick auf die unterschiedlichen Bereiche, zeigt sich, dass 80 Prozent der Jungen in Betracht ziehen, Naturwissenschaftler, Mathematiker, Ingenieure zu werden oder im zukünftig weiter wachsenden Berufsfeld der Informations- und Kommunikationstechnologie tätig zu werden. Etwas mehr als ein Drittel der Mädchen findet die erstgenannten Berufe auch spannend, die Hälfte erwägt eher einen MINT-Beruf im Gesundheitsbereich, welche jedoch bereits in der Ausbildung mit vergleichsweise geringen Gehältern einhergehen (acatech, 2015). Von den Mädchen können sich gerade einmal 3 Prozent einen Beruf im Bereich der Informations- und Kommunikationstechnologie vorstellen. Dabei kommen – zum Teil ungerechtfertigte – Sichtweisen auf naturwissenschaftlich-technische Berufe zum Tragen. Mädchen nehmen Naturwissenschaften eher als schwierig, männlich und wenig sozial wahr und machen sie deshalb kaum zum Bestandteil ihrer eigenen Geschlechtsidentität (Kessels & Hannover, 2004, 2007). Darüber hinaus fehlen ihnen in der Familie häufig weibliche Vorbilder in diesen Berufen, mit denen sie sich identifizieren können. Insgesamt wenig Beachtung finden die Technikberufe, für die sich jedoch auch zukünftig ein Nachwuchskräftemangel abzeichnet und die ebenfalls gute Verdienstmöglichkeiten und Karriereperspektiven

© Waxmann Verlag GmbH

bieten (Bundesagentur für Arbeit, 2016). Ein Grund dafür kann die geringe Präsenz von Technik als Inhalt in den naturwissenschaftlichen Fächern sein.

Überraschend scheint – gerade im Vergleich zu den ernüchternden Befunden der motivationalen Orientierungen –, dass in Deutschland im Vergleich zu PISA 2006 die naturwissenschaftsbezogenen Berufserwartungen angestiegen sind. Dies lässt sich vor allem auf eine Zunahme der Berufswahltendenz bei den Jungen zurückführen. Dabei scheinen jedoch weniger die Schule und der Unterricht eine Rolle zu spielen, sondern eher die gestiegene Medienaufmerksamkeit sowie die zahlreichen Initiativen und Netzwerke zur Förderung von MINT, welche Jugendliche mit zunehmendem Erfolg für Ausbildungs- und Berufsentscheidungen im MINT-Bereich begeistern. Taskinen et al. (2013) fanden heraus, dass das Aufzeigen von Anwendungsbezügen im naturwissenschaftlichen Unterricht sowie extracurriculare naturwissenschaftliche Angebote der Schule mit mehr Freude und Interesse sowie einem höheren Selbstkonzept zusammenhängen, die wiederum ihrerseits mit einer höheren zukunftsorientierten Motivation einhergingen. Auch in Deutschland hing bei PISA 2015 die Bereitschaft, einen MINT-Beruf zu ergreifen, mit Freude und Interesse an Naturwissenschaften sowie der instrumentellen Motivation zusammen. Dies scheinen wichtige Ansatzpunkte zu sein, um Jugendliche für einen naturwissenschaftlichen Beruf zu begeistern. Darüber hinaus sollten Schulen ihr bestehendes Angebot der Berufsinformation und -orientierung ausbauen, um Jugendliche über MINT-Berufe zu informieren und ihnen realistische Vorstellungen über deren Tätigkeitsprofile zu vermitteln (vgl. acatech, 2015). Hier können Schulen als Vorbild dienen, die sich durch einen ausgeprägten MINT-Schwerpunkt auszeichnen und denen es gelingt, gleichzeitig kognitive und motivational-affektive Bildungsziele zu fördern sowie ihre Schülerinnen und Schüler über naturwissenschaftliche Ausbildungs- und Berufsmöglichkeiten zu informieren und für diese zu motivieren (vgl. Seidel et al., 2016).

Literatur

acatech – Deutsche Akademie der Technikwissenschaften & Körber-Stiftung. (2015). *MINT Nachwuchsbarometer 2015: Fokusthema: Berufliche Ausbildung.* Zugriff am 03.11.2016. Verfügbar unter http://www.koerber-stiftung.de/fileadmin/user_upload/allgemein/schwerpunkte/2015/wissenschaft/mint-nachwuchsbarometer/MINT-Nachwuchsbarometer-2015-Broschuere.pdf

Aktionsrat Bildung. (2015). *Bildung. Mehr als Fachlichkeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215. http://dx.doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*, 117–148. http://dx.doi.org/10.1207/s15326985ep2802_3

Bundesagentur für Arbeit, Statistik/Arbeitsmarktberichterstattung. (2016). *Der Arbeitsmarkt in Deutschland – MINT-Berufe*. Abgerufen am 03.11.2016 von http://statistik.arbeitsagentur.de/Statischer-Content/Arbeitsmarktberichte/Branchen-Berufe/generische-Publikationen/Broschuere-MINT-2016.pdf

© Waxmann Verlag GmbH

Croll, P. (2008). Occupational choice, socio-economic status and educational attainment: A study of the occupational choices and destinations of young people in the british household panel survey. *Research Papers in Education*, *23*, 243–268. http://dx.doi.org/10.1080/02671520701755424

Daniels, Z. (2008). *Entwicklung schulischer Interessen im Jugendalter*. Münster: Waxmann.

Deci, E. L. & Ryan, R. M. (Hrsg.) (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.

DeWitt, J. & Archer, L. (2015). Who aspires to a science career? A comparison of survey responses from primary and secondary school students. *International Journal of Science Education*, *37*, 2170–2192. http://dx.doi.org/10.1080/09500693.2015.1071899

Eccles, J. S. (2011). Understanding educational and occupational choices. *Journal of Social Issues, 67*, 644–648. http://dx.doi.org/10.1111/j.1540-4560.2011.01718.x

Frey, A., Taskinen, P., Schütte, K., Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. & Pekrun, R. (2009). *Pisa 2006 Skalenhandbuch: Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.

Furtak, E. M., Seidel, T., Iverson, H. & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research*, *82*, 300–329. http://doi.org/10.3102/0034654312457206

Häußler, P. & Hoffmann, L. (2000). A curricular frame for physics education: Development, comparison with students' interests, and impact on students' achievement and self-concept. *Science Education*, *84*, 689–705. http://doi.org/10.1002/1098-237X(200011)84:6<689::AID-SCE1>3.0.CO;2-L

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*, 903–918. http://doi.org/10.1037/0022-3514.82.6.903

International Labour Office. (2012). *International standard classification of occupations ISCO-08*. Genf: ILO.

Jansen, M., Scherer, R. & Schroeders, U. (2015). Students' self-concept and self-efficacy in the sciences: Differential relations to antecedents and educational outcomes. *Contemporary Educational Psychology*, *41*, 13–24. http://doi.org/10.1037/0022-3514.82.6.90310.1016/j.cedpsych.2014.11.002

Johnson, T. P., Shavitt, S. & Holbrook, A. L. (2011). Survey response styles across cultures. In D. R. Matsumoto & F. J. R. d. van Vijver (Hrsg.), *Culture and psychology. Cross-cultural research methods in psychology* (S. 130–178). New York: Cambridge University Press.

Kessels, U. & Hannover, B. (2004). Empfundene „Selbstnähe" als Mediator zwischen Fähigkeitsselbstkonzept und Leistungskurswahlintentionen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *36*, 130–138. http://doi.org/10.1026/0049-8637.36.3.130

Kessels, U. & Hannover, B. (2007). How the image of math and science affects the development of academic interests. In M. Prenzel (Hrsg.), *Studies on the educational quality of schools. The final report on the DFG priority programme* (S. 283–297). Münster: Waxmann.

Kjærnsli, M. & Lie, S. (2011). Students' preference for science careers: International comparisons based on PISA 2006. *International Journal of Science Education, 33*, 121–144. http://dx.doi.org/10.1080/09500693.2010.518642

Klassen, R. M. & Usher, E. L. (2010). Self-efficacy in educational settings: Recent research and emerging direction. In T. C. Urdan & S. A. Karabenick (Hrsg.), *Advances in moti-*

© Waxmann Verlag GmbH

vation and achievement: Vol. 16. The decade ahead: Theoretical perspectives on motivation and achievement (S. 1–33). Bingley, U.K.: Emerald.

Köller, O., Trautwein, U., Lüdtke, O. & Baumert, J. (2006). Zum Zusammenspiel von schulischer Leistung, Selbstkonzept und Interesse in der gymnasialen Oberstufe. *Zeitschrift für Pädagogische Psychologie*, *20*, 27–39. http://doi.org/10.1024/1010-0652.20.12.27

Krapp, A. (2002). An educational-psychological theory of interest and its relation to self-determination theory. In E. L. Deci & R. M. Ryan (Hrsg.), *Handbook of self-determination research* (S. 405–427). Rochester, NY: University of Rochester Press.

Krapp, A. & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education*, *33*, 27–50. http://doi.org/10.1080/09500693.2010.518645

Labudde, P. & Möller, K. (2012). Stichwort: Naturwissenschaftlicher Unterricht. *Zeitschrift für Erziehungswissenschaft*, *15*, 11–36. http://doi.org/10.1007/s11618-012-0257-0

Nagengast, B. & Marsh, H. W. (2012). Big fish in little ponds aspire more: Mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *Journal of Educational Psychology*, *104*, 1033–1053. http://doi.org/10.1037/a0027697

OECD. (2016a). *PISA 2015 Assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. Paris: OECD.

OECD. (2016b). *PISA 2015 Results (Volume I): Excellence and equity in education*. Paris: OECD.

Osborne, J., Simon, S. & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*, 1049–1079. http://doi.org/10.1080/0950069032000032199

Pajares, F., Britner S. L. & Valiante G. (2000). Relation between achievement goals and self-beliefs of middle school students in writing and science. *Contemporary Educational Psychology*, *25*, 406–422. http://doi.org/10.1006/ceps.1999.1027

Pant, H. A., Stanat, P., Pöhlmann, C., Hecht, M., Jansen, M., Kampa, N., Lenski, A., Penk, C., Radmann, S., Roppelt, A., Schroeders, U., Siegle, T. & Ziemke, A. (2013). Der Blick in die Länder. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 159–248). Münster: Waxmann.

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S. & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology*, *34*, 29–48. http://doi.org/10.1080/01443410.2013.797339

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*, 315–341. http://doi.org/10.1007/s10648-006-9029-9

Pekrun, R. & Linnenbrink-Garcia, L. (2014). *Handbook of emotions in education*. New York: Taylor & Francis / Routledge.

Prenzel, M., Geiser, H., Langeheine, R. & Lobemeier, K. (2003). Das naturwissenschaftliche Verständnis am Ende der Grundschule. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 143–188). Münster: Waxmann.

© Waxmann Verlag GmbH

Prenzel, M., Lankes, E.-M. & Minsel, B. (2000). Interessenentwicklung in Kindergarten und Grundschule: Die ersten Jahre. In U. Schiefele & K. P. Wild (Hrsg.), *Interesse und Lernmotivation. Neue Studien zu Entwicklung und Wirkungen* (S. 11–30). Münster: Waxmann.

Schiepe-Tiska, A., Roczen, N., Müller, K., Prenzel, M. & Osborne, J. (im Druck). Science-related outcomes: Attitudes, motivation, value beliefs, strategies. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing contexts of learning world-wide.* Berlin: Springer.

Schneider, B., Krajcik, J., Lavonen, J., Salmela-Aro, K., Broda, M., Spicer, J., Bruner, J., Moeller, J., Linnansaari, J., Juuti, K. & Viljaranta, J. (2015). Investigating optimal learning moments in U. S. and Finnish science classes. *Journal of Research in Science Teaching, 53*, 400–421. http://doi.org/10.1002/tea.21306

Schütte, K., Frenzel, A. C., Asseburg, R. & Pekrun, R. (2007). Schülermerkmale, naturwissenschaftliche Kompetenz und Berufserwartung. 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. 125–146). Münster: Waxmann.

Seidel, T., Reinhold, S., Holzberger, D., Mok, S.-Y., Schiepe-Tiska, A. & Reiss, K. (2016). *Wie gelingen MINT-Schulen? Anregungen aus Forschung und Praxis*. Münster: Waxmann.

Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, *77*, 454–499. http://doi.org/10.3102/0034654307310317

Tai, R. H., Qi Liu, C., Maltese, A. V. & Fan, X. (2006). Career choice. Planning early for careers in science. *Science*, *312*, 1143–1144. http://doi.org/10.1126/science.1128690

Taskinen, P., Asseburg, R. & Walter, O. (2008). Wer möchte später einen naturwissenschaftsbezogenen oder technischen Beruf ergreifen? Kompetenzen, Selbstkonzept und Motivationen als Prädiktoren der Berufserwartungen in PISA 2006. In M. Prenzel & J. Baumert (Hrsg.), *Vertiefende Analysen zu PISA 2006* (S. 79–105). Wiesbaden: VS Verlag für Sozialwissenschaften. http://doi.org/10.1126/science.112869010.1007/978-3-531-91815-0_5

Taskinen, P. H., Schütte, K. & Prenzel, M. (2013). Adolescents' motivation to select an academic science-related career: The role of school factors, individual interest, and science self-concept. *Educational Research and Evaluation*, *19*, 717–733. http://doi.org/10.1080/13803611.2013.853620

van de Gaer, E., Grisay, A., Schulz, W. & Gebhardt, E. (2012). The reference group effect: An explanation of the paradoxical relationship between academic achievement and self-confidence across countries. *Journal of Cross-Cultural Psychology*, *43*, 1205–1228. http://doi.org/10.1177/0022022111428083

© Waxmann Verlag GmbH

4 Naturwissenschaftlicher Unterricht in Deutschland in PISA 2015 im internationalen Vergleich

Anja Schiepe-Tiska, Stefanie Schmidtner, Katharina Müller, Jörg-Henrik Heine, Knut Neumann & Oliver Lüdtke

Dieser Beitrag geht der Frage nach, wie sich der Unterricht aus der Sicht der Fünfzehnjährigen in Deutschland im Vergleich zu ausgewählten OECD-Staaten charakterisieren lässt und mit mehrdimensionalen Bildungszielen wie naturwissenschaftlicher Kompetenz sowie Freude und Interesse an den Naturwissenschaften zusammenhängt. Die Ergebnisse zeigen für Deutschland, dass der Unterricht wenige Störungen aufweist – jedoch berichten die Fünfzehnjährigen auch eher wenig wahrgenommene Unterstützung, individuelle Rückmeldungen und Differenzierung. In Bezug auf die fachspezifischen Lernaktivitäten zeichnet sich der Naturwissenschaftsunterricht in Deutschland vor allem durch „*Minds-on*"-Aktivitäten aus, die auf eine kognitive Anregung der Schülerinnen und Schüler abzielen. „*Hands-on*"-Aktivitäten kommen hingegen seltener vor. Das Anwenden von Prinzipien auf naturwissenschaftliche Phänomene sowie das Herstellen eines Bezugs zur alltäglichen Lebenswelt der Schülerinnen und Schüler scheint im naturwissenschaftlichen Unterricht in Deutschland ebenfalls eine eher geringe Rolle zu spielen. Im Vergleich zu PISA 2006 hat sich der naturwissenschaftliche Unterricht in Deutschland in seiner Grundstruktur kaum verändert. Eine Betrachtung von Unterrichtsmustern im Zusammenhang mit den domänenübergreifenden Qualitätsmerkmalen verdeutlicht allerdings, dass sich ein guter Naturwissenschaftsunterricht durch eine Kombination hoher kognitiver Anregung („*Minds-on*"), regelmäßiger „*Hands-on*"-Aktivitäten, wie die Entwicklung eigener Experimente oder die Durchführung strukturierter Laborexperimente, sowie dem Aufzeigen von Anwendungsbezügen auszeichnet. Am Gymnasium ist dieser Unterrichtstyp zudem mit einer hohen naturwissenschaftlichen Kompetenz und einer ausgeprägten Freude und Interesse an den Naturwissenschaften verbunden.

Bereits im Vorschulalter sollen Kinder ein erstes Verständnis naturwissenschaftlicher Konzepte, grundlegende naturwissenschaftliche Fähigkeiten und insbesondere ein Interesse an den Naturwissenschaften entwickeln (z. B. im *Haus der kleinen Forscher*[1]). Jedoch ist erst die Schule der Ort, an dem die Entwicklung einer naturwissenschaftlichen

1 Weiterführende Informationen zum Haus der kleinen Forscher können unter http://www.haus-der-kleinen-forscher.de/ aufgerufen werden.

© Waxmann Verlag GmbH

Grundbildung systematisch betrieben wird. Der naturwissenschaftliche Unterricht soll dabei nicht nur ein solides naturwissenschaftliches Fachwissen und die Beherrschung gängiger naturwissenschaftlicher Arbeitsweisen vermitteln. Im Sinne einer umfassenden naturwissenschaftlichen Grundbildung soll er darüber hinaus Freude und Interesse an den Naturwissenschaften fördern sowie die Entwicklung positiver Selbstbilder unterstützen (Aktionsrat Bildung, 2015; Schiepe-Tiska, Rozcen, Müller, Prenzel & Osborne, im Druck; siehe auch Kapitel 3).

Ein solcher Unterricht folgt nicht der Annahme, dass Lehren automatisch zu einem Kompetenzzuwachs bei Schülerinnen und Schülern führt. Vielmehr stellt die Lehrkraft ein Angebot bereit, welches von den Lernenden auf der Basis ihrer individuellen Voraussetzungen und Erfahrungen aktiv wahrgenommen und genutzt werden muss (Seidel & Reiss, 2014). Dabei sind letztlich auch nicht einzelne Merkmale (z. B. bestimmte Organisationsformen oder Methoden) entscheidend, sondern deren Arrangement und Zusammenspiel. Vielfältige Lernaktivitäten mit variierenden methodischen Zugängen, die situationsangemessen eingesetzt werden, führen zu substanziellen Lernerfolgen (Cohen, Raudenbush & Ball, 2003).

Ein entsprechendes Verständnis von gutem Unterricht liegt auch der Erfassung der Unterrichtsqualität in PISA 2015 zugrunde. Die Unterrichtsmerkmale werden dabei aus der Sicht der Schülerinnen und Schüler erhoben. Im Folgenden wird zunächst der aktuelle Forschungsstand zur Qualität des Unterrichts in den Naturwissenschaften zusammengefasst. Dabei wird auf die Bedeutsamkeit domänenübergreifender Qualitätsmerkmale und naturwissenschaftsbezogener Lernaktivitäten eingegangen. Anschließend wird dargestellt, wie die Unterrichtsqualität im Rahmen von PISA 2015 erfasst wurde. Darauf aufbauend werden die Ergebnisse für Deutschland berichtet und zu denen ausgewählter OECD-Staaten in Beziehung gesetzt. Es folgt eine differenzierte Betrachtung der Schularten. Da bei PISA 2006, als die Naturwissenschaften das letzte Mal Hauptdomäne waren, ein Teil der Unterrichtsmerkmale bereits erhoben wurde, wird zusätzlich die Veränderung dieser Merkmale im Vergleich zu 2006 untersucht. Abschließend werden Muster des forschend-entdeckenden Unterrichts in Deutschland präsentiert und Zusammenhänge mit domänenübergreifenden Qualitätsmerkmalen betrachtet. Da die Schülerschaft am Gymnasium in Bezug auf die naturwissenschaftliche Kompetenz relativ homogen ist, werden für diese darüber hinaus Zusammenhänge zwischen den Unterrichtsmustern und naturwissenschaftlicher Kompetenz sowie Freude und Interesse an den Naturwissenschaften berichtet. Für die nicht gymnasialen Schularten wird auf solche Analysen verzichtet, da die Ergebnisse aufgrund der Überführung in neue oder neu bezeichnete Schularten in verschiedenen Bundesländern und der weniger homogen zusammengesetzten Schülerschaft schwieriger zu interpretieren wären.

© Waxmann Verlag GmbH

4.1 Empirische Befunde zu lernwirksamen Merkmalen des naturwissenschaftlichen Unterrichts

Die empirische Unterrichtsforschung belegt die Wirksamkeit einer ganzen Reihe von Unterrichtsmerkmalen im Hinblick auf das Erreichen kognitiver und motivational-affektiver Lernergebnisse. Als Voraussetzungen für einen guten Unterricht wurden auf der Grundlage von Analysen von Unterrichtsvideos die sogenannten Basisdimensionen der Unterrichtsqualität identifiziert (Klieme, Schümer & Knoll, 2001). Zu diesen domänenübergreifenden Merkmalen der Instruktionsqualität gehören die Klarheit und Strukturiertheit des Unterrichts, die Klassenführung, die konstruktive Unterstützung durch die Lehrkraft sowie das kognitive Aktivierungspotenzial des Unterrichts. Im Zusammenhang mit der Unterstützung stehen auch intentionale Unterrichtshandlungen, wie Schülerinnen und Schülern individuelle Leistungsrückmeldungen zu geben oder den Unterricht differenziert an die Bedürfnisse und den Wissensstand der Klasse anzupassen.

Obwohl sich die konkrete Operationalisierung dieser Dimensionen in Abhängigkeit vom jeweiligen Unterrichtsfach unterscheiden kann, sind sie dennoch auf den Unterricht in verschiedenen Fächern anwendbar. So ist zum Beispiel gut belegt, dass ein strukturierter, störungsarmer und disziplinierter Mathematikunterricht, in dem die Lehrkraft die Schülerinnen und Schüler unterstützt und positives und konstruktives Feedback gibt, zu besseren motivationalen und kognitiven Lernergebnissen führt (etwa Baumert et al., 2010; Decristan et al., 2015; Klieme et al., 2001; Klieme, Pauli & Reusser, 2009; Kunter & Voss, 2011). Im Kontext von PISA zeigte sich für diese Basisdimensionen der Unterrichtsqualität im Bereich Mathematik ein robuster Zusammenhang mit unterschiedlichen Lernergebnissen (z. B. Klieme & Rakoczy, 2003; Schiepe-Tiska, Heine, Lüdtke, Seidel & Prenzel, 2016). Auch im Bereich des naturwissenschaftlichen Unterrichts haben sich diese Merkmale als relevant für das Lernen erwiesen (Neumann, Kauertz & Fischer, 2012). So fand sich in einer Videostudie zum Physikunterricht unter anderem ein positiver Zusammenhang zwischen der Klassenführung und dem Fachwissen (Junge, von Arx & Labudde, 2014).

Videostudien zum naturwissenschaftlichen Unterricht zeigen allerdings auch, dass neben diesen domänenübergreifenden, eher distalen Merkmalen der Unterrichtsqualität, die sich auf den Unterricht als Ganzes beziehen, auch Merkmale fachspezifischer Lehr-Lernaktivitäten, z. B. im Bereich des naturwissenschaftlichen Arbeitens, berücksichtigt werden müssen (Seidel & Prenzel, 2006b). Diese Merkmale sind näher am konkreten Lernprozess und werden deshalb auch als proximale Unterrichtsmerkmale bezeichnet. In einer Metaanalyse zur Effektivität verschiedener distaler und proximaler Unterrichtsmerkmale zeigten diese fachspezifischen Lernaktivitäten die höchsten Effekte sowohl auf kognitive als auch auf motivational-affektive Lernergebnisse (Seidel & Shavelson, 2007).

Als eine Form des Unterrichts in den Naturwissenschaften, der fachspezifisch aktivierende Tätigkeiten in besonderem Maße unterstützt, gilt der forschend-entdeckende Unterricht (*Scientific Inquiry*; NRC, 2012). Dieses Unterrichtskonzept berücksich-

© Waxmann Verlag GmbH

tigt die zentrale Rolle des Experiments für die Naturwissenschaften und stellt die Entwicklung eines Verständnisses des Gesamtprozesses naturwissenschaftlicher Erkenntnisgewinnung in den Mittelpunkt des Lernens. Dabei sollen Schülerinnen und Schüler nicht nur lernen, Experimente durchzuführen und deren Ergebnisse zu interpretieren, sondern auch z. B. naturwissenschaftliche Problem- oder Fragestellungen zu formulieren, Vermutungen anzustellen, passende Experimente selbst zu planen und zu entwickeln oder Beobachtungen durchzuführen. Darüber hinaus lernen sie, die gefundenen Ergebnisse sachgerecht darzustellen, zu interpretieren und sie im Kontext ihres bereits bekannten Wissens zu diskutieren. Darauf aufbauend entwickeln sie neue Frage- und Problemstellungen, die sie in weiterführenden Untersuchungen bearbeiten können (Labudde & Börlin, 2013; NRC, 2012). Ein solcher Unterricht soll Schülerinnen und Schülern zum einen ein vertieftes Verständnis alltäglicher und nicht alltäglicher naturwissenschaftlicher Phänomene vermitteln. Zum anderen soll den Lernenden in anwendungs- und lebensweltbezogenen Kontexten die Möglichkeit gegeben werden, Fähigkeiten im Bereich der Methoden naturwissenschaftlicher Erkenntnisgewinnung zu erlangen (Abd-el-Khalik et al., 2004).

Obwohl dem forschend-entdeckenden Lernen ein positiver Einfluss auf die Leistung und motivational-affektiven Lernergebnisse der Schülerinnen und Schüler bescheinigt wurde (Schroeder, Scott, Tolson, Huang & Lee, 2007; Seidel & Shavelson, 2007), ist dieses Unterrichtskonzept nicht unumstritten. Kritiker führen an, Schülerinnen und Schüler würden im forschend-entdeckenden Unterricht oftmals zu wenig durch die Lehrkraft angeleitet und strukturiert (Kirschner, Sweller & Clark, 2006; Klahr & Nigam, 2004). In einer Metaanalyse zur Wirksamkeit verschiedener Dimensionen des forschend-entdeckenden Unterrichts wurde daher zum einen zwischen Lernaktivitäten seitens der Lernenden und zum anderen zwischen dem Ausmaß, mit dem diese Lernaktivitäten durch die Lehrkraft angeleitet wurden, unterschieden (Furtak, Seidel, Iverson & Briggs, 2012). Dabei wurden vier grundlegend verschiedene Lernaktivitäten beschrieben: (1) *Konzeptuelle Aktivitäten* umfassen kognitive Prozesse wie das wissenschaftliche Begründen oder die Verwendung theoretischen Wissens. (2) *Prozedurale Aktivitäten* beziehen sich vor allem auf das Experimentieren als zentrale Aktivität in den Naturwissenschaften (NRC, 2012). Dazu zählen neben dem Durchführen von Experimenten und dem Austesten von Ideen auch das selbstständige Planen und Entwickeln eigener Untersuchungen zur Überprüfung einer Fragestellung. (3) *Epistemische Aktivitäten* umfassen das Sammeln, Evaluieren und Interpretieren von Befunden aus eigenen wissenschaftlichen Untersuchungen. (4) *Soziale Aktivitäten* beziehen sich auf das Kommunizieren als bedeutsamen Aspekt der Naturwissenschaften, um Informationen zu vermitteln. Dazu gehören Aktivitäten wie Ideen Erklären, Argumentieren, Diskutieren oder das gemeinsame Begründen. Die Ergebnisse von Furtak et al. (2012) zeigen, dass sich vor allem Lernaktivitäten im epistemischen Bereich sowie eine Kombination von Aktivitäten aus dem prozeduralen, epistemischen und sozialen Bereich positiv auf das Lernen auswirken. Gestützt werden diese Befunde durch Arbeiten, die vertieft einzelne Lernaktivitäten in den Blick nehmen und etwa die Rolle des Argumentierens (z. B. Berland & McNeill, 2010; Jiménez-Aleixandre

© Waxmann Verlag GmbH

& Puig, 2012), des kritischen Diskurses (z. B. Osborne, 2010; Henderson, MacPherson, Osborne & Wild, 2015) oder die Bedeutung der von Schülerinnen und Schülern gestellten Fragen im Unterricht (z. B. Andersson-Bakken & Klette, 2016; Ødegaard & Klette, 2012) als wichtige Ressource betonen. In Bezug auf das Ausmaß, mit dem diese Aktivitäten durch die Lehrperson angeleitet werden, zeigten sich positive Effekte bei einem stärker von der Lehrkraft angeleiteten forschend-entdeckenden Unterricht (Furtak et al., 2012). Als wesentliches Kriterium für die Wirksamkeit dieser Aktivitäten im Hinblick auf den Aufbau von Wissen und vor allem die Entwicklung motivational-affektiver Lernergebnisse hat sich darüber hinaus die Einbettung in authentische und für die Lebenswelt der Schülerinnen und Schüler relevante Kontexte erwiesen, in denen ein *Anwendungsbezug* hergestellt wird (z. B. Parchmann et al., 2007; Smith & Matthews, 2000; Tsai, 2000).

Merkmale des forschend-entdeckenden Unterrichts wurden bereits im Rahmen von PISA 2006 in den Blick genommen (Seidel, Prenzel, Wittwer & Schwindt, 2007). Dabei zeigte sich, dass in den meisten Staaten interaktive und kooperative Aktivitäten sowie der Transfer naturwissenschaftlicher Konzepte auf die Lebenswelt der Schülerinnen und Schüler im Vordergrund standen. Von praktischen Tätigkeiten wie etwa der Durchführung eigener Experimente wurde seltener berichtet (Seidel et al., 2007). In Staaten, in denen Schülerinnen und Schüler hohe Kompetenzwerte in den Naturwissenschaften erzielten, fanden sich häufiger Aktivitäten, in denen sie angehalten wurden, ihre Ideen zu erklären oder eigene Schlussfolgerungen aus Experimenten zu ziehen (Taylor, Stuhlsatz & Bybee, 2009). Das Schlussfolgern scheint jedoch nicht nur mit einer hohen naturwissenschaftlichen Kompetenz (Lavonen & Laaksonen, 2009), sondern auch mit günstigen motivational-affektiven Lernergebnissen (Kjærnsli & Lie, 2011) im Zusammenhang zu stehen.

Dass aber nicht nur einzelne Aspekte des forschend-entdeckenden Lernens, sondern insbesondere deren Zusammenspiel und Orchestrierung bedeutsam sind, konnten Kobarg et al. (2011) mithilfe einer Analyse von Unterrichtsmustern zeigen. In einer vertieften Analyse der in PISA 2006 verwendeten Fragen zum naturwissenschaftlichen Unterricht wurden anhand ausgewählter Items drei Muster identifiziert: (1) ein kognitiv aktivierender, forschungsbezogener „*Minds-on*"-Unterricht, der positiv mit den kognitiven Lernergebnissen und dem Interesse zusammenhing, (2) ein stärker auf die Vermittlung von Wissen ausgerichteter Unterricht, der positiv mit der naturwissenschaftlichen Kompetenz zusammenhing, aber mit weniger Interesse verbunden war, und (3) ein motivierender, aber mit niedrigeren Lernergebnissen verbundener, von „*Hands on*"-Aktivitäten geprägter Unterricht. Diese Muster forschend-entdeckenden Lernens, die sich in allen OECD-Staaten fanden, zeigen, dass ein gezielt auf vertiefte kognitive Aktivitäten ausgerichteter Unterricht nicht nur mit ausgeprägten kognitiven Lernergebnissen, sondern auch mit einem entsprechenden Interesse an naturwissenschaftlichen Themen einhergeht (Kobarg et al., 2011).

Der Naturwissenschaftsunterricht in Deutschland zeichnete sich bei PISA 2006 vor allem dadurch aus, dass Schülerinnen und Schüler zwar ab und zu Laborexperimente

© Waxmann Verlag GmbH

durchführen durften, aber selten die Gelegenheit erhielten, eigene Experimente zu entwickeln. Sie waren jedoch regelmäßig aufgefordert, Schlüsse aus Experimenten zu ziehen, eigene Ideen zu erklären und Bezüge zur alltäglichen Lebenswelt herzustellen (Seidel et al., 2007).

Fasst man domänenübergreifende Qualitätsmerkmale des Unterrichts und fachspezifische naturwissenschaftsbezogene Lernaktivitäten zusammen, so zeichnet sich ein naturwissenschaftlicher Unterricht, der den Aufbau einer naturwissenschaftlichen Grundbildung wirkungsvoll fördert sowie Freude und Interesse an den Naturwissenschaften unterstützt, dadurch aus, dass er im Hinblick auf die Organisation des Unterrichts einen strukturierten, störungsarmen und disziplinierten Kontext zur Verfügung stellt. Dieser ist einerseits kognitiv anregend gestaltet und bietet Gelegenheit zu forschend-entdeckendem Lernen und andererseits unterstützt die Lehrkraft gezielt die fachlichen Lernaktivitäten der Schülerinnen und Schüler. Die Forschung der letzten Jahrzehnte hat gezeigt, dass effektives forschend-entdeckendes Lernen ein Gleichgewicht zwischen inhaltlicher Offenheit einerseits sowie passender Lernbegleitung, Strukturierung und Unterstützung durch die Lehrkraft andererseits aufrechterhalten muss (Alfieri, Brooks, Aldrich & Tenenbaum, 2011; Blanchard et al., 2010; Lazonder & Harmsen, 2016; Seidel et al., 2007). Die Untersuchung der Unterrichtsqualität mit Blick sowohl auf kognitive als auch auf motivational-affektive Lernergebnisse erfordert demnach die Berücksichtigung domänenübergreifender und naturwissenschaftsbezogener Unterrichtsmerkmale.

4.2 Die Erfassung der Unterrichtsqualität in PISA 2015

4.2.1 Methodische Herausforderungen

In PISA werden die Merkmale der Unterrichtsqualität aus einer international vergleichenden Perspektive in den Blick genommen. Dabei werden die fünfzehnjährigen Schülerinnen und Schüler zu Merkmalen ihres naturwissenschaftlichen Unterrichts befragt, die sich in der empirischen Schul- und Unterrichtsforschung sowohl hinsichtlich kognitiver als auch motivational-affektiver Bildungsziele als wirksam erwiesen haben (Müller, Prenzel, Seidel, Schiepe-Tiska & Kjærnsli, im Druck). Das Ziel ist es, den Unterricht international zu beschreiben und den Zusammenhang zwischen Merkmalen der Unterrichtsqualität und mehrdimensionalen Bildungszielen zu untersuchen.

Die Unterrichtsmerkmale werden in PISA ausschließlich durch Befragungen der Schülerinnen und Schüler erfasst. Es wird demnach das von ihnen wahrgenommene Unterrichtsangebot erhoben. Es kann davon ausgegangen werden, dass Schülerinnen und Schüler Unterrichtsprozesse auf aggregierter Ebene durchaus reliabel beschreiben und bestimmte instruktionale Unterrichtsmerkmale auch valide beurteilen können (z. B. Lüdtke, Robitzsch, Trautwein & Kunter, 2009). Dazu müssen die Merkmale ereignisbezogen und verhaltensnah erfragt werden (Prenzel & Lankes, 2013). Dennoch liefern die

© Waxmann Verlag GmbH

Angaben der Schülerinnen und Schüler eine eingeschränkte Perspektive im Hinblick auf die prinzipiell beobacht- und wahrnehmbaren Unterrichtsmerkmale, die beispielsweise durch Angaben der Lehrkraft zum Unterrichtsangebot oder Beobachtungen aus Videostudien ergänzt werden können, um die didaktische Qualität des Unterrichts einschätzen sowie Leistungsunterschiede besser vorhersagen zu können (Kunter & Baumert, 2006).

Zusätzlich hängt die Wahrnehmung der unterrichtlichen Prozesse stark vom jeweiligen Lernkontext ab. So können sich die Beschreibungen der Schülerinnen und Schüler sowohl innerhalb von Staaten, etwa aufgrund unterschiedlicher Schularten, als auch zwischen den Staaten, etwa aufgrund des kulturellen Hintergrundes, unterscheiden. Speziell im naturwissenschaftlichen Unterricht kommt hinzu, dass der Unterricht zwischen und zum Teil auch innerhalb der Staaten ausgesprochen heterogen organisiert sein kann und sich der Referenzrahmen für die Schülerinnen und Schüler unterscheidet. Während in manchen Staaten in einzelnen Disziplinen wie etwa Biologie oder Chemie unterrichtet wird, findet der Unterricht in anderen Staaten in disziplinübergreifenden Fächerverbünden statt. Zudem ist der Umfang des wöchentlichen naturwissenschaftlichen Unterrichts zwischen den Staaten uneinheitlich (vgl. Kapitel 5).

Außerdem können die in PISA erfassten Unterrichtsmerkmale aufgrund des Stichprobendesigns nicht auf der Ebene der Klassen aggregiert werden. Die vorliegenden Analysen beruhen auf der internationalen altersbasierten Stichprobe. Befragt wurden daher, anders als bei einem klassenbasierten Sampling, keine intakten Schulklassen, sondern eine Auswahl Fünfzehnjähriger innerhalb einer Schule, unabhängig von ihrer Klassenzugehörigkeit. Die Analysen zum Unterricht beziehen sich daher nicht auf die Ebene der Klasse, sondern den Unterricht der fünfzehnjährigen Schülerinnen und Schüler. Infolgedessen können die Analysen nicht auf Klassenebene zusammengefasst werden. Die in Deutschland zusätzlich erhobenen Daten von Neuntklässlerinnen und Neuntklässlern (vgl. Kapitel 1) ermöglichen solche Mehrebenenanalysen und können das Bild des naturwissenschaftlichen Unterrichts in Deutschland ergänzen. Im vorliegenden Beitrag soll der Unterricht in Deutschland jedoch international eingeordnet werden und konzentriert sich deshalb auf den internationalen Datensatz der Fünfzehnjährigen. Auch die zusätzliche Betrachtung des erstmals international eingesetzten Lehrerfragebogens ist Bestandteil weiterführender Analysen und wird im Folgenden nicht näher in den Blick genommen.

Schließlich sind die Analysen zum Zusammenhang zwischen Merkmalen des naturwissenschaftlichen Unterrichts und der naturwissenschaftlichen Kompetenz der Schülerinnen und Schüler insofern problematisch, als dass Kompetenz das Resultat eines über mehrere Schuljahre hinweg andauernden kumulativen Lernprozesses ist (Klieme, 2013). Die im Fragebogen erfassten Unterrichtsmerkmale beziehen sich hingegen lediglich auf das laufende Schuljahr. Hohe Kompetenzwerte hängen unter Umständen nur bedingt mit dem Unterricht zusammen und resultieren vielmehr aus mehreren vorangegangenen Schuljahren.

Für die Analyse der Unterrichtsmerkmale in PISA folgt daraus, dass sich auf der Grundlage der Unterrichtsmerkmale motivational-affektive sowie Leistungsunterschiede

© Waxmann Verlag GmbH

nur bedingt vorhersagen lassen. Der Schwerpunkt der vorliegenden Analysen liegt daher auf deskriptiven Befunden. Zudem werden einzelne Merkmale anhand eines typologischen Ansatzes genauer in den Blick genommen. Es werden keine einfachen Korrelationen zwischen Unterrichtsmerkmalen und mehrdimensionalen Bildungszielen dargestellt, die eine Fehlinterpretation der Zusammenhänge nahelegen könnten. Wir berichten vielmehr einige anspruchsvollere Analysen, welche die Beziehungen zwischen Unterrichtsmerkmalen und naturwissenschaftlicher Kompetenz sowie Freude und Interesse an Naturwissenschaften untersuchen.

4.2.2 Eingesetzte Skalen

Disziplin im Klassenzimmer

Als eine Facette der Klassenführung wurde das Ausmaß von Störungen im Klassenzimmer erfasst – ein Merkmal, das in PISA als „Disziplin im Klassenzimmer" bezeichnet wird. Mit fünf Fragen wurden die Schülerinnen und Schüler gebeten, Aussagen zur Disziplin in ihrem Naturwissenschaftsunterricht auf einer vierstufigen Antwortskala („in allen Stunden", „in den meisten Stunden", „in einigen Stunden", „nie oder fast nie") einzuschätzen (Cronbachs α = .88[2]). Dazu sollten sie angeben, wie häufig Folgendes im Unterricht vorkommt: (a) „Die Schülerinnen und Schüler hören der Lehrerin/dem Lehrer nicht zu.", (b), „Im Klassenzimmer ist es oft laut und es geht drunter und drüber.", (c) „Die Lehrerin/der Lehrer muss lange warten, bis die Schülerinnen und Schüler ruhig werden.", (d) „Die Schülerinnen und Schüler können nicht ungestört arbeiten." und (e) „Die Schülerinnen und Schüler fangen erst lange nach Beginn der Stunde an zu arbeiten." (Frey et al., 2009; Klieme et al., 2001; OECD, 2009).

Unterstützung durch die Lehrkraft

Um zu erfahren, inwieweit sich die Schülerinnen und Schüler durch ihre Lehrkraft im naturwissenschaftlichen Unterricht unterstützt fühlen, schätzten sie folgende Aussagen ein (Cronbachs α = .89; Antwortskala „in allen Stunden", „in den meisten Stunden", „in einigen Stunden", „nie oder fast nie"): (a) „Unsere Lehrerin/unser Lehrer interessiert sich für den Lernfortschritt jeder einzelnen Schülerin/jedes einzelnen Schülers.", (b) „Unsere Lehrerin/unser Lehrer unterstützt uns zusätzlich, wenn wir Hilfe brauchen.", (c) „Unsere Lehrerin/unser Lehrer unterstützt uns beim Lernen.", (d) „Unsere Lehrerin/unser Lehrer erklärt etwas so lange, bis wir es verstehen." und (e) „Unsere Lehrerin/unser Lehrer gibt uns Gelegenheit, unsere Meinung zu sagen." (Frey et al., 2009; OECD, 2009).

2 Cronbachs Alpha wurde jeweils auf der Basis der Stichprobe in Deutschland ermittelt.

© Waxmann Verlag GmbH

Rückmeldung (formatives Assessment)

Wie häufig Jugendliche über Rückmeldungen zu ihren individuellen Stärken und Schwächen sowie Hinweise zum Weiterlernen durch die Lehrkraft berichten, wurde mithilfe der folgenden Aussagen eingeschätzt (Cronbachs α = .90; Antwortskala „nie oder fast nie", „in einigen Stunden", „in den meisten Stunden", „in allen Stunden"): (a) „Die Lehrkraft informiert mich über meine Leistung in diesem Fach.", (b) „Die Lehrkraft sagt mir, wo meine Stärken im naturwissenschaftlichen Unterricht liegen." (c) „Die Lehrkraft sagt mir, in welchen Punkten ich mich noch verbessern kann.", (d) „Die Lehrkraft sagt mir, wie ich meine Leistungen verbessern kann." und (e) „Die Lehrkraft berät mich, wie ich meine Lernziele erreichen kann." (Bayer, Klieme & Jude, im Druck).

Differenzierung

Mit drei Fragen wurde erhoben, wie häufig die Lehrkraft den Unterricht an die individuellen Bedürfnisse der Schülerinnen und Schüler anpasst (Cronbachs α = .79; Antwortskala „nie oder fast nie", „in einigen Stunden", „in den meisten Stunden", „in allen Stunden"): (a) „Die Lehrkraft passt den Unterricht den Bedürfnissen und dem Wissensstand der Klasse an.", (b) „Die Lehrkraft bietet individuelle Unterstützung an, wenn eine Schülerin oder ein Schüler Schwierigkeiten hat, ein Thema oder eine Aufgabe zu verstehen." sowie (c) „Die Lehrkraft ändert den Unterricht, wenn die meisten Schülerinnen und Schüler Mühe haben, das Thema zu verstehen." (Bayer, Klieme & Jude, im Druck).

Forschend-entdeckender Unterricht

Mit insgesamt neun Fragen wurden unterschiedliche Aktivitäten des forschend-entdeckenden Unterrichts erhoben (Cronbachs α = .85), die auf einer vierstufigen Antwortskala („in allen Stunden", „in den meisten Stunden", „in einigen Stunden", „nie oder fast nie") eingeschätzt werden mussten. S*oziale Aktivitäten* wurden durch folgende Aussagen erhoben: (a) „Schülerinnen und Schüler bekommen Gelegenheit, ihre Ideen zu erklären." (b) „Die Klasse diskutiert über Untersuchungen." und (c) „Die Schülerinnen und Schüler werden aufgefordert, über naturwissenschaftliche Fragen zu diskutieren.". Für *prozedurale Aktivitäten* nahmen die Jugendlichen Stellung zu folgenden Aussagen: (a) „Die Schülerinnen und Schüler führen praktische Experimente im Labor durch.", (b) „Die Schülerinnen und Schüler dürfen ihre eigenen Experimente entwickeln." und (c) „Die Schülerinnen und Schüler sollen eine Untersuchung durchführen, um Ideen zu überprüfen". Im Bereich *epistemische Aktivitäten* beurteilten die Fünfzehnjährigen folgende Aussage: (a) „Die Schülerinnen und Schüler sollen Schlüsse aus einem Experiment ziehen, das sie durchgeführt haben". Mit zwei weiteren Fragen wurden das *Modellieren naturwissenschaftlicher Anwendungen* sowie der *Lebensweltbezug* erfasst: (a) „Der Lehrer/die Lehrerin erklärt, wie ein im Unterricht erlerntes naturwissenschaftliches Prinzip auf eine Reihe verschiedener Phänomene angewendet werden kann (z. B. auf

© Waxmann Verlag GmbH

die Bewegung von Objekten oder auf Substanzen mit ähnlichen Eigenschaften)." und (b) „Der Lehrer/die Lehrerin erklärt deutlich die Wichtigkeit von naturwissenschaftlichen Konzepten für unser Leben." (Frey et al., 2009; Kobarg et al., 2011; Müller et al., im Druck; OECD, 2009; Seidel & Prenzel, 2006a).

4.3 Ergebnisse

Im Folgenden wird zunächst der naturwissenschaftliche Unterricht in Deutschland beschrieben und international eingeordnet. Vertiefende Analysen betreffen Unterschiede zwischen dem Unterricht an Gymnasien und dem an nicht gymnasialen Schularten. Es schließt sich ein Vergleich von Merkmalen des forschend-entdeckenden Lernens zwischen PISA 2006 und PISA 2015 an. Weiterführende Analysen identifizieren Muster des forschend-entdeckenden Unterrichts in Deutschland und gehen auf das Zusammenspiel mit domänenübergreifenden Qualitätsmerkmalen ein. Diese Muster werden anschließend mit mehrdimensionalen Bildungszielen – naturwissenschaftlicher Kompetenz und Freude und Interesse an Naturwissenschaften – in Verbindung gebracht.

4.3.1 Naturwissenschaftlicher Unterricht im internationalen Vergleich

Um der Frage nachzugehen, ob es im Vergleich zu anderen Staaten Besonderheiten in der Unterrichtswahrnehmung der Schülerinnen und Schüler in Deutschland gibt, wurden die Angaben der Jugendlichen im Vergleich zum Mittelwert aller OECD-Staaten betrachtet. Dieser ist auf 0 mit einer Standardabweichung von 1 festgelegt (vgl. Kapitel 12). Da der Durchschnitt Kennwerte von Staaten unterschiedlicher Kulturräume zusammenfasst, ist dieser vorsichtig zu interpretieren (vgl. Abschnitt 4.2.1 sowie Kapitel 13). Daher wird in diesem Kapitel auf einen Vergleich mit allen OECD-Staaten verzichtet. Stattdessen wurden für einen Vergleich Staaten herangezogen, die einem ähnlichen Kulturraum angehören wie Deutschland, da hier weniger unterschiedliche kulturspezifische Antworttendenzen zu erwarten sind (Johnson, Shavitt & Holbrook, 2011). Von besonderem Interesse sind dabei Staaten, die überdurchschnittliche Kompetenzwerte erzielen und denen die Förderung naturwissenschaftlicher Kompetenz sowohl in der Breite als auch mit Blick auf die Gruppe hochkompetenter Jugendlicher besonders gut gelingt. Dazu gehören Kanada, Finnland und Estland. In Estland werden außerdem die Naturwissenschaften – ähnlich wie an vielen Schulen in Deutschland – nach Fächern getrennt unterrichtet.

Darüber hinaus wurden für den Vergleich Staaten ausgewählt, die unterschiedliche Unterrichtstraditionen aufweisen. Der Unterricht in Kanada und dem Vereinigten Königreich zeichnet sich zum Beispiel durch eine starke Orientierung an internalen („*Minds-on*") und externalen („*Hands-on*") Zielen aus. Bisherige Studien zeigen, dass in diesen Staaten ein Schwerpunkt im Unterricht auf dem Experimentieren liegt und Schü-

© Waxmann Verlag GmbH

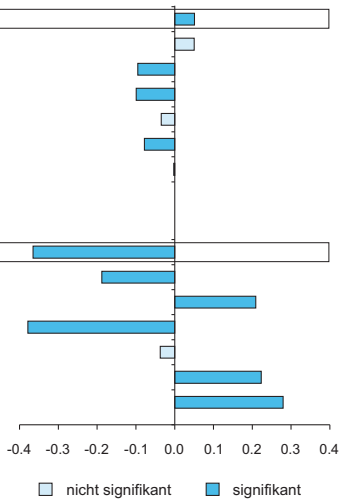
lerinnen und Schüler regelmäßig aufgefordert sind, aus diesen eigene Schlüsse zu ziehen (Seidel et al., 2007). In der Schweiz, Finnland und den Niederlanden ist der naturwissenschaftliche Unterricht dagegen eher lehrerzentriert. Hier liegt der Schwerpunkt ähnlich wie in Deutschland eher auf dem Klassengespräch (vgl. auch Fischer, Labudde, Neumann & Virii, 2014; Kobarg et al., 2011).

Domänenübergreifende Qualitätsmerkmale

Die Ausprägungen der domänenübergreifenden Unterrichtsmerkmale zeigen die Abbildungen 4.1 und 4.2. Positive Skalenwerte geben an, dass das Merkmal im OECD-Vergleich überdurchschnittlich ausgeprägt ist, negative Skalenwerte weisen auf eine unterdurchschnittliche Merkmalsausprägung hin. Damit ist ein Vergleich der Staaten untereinander möglich, es kann jedoch keine Aussage über die absolute Höhe der Merkmalsausprägungen getroffen werden. Deshalb werden zusätzlich für jedes Unterrichtsmerkmal ein charakteristisches Beispielitem und die dazugehörige prozentuale Zustimmung zu den entsprechenden Antwortkategorien angegeben, um das Antwortverhalten genauer in den Blick nehmen zu können. Die Items wurden so ausgewählt, dass sie die Verteilung der relativen Häufigkeitsangaben für die jeweilige Skala möglichst gut repräsentieren.

Vergleicht man die domänenübergreifenden Basisdimensionen der Unterrichtsqualität, so fällt auf, dass die *Disziplin* im naturwissenschaftlichen Unterricht in Deutsch-

Unterrichtsmerkmal	Zustimmung in %[1]	OECD-Staaten	*M*	*(SE)*	*SD*
Disziplin Im Klassenzimmer ist es oft laut, und es geht drunter und drüber.	26.4	Deutschland	0.05	(0.02)	0.94
	33.2	Schweiz	0.05	(0.03)	1.04
	38.1	Finnland	-0.10	(0.02)	0.88
	33.0	Niederlande	-0.10	(0.02)	0.81
	30.2	Estland	-0.04	(0.02)	0.91
	39.0	Vereinigtes Königreich	-0.08	(0.02)	1.03
	36.0	Kanada	-0.01	(0.02)	1.02
Unterstützung Unsere Lehrerin/unser Lehrer erklärt etwas so lange, bis wir es verstehen.	60.7	Deutschland	-0.37	(0.02)	0.97
	65.5	Schweiz	-0.20	(0.02)	0.96
	74.5	Finnland	0.20	(0.02)	0.89
	61.8	Niederlande	-0.39	(0.02)	0.86
	65.5	Estland	-0.05	(0.02)	0.96
	74.4	Vereinigtes Königreich	0.21	(0.02)	0.97
	75.4	Kanada	0.27	(0.02)	0.99



[1] Die Antwortkategorien „in allen Stunden“ und „in den meisten Stunden“ wurden zusammengefasst.

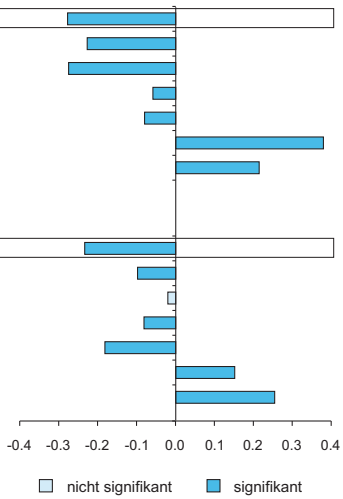
Abbildung 4.1: Disziplin im Klassenzimmer und wahrgenommene Unterstützung im internationalen Vergleich

© Waxmann Verlag GmbH

land leicht über dem OECD-Durchschnitt liegt (Abbildung 4.1). Dennoch stimmt noch immer ein Viertel der Schülerinnen und Schüler der Aussage zu, dass es in allen oder den meisten Stunden im Klassenzimmer „laut ist und drunter und drüber geht". In den Vergleichsstaaten berichten die Fünfzehnjährigen in Finnland, den Niederlanden und dem Vereinigten Königreich von einer eher unterdurchschnittlich ausgeprägten Disziplin im Klassenzimmer. Auch wenn die Unterschiede statistisch signifikant ausfallen, so sind sie im Vergleich zu den Unterschieden in der *wahrgenommenen Unterstützung* durch die Lehrkraft gering. Hier liegt der Mittelwert für Deutschland deutlich unterhalb des OECD-Durchschnitts. Die Fünfzehnjährigen in Deutschland fühlen sich demnach im Naturwissenschaftsunterricht deutlich seltener durch ihre Lehrkraft unterstützt als Gleichaltrige in anderen OECD-Staaten. Ähnlich negativ fällt diese Wahrnehmung in den Niederlanden aus. Auch die Jugendlichen aus der Schweiz berichten über eine unterdurchschnittlich wahrgenommene Unterstützung. In Finnland, Kanada und dem Vereinigten Königreich hingegen fühlen sich die Fünfzehnjährigen von ihrer Lehrkraft angemessen unterstützt. Betrachtet man die absolute Höhe der Zustimmung, so relativiert sich dieses Bild etwas. In Deutschland berichten ungefähr 60 Prozent der Schülerinnen und Schüler, dass in allen oder den meisten Stunden die Lehrkraft etwas so lange erklärt, bis sie es verstanden haben. Dennoch zeigt der Blick in die Vergleichsstaaten, dass es hier Verbesserungspotenzial gibt. In Finnland, dem Vereinigten Königreich und Kanada stimmen dieser Aussage drei Viertel aller Schülerinnen und Schüler zu.

Die Unterrichtsmerkmale *Rückmeldung* und *Differenzierung* liegen in Deutschland ebenfalls unterhalb des OECD-Durchschnitts (Abbildung 4.2). Jugendliche in Deutsch-

Unterrichtsmerkmal	Zustimmung in %[1]	OECD-Staaten	*M*	*(SE)*	*SD*
Rückmeldung Die Lehrkraft sagt mir, wie ich meine Leistungen verbessern kann.	21.0	Deutschland	-0.28	(0.01)	0.89
	25.7	Schweiz	-0.23	(0.02)	1.02
	18.8	Finnland	-0.27	(0.02)	0.92
	27.1	Niederlande	-0.06	(0.02)	0.88
	29.6	Estland	-0.08	(0.02)	0.93
	42.3	Vereinigtes Königreich	0.37	(0.02)	0.93
	38.9	Kanada	0.21	(0.02)	1.00
Differenzierung Die Lehrkraft passt den Unterricht den Bedürfnissen und dem Wissensstand der Klasse an.	42.7	Deutschland	-0.22	(0.02)	0.98
	45.9	Schweiz	-0.09	(0.02)	1.01
	37.5	Finnland	-0.01	(0.02)	0.92
	31.3	Niederlande	-0.07	(0.02)	0.86
	30.4	Estland	-0.17	(0.02)	0.92
	47.5	Vereinigtes Königreich	0.15	(0.02)	0.97
	51.6	Kanada	0.26	(0.02)	1.01



[1] Die Antwortkategorien „in allen Stunden" und „in den meisten Stunden" wurden zusammengefasst.

Abbildung 4.2: Rückmeldung und Differenzierung im internationalen Vergleich

© Waxmann Verlag GmbH

land nehmen weniger individuelle Rückmeldungen zu ihrer Leistung wahr. Nur ein Fünftel berichtet, dass ihnen die Lehrkraft im Naturwissenschaftsunterricht in allen oder den meisten Stunden Rückmeldung gibt, wie sie sich verbessern können. Über vergleichbar wenig individuelle Rückmeldung berichten die Jugendlichen in Finnland und der Schweiz. Entsprechend verhält es sich bei der wahrgenommenen *Differenzierung* im naturwissenschaftlichen Unterricht, denn auch diese scheint an Schulen in Deutschland weniger verbreitet zu sein. Etwas mehr als die Hälfte der Fünfzehnjährigen gibt an, dass die Lehrkraft nur in einigen Stunden oder sogar nie den Unterricht an die Bedürfnisse und den Wissensstand der Klasse anpasst. Im Einklang mit den Ergebnissen zur individuellen Rückmeldung zeichnen sich vor allem das Vereinigte Königreich und Kanada durch häufige Differenzierung im Naturwissenschaftsunterricht aus.

Insgesamt zeigt sich für Deutschland bei den domänenübergreifenden Qualitätsmerkmalen, dass die Schülerinnen und Schüler den naturwissenschaftlichen Unterricht als eher störungsarm und diszipliniert wahrnehmen. Es berichtet aber dennoch immerhin ein Viertel, dass es im Unterricht häufiger laut und wenig diszipliniert zugeht. Hinsichtlich der wahrgenommenen Unterstützung, Rückmeldungen und der Anpassung des Unterrichts an die individuellen Bedürfnisse der Schülerinnen und Schüler zeichnet sich im OECD-Vergleich ein wenig positives Bild ab: Diese Merkmale sind in Deutschland unterdurchschnittlich ausgeprägt. Zwar relativiert die Betrachtung der charakteristischen Items dieses Bild für die wahrgenommene Unterstützung, denn 60 Prozent geben an, dass sie sich durch ihre Lehrkraft unterstützt fühlen, dennoch bleiben 40 Prozent, die sich nur in einigen Stunden oder nie unterstützt fühlen. Für die wahrgenommene Differenzierung und die individuellen Rückmeldungen bleibt ein nicht unproblematisches Bild bestehen. Nicht einmal die Hälfte der Jugendlichen berichtet, dass der naturwissenschaftliche Unterricht in allen oder den meisten Stunden an die Bedürfnisse und den Wissensstand der Klasse angepasst wird. Nur ein Fünftel bekommt darüber hinaus regelmäßig Rückmeldung dazu, wie sie ihre Leistungen verbessern können. In den Vergleichsstaaten finden sich besonders positiv ausgeprägte domänenübergreifende Unterrichtsmerkmale in Kanada und im Vereinigten Königreich. Diese beiden Länder verdeutlichen, dass – fragt man die Schülerinnen und Schüler – durchaus Verbesserungspotenzial für den naturwissenschaftlichen Unterricht in Deutschland gegeben ist.

Forschend-entdeckender Unterricht

Abbildung 4.3 zeigt die Ausprägungen der Skalenmittelwerte für die Gesamtskala forschend-entdeckendes Lernen in Deutschland sowie in den ausgewählten Vergleichsstaaten. Insgesamt kommt forschend-entdeckender Unterricht in Deutschland im Vergleich zum OECD-Durchschnitt etwas häufiger vor. Auch in der Schweiz und vor allem in Kanada berichten die Jugendlichen über einen wesentlichen Anteil forschend-entdeckenden Lernens in ihrem naturwissenschaftlichen Unterricht. Geringer im Vergleich zum OECD-Durchschnitt ist der Anteil in Estland sowie den Niederlanden und Finnland.

© Waxmann Verlag GmbH

Unterrichtsmerkmal	OECD-Staaten	*M*	*(SE)*	*SD*
Forschend-entdeckender Unterricht	Deutschland	0.06	(0.02)	0.88
	Schweiz	0.15	(0.02)	0.95
	Finnland	-0.30	(0.02)	0.86
	Niederlande	-0.25	(0.02)	0.95
	Estland	-0.07	(0.02)	0.83
	Vereinigtes Königreich	-0.01	(0.01)	0.84
	Kanada	0.27	(0.01)	0.97

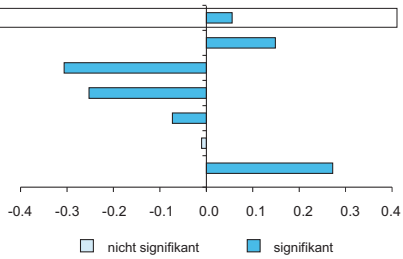


Abbildung 4.3: Forschend-entdeckender Unterricht im internationalen Vergleich

Einen differenzierten Blick auf die verschiedenen Dimensionen des forschend-entdeckenden Unterrichts – soziale Aktivitäten, prozedurale Aktivitäten, epistemische Aktivitäten und Anwendungsbezug – ermöglicht die Betrachtung der absoluten Höhe der Merkmalsausprägungen. Dazu werden die prozentualen Anteile der Fünfzehnjährigen dargestellt, die in allen oder den meisten Unterrichtsstunden diese Aktivitäten wahrnehmen.

Tabelle 4.1 gibt an, wie häufig die Schülerinnen und Schüler über *soziale Lernaktivitäten* in ihrem naturwissenschaftlichen Unterricht berichten. Kommunikation ist ein zentraler Aspekt der Naturwissenschaften, um Informationen zu vermitteln, Vorstellungen anderer kennenzulernen, eigene Vorstellungen zu entwickeln sowie Diskussionen zu führen und Rückmeldungen zu geben. So können Annahmen und Hypothesen formuliert, Ergebnisse dargestellt und anschließend interpretiert und diskutiert werden (KMK, 2005a, 2005b, 2005c).

Die Schülerinnen und Schüler wurden daher im Rahmen von PISA 2015 gefragt, in welchem Ausmaß sie eigene Ideen erklären können und wie häufig naturwissenschaftliche Fragen sowie Experimente diskutiert werden. Im Durchschnitt geben 69 Prozent der Fünfzehnjährigen in den OECD-Staaten an, dass sie in allen oder den meisten Stunden ihre eigenen Ideen äußern dürfen. Diskussionen über Fragen oder Untersuchungen finden hingegen seltener statt. 30 Prozent der Jugendlichen berichten, dass sie aufgefordert werden, über naturwissenschaftliche Fragen zu sprechen, ungefähr ein Viertel gibt an, dass man in allen oder den meisten Stunden Untersuchungen diskutiert.

Auch der naturwissenschaftliche Unterricht in Deutschland ist durch soziale Lernaktivitäten geprägt. 70 Prozent der Schülerinnen und Schüler geben an, dass sie in allen oder den meisten Stunden eigene Ideen und Meinungen äußern können. Darüber hinaus finden 30 Prozent, dass sie in allen oder den meisten Stunden aufgefordert sind, über naturwissenschaftliche Fragen zu diskutieren. Experimente zu diskutieren kommt in Deutschland im Vergleich zum OECD-Durchschnitt mit 38 Prozent zustimmenden Antworten sogar signifikant häufiger vor.

© Waxmann Verlag GmbH

Tabelle 4.1: *Soziale Lernktivitäten* im internationalen Vergleich

OECD-Staaten	Ideen erklären		Naturwissenschaftliche Fragen diskutieren		Experimente diskutieren	
	Zustimmung in %[1]	*(SE)*	Zustimmung in %[1]	*(SE)*	Zustimmung in %[1]	*(SE)*
Deutschland	70.0	(0.8)	30.1	(0.8)	38.0	(0.8)
Schweiz	71.5	(0.9)	36.3	(1.1)	40.6	(1.1)
Finnland	71.6	(0.8)	15.0	(0.7)	12.3	(0.7)
Niederlande	52.9	(0.9)	18.8	(0.6)	14.7	(0.7)
Estland	70.3	(0.8)	15.0	(0.6)	30.5	(0.9)
Vereinigtes Königreich	73.9	(0.8)	17.7	(0.7)	14.6	(0.6)
Kanada	77.1	(0.6)	30.0	(0.6)	25.1	(0.6)
OECD-Durchschnitt	68.5	(0.1)	29.7	(0.1)	26.2	(0.1)

[1] Die Antwortkategorien „in allen Stunden“ und „in den meisten Stunden“ wurden zusammengefasst.

Auch in den meisten der hier einbezogenen Vergleichsstaaten nennen die Jugendlichen besonders häufig die Möglichkeit, eigene Ideen im Unterricht erklären zu können (über 70 Prozent). In Kanada ist die Quote mit 77 Prozent am höchsten, in den Niederlanden mit 53 Prozent am geringsten. Auch die sozialen Aktivitäten, also das Diskutieren von naturwissenschaftlichen Fragen und Experimenten, kommen in den Niederlanden im Vergleich zum OECD-Durchschnitt signifikant seltener vor. Auch die Fünfzehnjährigen in Finnland und dem Vereinigten Königreich führen im Vergleich zum OECD-Durchschnitt signifikant seltener Diskussionen in ihrem Unterricht. Besonders häufig sind aus Sicht der Jugendlichen soziale Aktivitäten im Naturwissenschaftsunterricht in der Schweiz vorzufinden. Die prozentualen Häufigkeiten liegen alle über dem OECD-Durchschnitt. Ähnlich wie in Deutschland geben 40 Prozent der Fünfzehnjährigen an, dass sie im Unterricht über Untersuchungen diskutieren. Etwas häufiger als in Deutschland werden auch naturwissenschaftliche Fragen der Schülerinnen und Schüler diskutiert (36 Prozent).

Prozedurale Lernaktivitäten beziehen sich vor allem auf das Experimentieren als bedeutsame Aktivität in den Naturwissenschaften und sollten deshalb ein zentrales Element im naturwissenschaftlichen Unterricht sein (KMK, 2005a, 2005b, 2005c; Labudde & Möller, 2015). Experimente können dabei unterschiedliche Ziele verfolgen: Sie können ein naturwissenschaftliches Phänomen demonstrieren, eine Fragestellung umsetzen, Daten generieren, um daraus Schlüsse und Interpretationen abzuleiten, oder den Umgang mit naturwissenschaftlichen Arbeitsweisen und Geräten zeigen und einüben (Hofstein & Lunetta, 2004). Im forschend-entdeckenden Unterricht geht es jedoch nicht nur darum, dass Schülerinnen und Schüler Gelegenheiten haben, praktische Experimente selbst durchzuführen, sondern auch, dass sie eigene naturwissenschaftliche Fragestellungen formulieren, und selber Untersuchungen planen, um Ideen zu überprüfen (Labudde & Börlin, 2013).

© Waxmann Verlag GmbH

Welchen Stellenwert forschendes Lernen im naturwissenschaftlichen Unterricht und insbesondere Experimente in den ausgewählten Staaten besitzen, zeigt Tabelle 4.2. Im OECD-Durchschnitt berichtet ein Fünftel der Fünfzehnjährigen, dass man in allen oder den meisten Stunden praktische Experimente im Labor durchführt. Eigene Ideen zu überprüfen, kommt in einem Viertel der Stunden vor. Am seltensten in ihrem Unterrichtsalltag dürfen die Jugendlichen eigene Experimente entwickeln.

Tabelle 4.2: *Prozedurale Lernktivitäten* im internationalen Vergleich

OECD-Staaten	Im Labor experimentieren		Ideen austesten		Experimente selbst entwickeln	
	Zustimmung in %[1]	*(SE)*	Zustimmung in %[1]	*(SE)*	Zustimmung in %[1]	*(SE)*
Deutschland	21.6	(0.8)	28.7	(0.9)	12.8	(0.5)
Schweiz	24.5	(1.0)	27.4	(1.1)	17.3	(0.7)
Finnland	20.8	(0.8)	12.8	(0.5)	6.2	(0.4)
Niederlande	29.4	(0.8)	16.7	(0.7)	11.3	(0.6)
Estland	9.0	(0.6)	14.6	(0.7)	11.6	(0.6)
Vereinigtes Königreich	18.6	(0.8)	29.7	(0.9)	9.4	(0.5)
Kanada	28.7	(0.6)	35.7	(0.7)	21.0	(0.7)
OECD-Durchschnitt	20.9	(0.1)	25.8	(0.1)	15.7	(0.1)

[1] Die Antwortkategorien „in allen Stunden“ und „in den meisten Stunden“ wurden zusammengefasst.

In Deutschland unterscheidet sich die Häufigkeit, mit der Experimente im Labor durchgeführt werden (22 Prozent), nicht vom OECD-Durchschnitt. Die Jugendlichen berichten jedoch, dass sie häufiger als im OECD-Durchschnitt Untersuchungen durchführen, um Ideen zu überprüfen. Fast 30 Prozent geben an, dass dies in allen oder den meisten Stunden vorkommt. Selbstständig Experimente zu entwickeln kommt im naturwissenschaftlichen Unterricht im Vergleich zu Laborexperimenten und Ideen austesten seltener vor. Nur 13 Prozent der Fünfzehnjährigen stimmen zu, dass dies in allen oder den meisten Stunden vorkommt. Dieser Wert liegt ebenfalls unterhalb des OECD-Durchschnitts.

In den ausgewählten Vergleichsstaaten dürfen besonders häufig die Schülerinnen und Schüler in Kanada und im Vereinigten Königreich eigene Experimente entwickeln. In Kanada – einem Bildungssystem, in dem Jugendliche herausragende naturwissenschaftliche Kompetenzen und hohe Werte im Bereich der motivational-affektiven Lernergebnisse erreichen (vgl. Kapitel 2 und 3) – berichtet ein Fünftel der Jugendlichen, dass dies in allen oder den meisten Stunden vorkommt. Darüber hinaus dürfen sie deutlich häufiger (36 Prozent) Untersuchungen durchführen, um Ideen zu überprüfen, und im Labor experimentieren (29 Prozent). Hier scheint der Prozess des naturwissenschaftlichen Forschens einen besonderen Stellenwert im Unterricht zu haben – entsprechend der Bedeutung, die dem forschend-entdeckenden Lernen international beigemessen wird (Hazelkorn et al., 2015; Osborne & Dillon, 2008). In Estland und Finnland – zwei weiteren

© Waxmann Verlag GmbH

Spitzenstaaten im Bereich naturwissenschaftlicher Kompetenz – sowie den Niederlanden stehen jedoch gerade prozedurale Lernaktivitäten, die über das Durchführen von praktischen Experimenten im Labor hinausgehen, wie Ideen überprüfen und für diese selbst passende Untersuchungen planen, seltener im Fokus des naturwissenschaftlichen Unterrichts.

Epistemische Aktivitäten beziehen sich darauf, inwieweit Schülerinnen und Schüler im Unterricht die Gelegenheit haben, Befunde aus naturwissenschaftlichen Untersuchungen selbstständig zu evaluieren und zu interpretieren. Diese für die Entwicklung mehrdimensionaler Bildungsziele besonders bedeutsamen Aktivitäten (Furtak et al., 2012) wurden im Rahmen von PISA 2015 mit der Frage erfasst, wie häufig Fünfzehnjährige aus einem Experiment, das sie durchgeführt haben, eigene Schlüsse ziehen sollen.

Im OECD-Durchschnitt geben 42 Prozent der Jugendlichen an, dass sie in allen oder den meisten Stunden aufgefordert werden, eigene Schlüsse zu ziehen. In Deutschland geben mit 59 Prozent signifikant mehr Schülerinnen und Schüler an, derartige Lerngelegenheiten im Unterricht vorzufinden (vgl. Tabelle 4.3). Auch in Kanada, der Schweiz und dem Vereinigten Königreich darf laut eigenen Aussagen ungefähr die Hälfte der Fünfzehnjährigen in allen oder den meisten Stunden Schlüsse aus Experimenten ziehen. In Finnland und den Niederlanden sowie in Estland kommt das Ziehen eigener Schlüsse im naturwissenschaftlichen Unterricht weniger häufig vor.

Tabelle 4.3: *Epistemische Lernaktivitäten*, *naturwissenschaftliches Modellieren* und *Lebensweltbezug* im internationalen Vergleich

OECD-Staaten	Eigene Schlüsse ziehen		Anwendung von Prinzipien		Relevanz für das Leben	
	Zustimmung in %[1]	*(SE)*	Zustimmung in %[1]	*(SE)*	Zustimmung in %[1]	*(SE)*
Deutschland	58.7	(1.0)	56.0	(1.0)	37.5	(0.7)
Schweiz	50.9	(1.1)	65.7	(0.9)	48.3	(1.1)
Finnland	36.7	(1.0)	53.7	(0.8)	48.3	(0.9)
Niederlande	38.5	(0.8)	45.7	(1.0)	39.2	(0.9)
Estland	29.7	(0.9)	57.8	(0.8)	56.8	(0.9)
Vereinigtes Königreich	48.5	(0.8)	60.6	(0.8)	48.0	(0.8)
Kanada	53.0	(0.4)	73.0	(0.6)	61.9	(0.7)
OECD-Durchschnitt	41.5	(0.2)	59.1	(0.1)	50.0	(0.1)

[1] Die Antwortkategorien „in allen Stunden“ und „in den meisten Stunden“ wurden zusammengefasst.

Das *Anwenden von Prinzipien auf naturwissenschaftliche Phänomene* sowie das Herstellen eines *Bezugs zur alltäglichen Lebenswelt* der Schülerinnen und Schüler sind wichtige Ansatzpunkte für den Aufbau naturwissenschaftlicher Kompetenz (Donovan & Bransford, 2005). Jugendliche sollten im Unterricht das Modellieren naturwissenschaftlicher Sachverhalte erfahren und den Umgang mit Modellvorstellungen kennenlernen. Die Fünfzehnjährigen wurden daher in PISA 2015 gefragt, wie häufig in ihrem Unterricht

© Waxmann Verlag GmbH

naturwissenschaftliche Prinzipien und Konzepte auf unterschiedliche Alltagsphänomene angewendet werden und deren Wichtigkeit für das persönliche Leben thematisiert wird.

Im OECD-Durchschnitt stellen das Anwenden von Prinzipien und das Herstellen eines Anwendungsbezugs wesentliche Bestandteile des naturwissenschaftlichen Unterrichts dar. Hier sagen 60 Prozent der Schülerinnen und Schüler, dass die Lehrkraft in allen oder den meisten Stunden erklärt, wie naturwissenschaftliche Prinzipien auf ein Phänomen angewendet werden können. Darüber hinaus gibt die Hälfte der Jugendlichen an, dass die Lehrkraft in allen oder den meisten Stunden erklärt, inwieweit naturwissenschaftliche Konzepte wichtig für das persönliche Leben sind.

In Deutschland kommt beides im naturwissenschaftlichen Unterricht signifikant seltener vor und scheint im Vergleich eine etwas geringere Rolle zu spielen. Zwar geben 56 Prozent der Fünfzehnjährigen an, dass in allen oder den meisten Stunden von der Lehrkraft erklärt wird, wie ein im Unterricht erlerntes naturwissenschaftliches Prinzip auf eine Reihe verschiedener Phänomene angewendet werden kann. Deutlich weniger Schülerinnen und Schüler – nämlich 38 Prozent – sind jedoch der Ansicht, dass ihnen die Wichtigkeit naturwissenschaftlicher Konzepte für das eigene Leben erklärt wird.

In den ausgewählten Vergleichsstaaten fällt erneut Kanada positiv auf. Hier meinen 73 Prozent der Jugendlichen, dass naturwissenschaftliches Modellieren in allen oder den meisten Unterrichtsstunden vorkommt, und 62 Prozent geben an, dass ihnen häufig erläutert wird, wie wichtig naturwissenschaftliche Konzepte für das persönliche Leben sind. Auch in Estland berichten dies 57 Prozent der Fünfzehnjährigen. In den Niederlanden stellen naturwissenschaftliches Modellieren und die Herstellung eines Anwendungsbezugs seltenere Bestandteile des Unterrichts dar. Das Modellieren wird von 46 Prozent der Jugendlichen, das Erläutern eines Anwendungsbezugs – ähnlich wie in Deutschland – von 39 Prozent wahrgenommen.

Zusammenfassend zeigt sich für Deutschland, dass forschend-entdeckender Unterricht gegenüber dem OECD-Durchschnitt signifikant öfter vorkommt. Ein Blick auf die Häufigkeiten einzelner fachspezifischer Lernaktivitäten zeigt ein differenzierteres Bild: Der naturwissenschaftliche Unterricht ist in Deutschland zumeist durch soziale Aktivitäten, wie das Einbringen eigener Ideen und das Diskutieren über naturwissenschaftliche Fragen oder Experimente, geprägt. Darüber hinaus experimentieren Schülerinnen und Schüler regelmäßig im Unterricht. Sie dürfen dabei Ideen austesten und vor allem strukturierte Laborexperimente durchführen. Im Vergleich dazu kommt das eigenständige Entwickeln von Experimenten seltener vor. Auch frühere nationale Analysen des naturwissenschaftlichen Unterrichts zeigten, dass der Schwerpunkt eher auf dem Durchführen von Experimenten unter Anleitung lag bzw. die Lehrkräfte naturwissenschaftliche Phänomene demonstrierten (z. B. Börlin & Labudde, 2014; Seidel et al., 2007). Das Ziehen eigener Schlüsse aus einem durchgeführten Experiment als ein zentrales Merkmal für den Aufbau naturwissenschaftlicher Kompetenz (Furtak et al., 2012) kommt hingegen häufiger vor als in den Vergleichsstaaten. Das Anwenden von Prinzipien auf naturwissenschaftliche Phänomene sowie das Herstellen eines Bezugs zur alltäglichen Lebens-

© Waxmann Verlag GmbH

welt der Schülerinnen und Schüler scheint im naturwissenschaftlichen Unterricht in Deutschland eine eher geringe Rolle zu spielen. Diese Befunde spiegeln damit nicht die verstärkten Bemühungen um den Lebensweltbezug in bundesweiten Innovationsprojekten wider (den sogenannten Kontext-Projekten), welche die Bedeutung des Anwendungsaspekts sowie die Einbettung in alltägliche Kontexte herausgestellt und Materialen zur Umsetzung für den Unterricht entwickelt haben (Bayrhuber et al., 2007; Demuth, Gräsel, Parchmann & Ralle, 2008: Mikelskis-Seifert & Duit, 2010).

4.3.2 Schulartspezifische Unterschiede in der Wahrnehmung des naturwissenschaftlichen Unterrichts in Deutschland

Im Folgenden werden die domänenübergreifenden und fachbezogenen Unterrichtsmerkmale im Gegensatz zu früheren PISA-Erhebungsrunden nur getrennt nach Gymnasium und nicht gymnasialen Schularten dargestellt. Außerhalb des gymnasialen Bereichs wurden zahlreiche Schulen mit einer Sekundarstufe I durch die Kombination von Bildungsgängen in neue oder neu bezeichnete Schularten überführt (Pant et al., 2013). Daher ist eine differenzierte Betrachtung der nicht gymnasialen, allgemeinbildenden Schularten nicht mehr tragfähig und sinnvoll (vgl. auch Kapitel 1).

Domänenübergreifende Qualitätsmerkmale

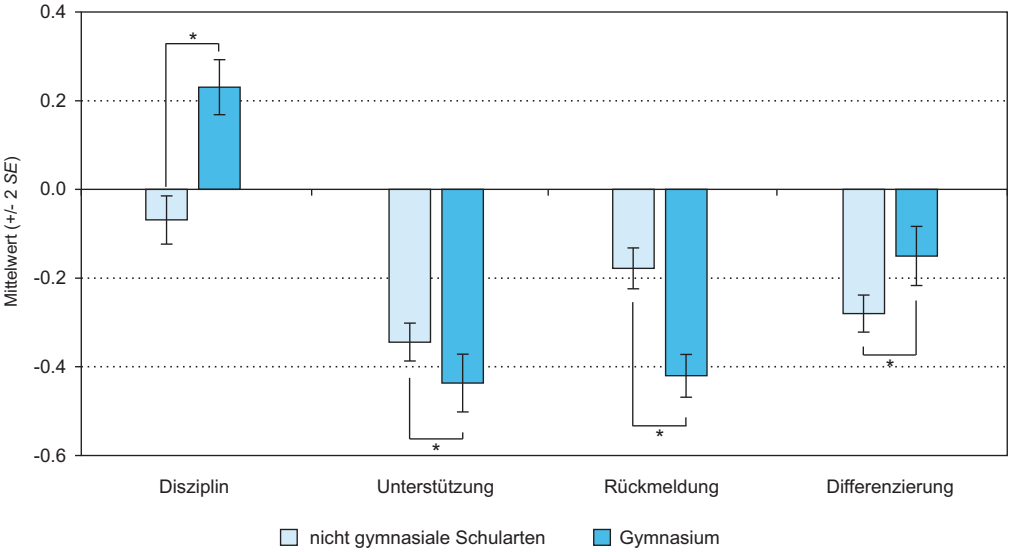
Abbildung 4.4 zeigt die Ausprägung der Skalen domänenübergreifender Unterrichtsmerkmale getrennt für das Gymnasium und nicht gymnasiale Schularten. Es fällt auf, dass der naturwissenschaftliche Unterricht am Gymnasium als deutlich störungsärmer wahrgenommen wird als an nicht gymnasialen Schularten.

Auch wenn in Deutschland im OECD-Vergleich die wahrgenommene Unterstützung, individuelle Rückmeldung und Differenzierung insgesamt unterdurchschnittlich ausgeprägt sind, zeigen sich doch Unterschiede, wenn man zwischen dem Gymnasium und nicht gymnasialen Schularten trennt. So berichten die Jugendlichen am Gymnasium häufiger über die Anpassung des Unterrichts an die Bedürfnisse und den Wissensstand der Klasse, sie fühlen sich jedoch weniger durch die Lehrperson unterstützt und erhalten seltener individuelle Rückmeldungen zu ihren Stärken und Schwächen als Jugendliche an nicht gymnasialen Schularten.

Forschend-entdeckender Unterricht

Betrachtet man die Aktivitäten des forschend-entdeckenden Unterrichts getrennt nach Schularten (Abbildung 4.5), so zeigt sich, dass am Gymnasium *soziale Aktivitäten* wie das Erklären von Ideen und das Diskutieren von naturwissenschaftlichen Fragestellungen häufiger vorkommen. Kein Unterschied findet sich bei der Diskussion von Experi-

© Waxmann Verlag GmbH

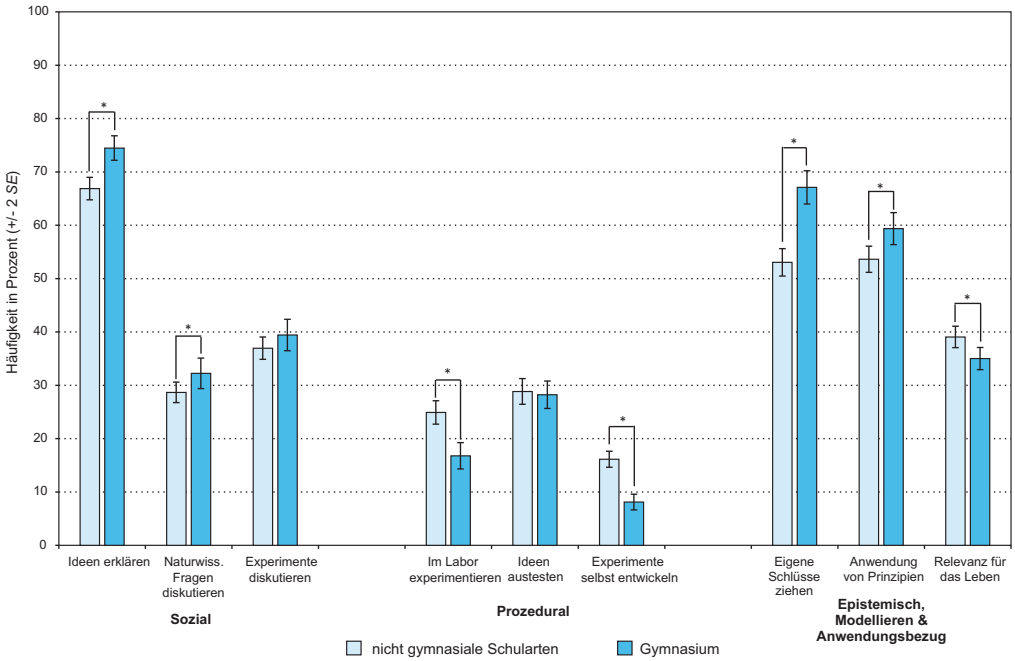


* Schularten unterscheiden sich signifikant ($p < .05$).

Abbildung 4.4: Domänenübergreifende Unterrichtsmerkmale am Gymnasium und an nicht gymnasialen Schularten

menten. Sie treten am Gymnasium und den nicht gymnasialen Schularten ähnlich häufig in allen oder den meisten Unterrichtsstunden auf. *Prozedurale Aktivitäten,* wie im Labor zu experimentieren und Experimente selbst zu entwickeln, werden hingegen von Schülerinnen und Schülern nicht gymnasialer Schularten häufiger berichtet. Hier scheint demnach das Experimentieren einen besonderen Stellenwert im Unterricht einzunehmen. Das Durchführen von Experimenten allein reicht jedoch für die Entwicklung naturwissenschaftlicher Kompetenz nicht aus. Schülerinnen und Schüler müssen darüber hinaus im Unterricht die Gelegenheit haben, auch kognitiv anspruchsvolle Tätigkeiten wie das Ziehen von Schlüssen aus eigenen Experimenten auszuführen. Diese *epistemischen Aktivitäten* werden von den Schülerinnen und Schülern am Gymnasium signifikant häufiger berichtet. Ebenso geben am Gymnasium die Fünfzehnjährigen häufiger an, dass in allen oder den meisten Unterrichtsstunden *naturwissenschaftliche Prinzipien* auf verschiedene Phänomene angewendet werden und Modellvorstellungen angesprochen werden. Der *Bezug zum alltäglichen Leben* wird hingegen von den Jugendlichen nicht gymnasialer Schularten häufiger gesehen. Hier scheinen die Lehrkräfte einen stärkeren Fokus darauf zu legen, zu erklären, dass naturwissenschaftliche Konzepte für das alltägliche Leben wichtig sind.

© Waxmann Verlag GmbH



* Schularten unterscheiden sich signifikant ($p < .05$).

Abbildung 4.5: Forschend-entdeckender Unterricht am Gymnasium und an nicht gymnasialen Schularten

4.3.3 Die Veränderung der Wahrnehmung des forschend-entdeckenden Unterrichts zwischen PISA 2006 und PISA 2015

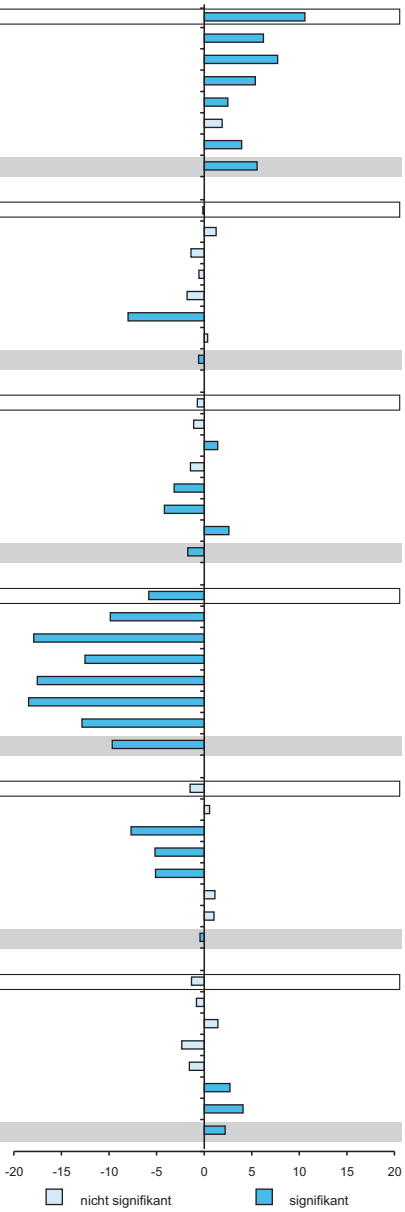
Mit den Naturwissenschaften als Hauptdomäne schließt PISA 2015 den zweiten vollständigen Erhebungszyklus ab. Bereits bei PISA 2006 wurde der naturwissenschaftliche Unterricht betrachtet und einige Fragen zu den Lernaktivitäten des forschend-entdeckenden Unterrichts kamen damals ebenfalls zum Einsatz. Für diese soll im Folgenden betrachtet werden, inwieweit sich die Wahrnehmung des naturwissenschaftlichen Unterrichts zwischen den beiden untersuchten Kohorten verändert hat (vgl. Abbildung 4.6).

Im Bereich *soziale Aktivitäten*, welcher Kommunikation als einen wichtigen Aspekt der Naturwissenschaften betont, berichten die Jugendlichen in PISA 2015 im OECD-Durchschnitt signifikant häufiger, dass sie ihre Ideen in allen oder den meisten Unterrichtsstunden erklären können. In Deutschland geben im Vergleich zu PISA 2006 sogar 11 Prozent Schülerinnen und Schüler mehr an, dass das Äußern eigener Meinungen und Ideen in allen oder den meisten Stunden vorkommt. Diese Verbesserung zeigt sich auch in den meisten der Vergleichsstaaten außer im Vereinigten Königreich, wo die Differenz nicht als statistisch bedeutsam abgesichert werden kann.

Für die *prozeduralen Aktivtäten*, das Experimentieren und Forschen im naturwissenschaftlichen Unterricht, lassen sich international nur geringe Veränderungen im Ver-

© Waxmann Verlag GmbH

		OECD-Staaten	2015 Zustimmung in %[1]	(SE)	2006 Zustimmung in %[1]	(SE)
Sozial	Ideen erklären	Deutschland	70.0	(0.8)	59.4	(0.9)
		Schweiz	71.5	(0.9)	65.3	(0.7)
		Finnland	71.6	(0.8)	63.9	(1.0)
		Niederlande	52.9	(0.9)	47.5	(1.0)
		Estland	70.3	(0.8)	67.8	(0.8)
		Vereinigtes Königreich	73.9	(0.8)	72.0	(0.8)
		Kanada	77.1	(0.6)	73.2	(0.7)
		OECD-Durchschnitt	68.5	(0.1)	63.0	(0.1)
Prozedural	Im Labor experimentieren	Deutschland	21.6	(0.8)	21.8	(0.8)
		Schweiz	24.5	(1.0)	23.3	(0.9)
		Finnland	20.8	(0.8)	22.2	(0.9)
		Niederlande	29.4	(0.8)	29.9	(1.0)
		Estland	9.0	(0.6)	10.8	(0.7)
		Vereinigtes Königreich	18.6	(0.8)	26.6	(1.1)
		Kanada	28.7	(0.6)	28.4	(0.8)
		OECD-Durchschnitt	20.9	(0.1)	21.5	(0.2)
	Experimente selbst entwickeln	Deutschland	12.8	(0.5)	13.6	(0.7)
		Schweiz	17.3	(0.7)	18.4	(0.7)
		Finnland	6.2	(0.4)	4.8	(0.4)
		Niederlande	11.3	(0.6)	12.8	(0.7)
		Estland	11.6	(0.6)	14.8	(0.7)
		Vereinigtes Königreich	9.4	(0.5)	13.6	(0.6)
		Kanada	21.0	(0.7)	18.4	(0.6)
		OECD-Durchschnitt	15.7	(0.1)	17.5	(0.1)
Epistemisch	Eigene Schlüsse ziehen	Deutschland	58.7	(1.0)	64.5	(1.1)
		Schweiz	50.9	(1.1)	60.8	(0.8)
		Finnland	36.7	(1.0)	54.6	(0.9)
		Niederlande	38.5	(0.8)	51.1	(0.8)
		Estland	29.7	(0.9)	47.2	(0.9)
		Vereinigtes Königreich	48.5	(0.8)	67.0	(0.9)
		Kanada	53.0	(0.4)	65.9	(0.6)
		OECD-Durchschnitt	41.5	(0.2)	51.2	(0.2)
Modellieren & Anwendungsbezug	Anwendung von Prinzipien	Deutschland	56.0	(1.0)	57.5	(1.0)
		Schweiz	65.7	(0.9)	65.1	(0.8)
		Finnland	53.7	(0.8)	61.4	(1.0)
		Niederlande	45.7	(1.0)	50.9	(0.8)
		Estland	57.8	(0.8)	63.0	(0.7)
		Vereinigtes Königreich	60.6	(0.8)	59.4	(0.7)
		Kanada	73.0	(0.6)	72.0	(0.6)
		OECD-Durchschnitt	59.1	(0.1)	59.5	(0.1)
	Relevanz für das Leben	Deutschland	37.5	(0.7)	38.8	(1.0)
		Schweiz	48.3	(1.1)	49.1	(0.7)
		Finnland	48.3	(0.9)	46.8	(0.8)
		Niederlande	39.2	(0.9)	41.6	(0.9)
		Estland	56.8	(0.9)	58.4	(0.9)
		Vereinigtes Königreich	48.0	(0.8)	45.3	(0.8)
		Kanada	61.9	(0.7)	57.9	(0.8)
		OECD-Durchschnitt	50.0	(0.1)	47.8	(0.1)



[1] Die Antwortkategorien „in allen Stunden" und „in den meisten Stunden" wurden zusammengefasst.

Abbildung 4.6: Veränderung des forschend-entdeckenden Unterrichts zwischen PISA 2006 und PISA 2015 im internationalen Vergleich

© Waxmann Verlag GmbH

gleich zu PISA 2006 feststellen. Am auffälligsten ist, dass im Vereinigten Königreich strukturierte Laborexperimente bei PISA 2015 seltener vorzukommen scheinen. Im naturwissenschaftlichen Unterricht in Deutschland zeigen sich gar keine signifikanten Veränderungen. Die Jugendlichen berichten bei PISA 2015 ähnlich häufig wie bei PISA 2006 über Möglichkeiten, im Labor zu experimentieren und Untersuchungen selbst zu entwickeln.

Epistemische Aktivitäten wie das Interpretieren und Schlussfolgern aus generierten Daten scheinen im OECD-Durchschnitt bei PISA 2015 eine deutlich geringere Rolle zu spielen als neun Jahre zuvor. In Deutschland sowie in allen ausgewählten Vergleichsstaaten geben die Fünfzehnjährigen inzwischen signifikant seltener an, dass sie im naturwissenschaftlichen Unterricht dazu aufgefordert werden, Schlüsse aus einem durchgeführten Experiment zu ziehen. Diese Abnahme ist besonders deutlich in Finnland, Estland und dem Vereinigten Königreich zu beobachten. Diese gleichmäßige Abnahme in allen ausgewählten OECD-Staaten könnte Ausdruck einer veränderten Wissenschaftsorientierung sein. Da epistemische Aktivitäten aber einen wichtigen Aspekt im Gesamtprozess der naturwissenschaftlichen Erkenntnisgewinnung darstellen und für den Erwerb naturwissenschaftlicher Kompetenz von zentraler Bedeutung sind (Furtak et al., 2012), sollte diese Entwicklung im Blick behalten werden.

Das *Anwenden naturwissenschaftlicher Prinzipien* auf Phänomene hat sich im OECD-Durchschnitt bei PISA 2015 im Vergleich zu 2006 aus Sicht der Schülerinnen und Schüler kaum verändert. In Deutschland, der Schweiz, dem Vereinigten Königreich und Kanada geben die Jugendlichen ähnlich häufig wie 2006 an, dass naturwissenschaftliches Modellieren in ihrem Unterricht vorkommt. In Finnland, den Niederlanden und Estland berichten bei PISA 2015 im Schnitt 6 Prozent der Jugendlichen weniger, dass die Lehrkraft erklärt, wie ein im Unterricht erlerntes naturwissenschaftliches Prinzip auf eine Reihe verschiedener Phänomene angewendet werden kann.

Die *Relevanz naturwissenschaftlicher Konzepte für das alltägliche Leben* scheint im OECD-Durchschnitt häufiger verdeutlicht zu werden als noch bei PISA 2006. In Deutschland wie in den meisten der ausgewählten Vergleichsstaaten zeigt sich allerdings keine Veränderung zwischen PISA 2006 und PISA 2015. Positiv fallen Kanada und das Vereinigte Königreich auf. Hier berichten mehr Jugendliche als noch 2006, dass die Lehrkraft ihnen erklärt, wie wichtig naturwissenschaftliche Konzepte für das eigene Leben sind. Diese Aktivität ist gerade für die Entwicklung motivational-affektiver Lernergebnisse wie Freude und Interesse an den Naturwissenschaften von zentraler Bedeutung.

Der naturwissenschaftliche Unterricht in Deutschland hat sich demnach seit PISA 2006 in Bezug auf die meisten forschend-entdeckenden Unterrichtsaktivitäten kaum verändert. Ein differenzierter Blick auf mögliche Schulartunterschiede offenbart ebenfalls keine differenziellen Veränderungen in den prozeduralen Aktivitäten oder dem Herstellen des Alltagsbezugs. Allerdings berichten Gymnasiastinnen und Gymnasiasten im Vergleich zu PISA 2006 weniger häufig, dass naturwissenschaftliches Modellieren in allen oder den meisten Unterrichtsstunden vorkommt. Positiv ist zu sehen, dass die Schü-

© Waxmann Verlag GmbH

lerinnen und Schüler in Deutschland häufiger ihre eigenen Ideen und Meinungen im Naturwissenschaftsunterricht äußern können. Diese Entwicklung zeigt sich sowohl in den gymnasialen als auch den nicht gymnasialen Schularten. Problematisch scheint, dass die Fünfzehnjährigen in Deutschland im Vergleich zu PISA 2006 seltener angeben, Schlüsse aus eigenen Untersuchungen zu ziehen. Dies zeigt sich ebenfalls in beiden Schularten (Gymnasium -11 Prozent, nicht gymnasiale Schularten -6 Prozent). Solche epistemischen Aktivitäten haben sich jedoch als besonders positiv für das Lernen erwiesen (Furtak et al., 2012). Gerade das Schlussfolgern hängt nicht nur mit einer hohen naturwissenschaftlichen Kompetenz zusammen (Lavonen & Laaksonen, 2009), sondern auch mit günstigen motivational-affektiven Lernergebnissen (Kjærnsli & Lie, 2011). Hier zeigt sich demnach Entwicklungspotenzial für den naturwissenschaftlichen Unterricht in Deutschland über die Schularten hinweg.

4.3.4 Muster des naturwissenschaftlichen Unterrichts in Deutschland

In den vorherigen Abschnitten wurden einzelne Unterrichtsmerkmale beschrieben, die sich in der Lehr-Lernforschung als wichtig für die Förderung kognitiver und motivational-affektiver Lernergebnisse erwiesen haben. Dies liefert wichtige Erkenntnisse zum Naturwissenschaftsunterricht in Deutschland im internationalen Vergleich sowie zu Schulartunterschieden und Veränderungen zwischen PISA 2006 und PISA 2015. Allerdings beschränken sich diese Analysen auf die Betrachtung einzelner Unterrichtsmerkmale. Die Unterrichtsforschung der letzten Jahrzehnte zeigt jedoch, dass der Erfolg wirksamen Unterrichts auf das Zusammenspiel verschiedener Unterrichtsmerkmale zurückzuführen ist (z. B. Aebli, 2011; Oser & Baeriswyl, 2001; Seidel & Reiss, 2014). Insbesondere im Zusammenhang mit mehrdimensionalen Lernergebnissen wie naturwissenschaftlicher Kompetenz und Freude an den Naturwissenschaften zeigt sich so, was „guten“ Naturwissenschaftsunterricht in Deutschland aktuell auszeichnet.

Im Folgenden werden daher klassifizierende Analysen beschrieben, die eine Identifikation von Unterrichtsmustern ermöglichen. Diese liefern eine anschauliche Darstellung der Unterrichtsrealität. Für diese Analysen wurden fünf Items der Skala zum forschend-entdeckenden Unterricht so ausgewählt, dass sie dessen unterschiedliche Lernaktivitäten (*sozial*: „Schülerinnen und Schüler bekommen Gelegenheit, ihre Ideen zu erklären.“; *prozedural*: „Die Schülerinnen und Schüler führen praktische Experimente im Labor durch.“ sowie „Die Schülerinnen und Schüler dürfen ihre eigenen Experimente entwickeln.“; *epistemisch*: „Die Schülerinnen und Schüler sollen Schlüsse aus einem Experiment ziehen, das sie durchgeführt haben.“) sowie *Anwendungsbezüge* auf die Lebenswelt der Schülerinnen und Schüler („Der Lehrer/die Lehrerin erklärt deutlich die Wichtigkeit von naturwissenschaftlichen Konzepten für unser Leben.“) abbilden. Auf Basis der ausgewählten Items wurden zunächst Muster des Naturwissenschaftsunterrichts in Deutschland identifiziert und die Häufigkeit des Vorkommens dieser Mus-

© Waxmann Verlag GmbH

ter analysiert. Anschließend wurden die Muster hinsichtlich des Zusammenhangs mit domänenübergreifenden Unterrichtsmerkmalen untersucht. Schließlich wurde analysiert, wie diese Unterrichtsmuster mit der naturwissenschaftlichen Kompetenz und dem Interesse an Naturwissenschaften in Deutschland am Gymnasium zusammenhängen.

Methodische Vorbemerkung – die Latent-Class-Analyse

Die Unterrichtsmuster wurden mithilfe einer Latent-Class-Analyse (LCA – Lazarsfeld, 1950; Formann, 1984) ermittelt. Die LCA klassifiziert die Schülerinnen und Schüler anhand ihrer Antworten (auf die fünf Items) nach einem probabilistischen Prinzip, sodass jede Schülerin, jeder Schüler mit unterschiedlichen Wahrscheinlichkeiten jeder latenten Klassen zugeordnet wird. Die Klassifikation erfolgt so, dass die Antworten von Schülerinnen und Schülern innerhalb einer latenten Klasse möglichst ähnlich und im Vergleich zwischen latenten Klassen möglichst unähnlich sind. Für jede Schülerin, jeden Schüler lässt sich anhand der maximalen Zuordnungswahrscheinlichkeit die Zugehörigkeit zu genau einer der latenten Klassen bestimmen. Die Anzahl der latenten Klassen wird durch die wiederholte Berechnung von LCA-Modellen (wobei eine unterschiedliche Anzahl latenter Klassen vorgegeben wird) und einem anschließenden (relativen) Vergleich der Modellpassungen, bestimmt. Als Kriterium beim Modellvergleich dient hier das Likelihood-basierte Fit-Kriterium BIC (Schwarz, 1978). Die Mittelwerte der Antworten (auf die fünf Items) lassen sich für jede latente Klasse als unterschiedliche Profile der Unterrichtswahrnehmung darstellen. Durchgeführt wurden die Analysen mit dem R-Paket poLCA (Linzer & Lewis, 2014) für die frei verfügbare Statistik-Software R (R Core Team, 2016). Der Datensatz für Deutschland in Bezug auf den Schülerfragebogen umfasst insgesamt n = 5 764 Schülerinnen und Schüler. Davon haben n = 1 138 Schülerinnen und Schüler keines der fünf Items beantwortet (sogenannte „*unit-non-responder*“). Diese Fälle wurden aus den Analysen ausgeschlossen, da von ihnen keine Aussagen über den naturwissenschaftlichen Unterricht getroffen wurden. Demnach wurden mit der LCA n = 4 626 Fälle klassifiziert. Der relative Modellvergleich indiziert ein Modell mit fünf latenten Klassen als am besten passend. Einen Überblick des relativen Modellvergleichs gibt Tabelle 4.4.

Die Reliabilität der Klassifikation nach der Fünf-Klassen-Lösung, operationalisiert als mittlere maximale Klassenzuordnungswahrscheinlichkeit, beträgt dabei r_p = 0.73 für die latente Klasse 1, r_p = 0.79 für die latente Klasse 2, r_p = 0.84 für die latente Klasse 3, r_p = 0.90 für die latente Klasse 4 sowie r_p = 0.69 für die latente Klasse 5. Die Mittelwerte der Schülerantworten auf die fünf Items für die beiden latenten Klassen 1 und 4 weisen dabei auf eine weitgehend ähnliche Wahrnehmung des Unterrichts hin. Aus diesem Grund wurden die Schülerinnen und Schüler dieser beiden latenten Klassen für die nachfolgenden Analysen, inhaltlichen Überlegungen folgend, in einer Klasse zusammengefasst. Diese so zusammengefasste Klasse weist das in Abbildung 4.7 dargestellte Profil „Typ 1“ der Unterrichtswahrnehmung auf.

© Waxmann Verlag GmbH

Tabelle 4.4: Relativer Modellvergleich für LCA-Modelle mit unterschiedlicher Anzahl latenter Klassen

Latente Klassen	*n*	Iterationen	Modellparameter	Likelihood	BIC
1	4253	1	15	-26612.7	53352.0
2	4253	140	31	-25384.7	51031.0
3	4253	371	47	-24915.2	50227.1
4	4253	416	63	-24654.8	49841.3
5	4253	1187	79	-24558.8	49784.4
6	4253	1409	95	-24527.6	49856.9

Unterrichtsmuster in Deutschland

Es konnten mithilfe der Latent-Class-Analyse vier Muster des naturwissenschaftlichen Unterrichts in Deutschland identifiziert werden (Abbildung 4.7).

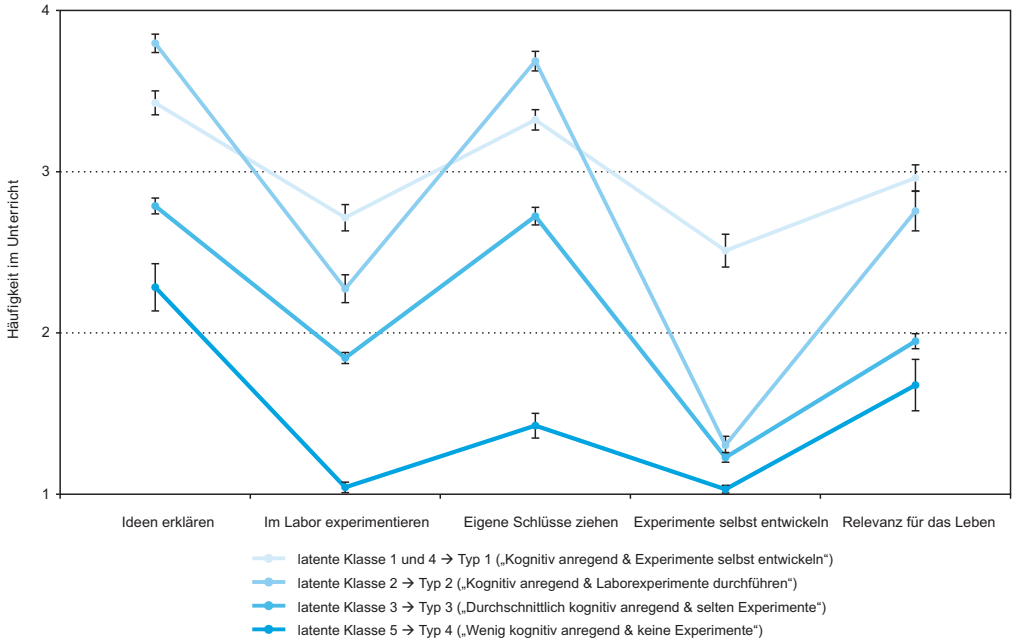


Abbildung 4.7: Profillinien der latenten Klassen für die Unterrichtswahrnehmung der Fünfzehnjährigen in Deutschland

Beim *ersten Unterrichtsmuster (Typ 1: „Kognitiv anregend & Experimente selbst entwickeln“)* berichten Schülerinnen und Schüler, dass alle Merkmale forschend-entdeckenden Lernens besonders häufig in jeder oder allen Unterrichtsstunden vorkommen. Sie haben die Gelegenheit, ihre eigenen Meinungen und Ideen zu äußern und dürfen regelmäßig praktische Experimente durchführen sowie eigene Experimente entwickeln. Darüber hinaus fordert die Lehrkraft die Jugendlichen auf, aus den Untersuchungen eigene Schlussfolgerungen zu ziehen. Schließlich stellt die Lehrperson in diesem Unterricht häufig die Relevanz und den Bezug zur Lebenswelt der Fünfzehnjährigen dar. Insgesamt erlaubt dieser kognitiv anregende, forschend-entdeckende Unterricht Schülerinnen und Schülern sowohl in Form von Experimenten praktisch tätig zu sein („*Hands-on*“) als auch sich mit eigenen Ideen und Begründungen („*Minds-on*“) am Unterricht zu beteiligen. Dabei berücksichtigt er die Lebenswelt der Schülerinnen und Schüler, indem die Relevanz naturwissenschaftlicher Konzepte für das Leben der Jugendlichen thematisiert wird.

Das *zweite Unterrichtsmuster* (*Typ 2: „Kognitiv anregend & Laborexperimente durchführen*“) beschreibt einen Naturwissenschaftsunterricht, der Schülerinnen und Schüler inhaltlich in fast allen Stunden einbezieht und kognitiv anregt, indem sie eigene Ideen äußern und Schlussfolgerungen aus Experimenten ziehen können („*Minds-on*“), sie jedoch seltener praktisch tätig sein lässt. Jugendliche mit Unterricht dieses Musters berichten, dass sie in einigen Stunden Experimente im Labor durchführen, jedoch fast nie oder in keiner Stunde eigene Experimente entwickeln („*Hands-on*“). Vergleichsweise häufig werden Bezüge zu ihrer Lebenswelt hergestellt und die Wichtigkeit naturwissenschaftlicher Konzepte für ihr Leben erläutert.

Das *dritte Unterrichtsmuster (Typ 3: „Durchschnittlich kognitiv anregend & selten Experimente“)* zeichnet sich dadurch aus, dass die Schülerinnen und Schüler in den meisten Stunden darüber berichten, dass sie eigene Ideen äußern können und ihre Begründungen und Schlussfolgerungen miteinbezogen werden. Seltener – nur in einigen Stunden – führen sie praktische Experimente im Labor durch, und eigene Experimente entwickeln sie in ihrer Wahrnehmung nie oder fast nie. Auch die Wichtigkeit naturwissenschaftlicher Konzepte für das Leben der Schülerinnen und Schüler wird nur in einigen Stunden erklärt. Während dieser Unterricht die kognitive und soziale Beteiligung von Schülerinnen und Schülern an inhaltlichen Überlegungen und Schlussfolgerungen zum Teil ermöglicht („*Minds-on*“), bietet er wenig Gelegenheiten für praktische Tätigkeiten („*Hands-on*“) und stellt kaum Bezüge zwischen naturwissenschaftlichen Inhalten und der Lebenswelt der Fünfzehnjährigen her.

Schülerinnen und Schüler, deren Antworten dem *vierten Unterrichtsmuster (Typ 4: „Wenig kognitiv anregend & keine Experimente“)* zuzuordnen sind, erleben die untersuchten Merkmale forschend-entdeckenden Lernens selten oder in keiner Unterrichtsstunde. Lediglich das Einbringen eigener Ideen berichten sie häufiger, jedoch im Vergleich zu den anderen Mustern ebenfalls in geringerem Ausmaß. Eigene Schlussfolgerungen aus Experimenten werden von der Lehrkraft nie oder fast nie gefordert. Gleichzeitig berichten die Schülerinnen und Schüler, in keiner Stunde die Möglichkeit zu haben, Experi-

© Waxmann Verlag GmbH

mente im Labor durchzuführen oder selbst zu entwickeln. Die Lehrerin oder der Lehrer erklärt zudem selten, warum naturwissenschaftliche Konzepte für das Leben der Schülerinnen und Schüler wichtig sind. Dieses Unterrichtsmuster beschreibt insgesamt einen wenig kognitiv aktivierenden Unterricht („*Minds-on*"), der fast keine Möglichkeiten für praktische Tätigkeiten („*Hands-on*") bietet. Dieser Unterricht scheint für das Erreichen kognitiver und motivational-affektiver Lernergebnisse besonders problematisch.

Tabelle 4.5: Prozentuale Häufigkeiten der Unterrichtsmuster insgesamt und getrennt für Gymnasien und andere Schularten

	Gesamt %	Gymnasium %	nicht gymnasiale Schularten %
Typ 1 („Kognitiv anregend & Experimente selbst entwickeln")	18.7	14.4	21.6
Typ 2 („Kognitiv anregend & Laborexperimente durchführen")	13.0	15.3	11.4
Typ 3 („Durchschnittlich kognitiv anregend & selten Experimente")	54.0	60.0	50.2
Typ 4 („Wenig kognitiv anregend & keine Experimente")	14.2	10.3	16.8

Diese vier Unterrichtsmuster sind im Naturwissenschaftsunterricht in Deutschland unterschiedlich häufig anzutreffen, und es zeigen sich Unterschiede zwischen dem Gymnasium und den nicht gymnasialen Schularten (vgl. Tabelle 4.5). Insgesamt berichtet etwas mehr als die Hälfte der Schülerinnen und Schüler, dass in ihrem Unterricht *Typ 3*, das heißt, *ein durchschnittlich kognitiv anregender Unterricht mit wenigen Experimenten* vorkommt. In diesem Naturwissenschaftsunterricht können sie zwar eigene Ideen und Schlussfolgerungen in den meisten oder einigen Stunden einbringen, sie haben jedoch wenige Gelegenheiten für das Durchführen und Entwickeln von Experimenten. Es zeigt sich, dass dieses Unterrichtsmuster am Gymnasium (60 Prozent) etwas häufiger anzutreffen ist als in den nicht gymnasialen Schularten (50 Prozent).

Das zweithäufigste Unterrichtsmuster, welches im Naturwissenschaftsunterricht in Deutschland vorkommt, ist *Typ 1* – ein *kognitiv anregender Unterricht mit selbst entwickelten Experimenten*. Rund ein Fünftel der Schülerinnen und Schüler berichtet über diesen auf der Basis der aktuellen Forschungslage günstigsten naturwissenschaftlichen Unterricht, der „*Minds-on*"- und „*Hands-on*"-Aktivitäten kombiniert. In diesem Unterricht können sich die Schülerinnen und Schüler sowohl häufig mit ihren Ideen und Begründungen am Unterricht beteiligen als auch regelmäßig Experimente durchführen und entwickeln. Es werden demnach soziale, prozedurale und epistemische Aktivitäten miteinander kombiniert – ein Muster, das positive Effekte auf multiple Lernziele erwarten lässt (Furtak et al., 2012). Dieses Unterrichtsmuster findet sich etwas seltener am Gymnasium und häufiger in nicht gymnasialen Schularten.

© Waxmann Verlag GmbH

Sowohl der *kognitiv anregende Unterricht, in dem vor allem Laborexperimente durchgeführt werden (Typ 2),* als auch der besonders problematische *wenig kognitiv anregende Unterricht ohne Experimente (Typ 4)* sind im Naturwissenschaftsunterricht in Deutschland weniger präsent und werden lediglich von 13 bzw. 14 Prozent der Schülerinnen und Schüler genannt. Betrachtet man diese Unterrichtsmuster getrennt nach Schularten, wird deutlich, dass die Gymnasiastinnen und Gymnasiasten etwas häufiger von einem *kognitiv anregenden Unterricht mit regelmäßigen Laborexperimenten (Typ 2)* berichten als Fünfzehnjährige nicht gymnasialer Schularten. Das problematischste Unterrichtsmuster (*Typ 4*), ein *wenig kognitiv anregender Unterricht ohne Experimente,* erleben Schülerinnen und Schüler nicht gymnasialer Schularten häufiger (17 Prozent) als Jugendliche am Gymnasium (10 Prozent).

Insgesamt überwiegen im Naturwissenschaftsunterricht in Deutschland mit 80 Prozent Unterrichtsmuster, die nur selten die Möglichkeit bieten, Experimente eigenständig zu planen und zu entwickeln. Mehrheitlich erwähnen die Schülerinnen und Schüler jedoch die Durchführung strukturierter Laborexperimente und die Gelegenheit, sich inhaltlich in Form von Ideen und Begründungen am Unterricht zu beteiligen. Naturwissenschaftsunterricht, der sich durch die Kombination von sozialen Aktivitäten wie der Erklärung von Ideen und prozeduralen Aktivitäten wie dem Durchführen und Entwickeln von Experimenten auszeichnet, findet sich deutlich seltener, jedoch häufiger als Unterricht, der weder kognitiv anregend gestaltet ist noch praktische Tätigkeit bietet. Der Schulartenvergleich legt nahe, dass Schülerinnen und Schüler an Gymnasien seltener forschend-entdeckende und häufiger kognitiv anregende Muster erleben. Die vorgenommene Typologie des Naturwissenschaftsunterrichts auf Basis von PISA 2015 ähnelt den Ergebnissen aus PISA 2006 (Kobarg et al., 2011). Dort wurden Unterrichtsmuster über die OECD-Staaten hinweg identifiziert und die Orchestrierung des Naturwissenschaftsunterrichts in Deutschland im internationalen Vergleich untersucht. Auch hier überwog ein Unterrichtsmuster mit wenigen Schülerexperimenten, aber starkem Fokus auf dem Ziehen von Schlussfolgerungen und Erklären von Ideen. In PISA 2006 zeigte sich, ähnlich wie in PISA 2015, in einem Zehntel der Fälle ein Unterricht, der sowohl praktische Tätigkeiten als auch das Schlussfolgern und Äußern von Ideen häufig ermöglicht. Wenngleich die Ergebnisse in PISA 2006 auf einer Identifikation von Unterrichtsmustern über Staaten hinweg beruhen, deutet der Vergleich darauf hin, dass der Naturwissenschaftsunterricht in Deutschland aktuell eine ähnliche Orchestrierung aufweist wie 2006.

© Waxmann Verlag GmbH

4.3.5 Zusammenhänge zwischen Unterrichtsmustern des forschend-entdeckenden Lernens und domänenübergreifenden Merkmalen der Unterrichtsqualität

Wie eingangs hergeleitet beeinflussen neben den fachspezifischen forschend-entdeckenden Lernaktivitäten auch domänenübergreifende Qualitätsmerkmale wie Unterstützung durch die Lehrkraft, individuelle Rückmeldungen, Differenzierung und Disziplin im Klassenzimmer die Effektivität von Unterricht (vgl. Abschnitt 4.1). In PISA 2015 wurden diese Merkmale der Unterrichtsqualität für die Naturwissenschaften erstmals zusätzlich zu den Lernaktivitäten des forschend-entdeckenden Unterrichts berücksichtigt. Dadurch kann untersucht werden, welche Aspekte der domänenübergreifenden Unterrichtsqualität in den einzelnen Mustern des Naturwissenschaftsunterrichts gegeben sind und inwieweit sich einzelne Unterrichtsmuster in Bezug auf die domänenübergreifenden Qualitätsmerkmale unterscheiden. Abbildung 4.8 gibt einen Überblick darüber, wie hoch die Unterstützung durch die Lehrkraft, die Rückmeldung, die Differenzierung sowie die Disziplin im Klassenzimmer im Mittel in den einzelnen Unterrichtsmustern ausgeprägt sind.

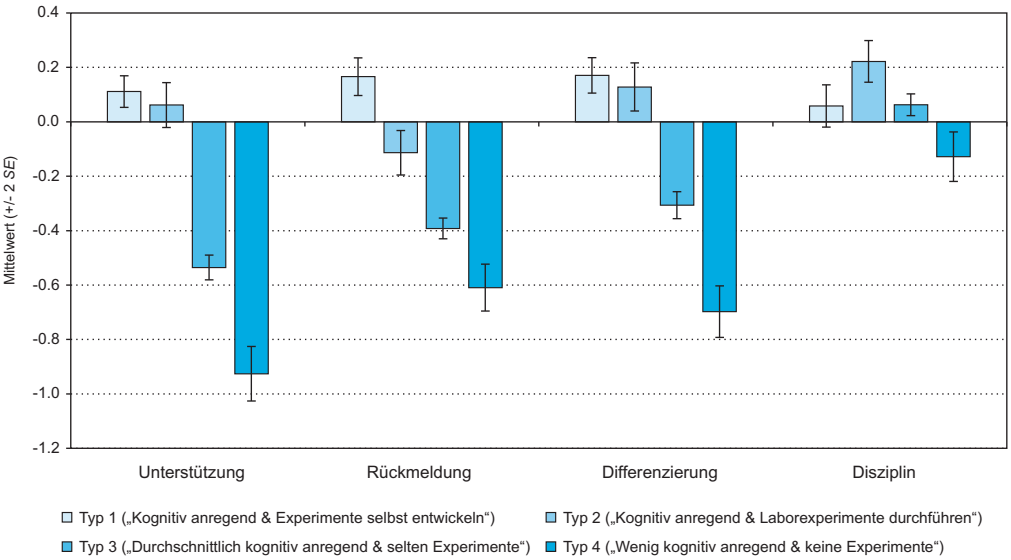


Abbildung 4.8: Vergleich der Unterrichtsmuster hinsichtlich domänenübergreifender Merkmale der Unterrichtsqualität

© Waxmann Verlag GmbH

In Bezug auf die *Unterstützung durch die Lehrkraft* geben Schülerinnen und Schüler des *kognitiv anregenden und Experimente selbst entwickelnden Unterrichts (Typ 1)* sowie des *kognitiv anregenden und Laborexperimente durchführenden Unterrichts (Typ 2)* die höchste Zustimmung. Der Unterschied in der wahrgenommenen Unterstützung zwischen diesen beiden Unterrichtsmustern ist nicht statistisch bedeutsam. Jeweils signifikant seltener nennen Jugendliche in den beiden anderen Unterrichtsmustern (*Typ 3 und 4*) Unterstützungsmaßnahmen. Der problematische *wenig kognitiv anregende und keine Experimente ermöglichende Unterricht* (*Typ 4*) geht zusätzlich mit einer als besonders gering wahrgenommenen Unterstützung durch die Lehrkraft einher. Diese Befunde zur wahrgenommenen Unterstützung in den Unterrichtsmustern finden sich über alle Schularten hinweg.

Eine regelmäßige *Rückmeldung* findet sich im *kognitiv anregenden und Experimente selbst entwickelnden Unterricht (Typ 1)*, während dieses Merkmal in allen anderen Unterrichtsmustern niedriger als im OECD-Durchschnitt ausgeprägt ist. Das bedeutet, dass Schülerinnen und Schüler in Deutschland vor allem in einem Unterricht, der sowohl das Einbringen eigener Ideen und Begründungen („*Minds-on*") als auch das Durchführen und Entwickeln von Experimenten („*Hands-on*") ermöglicht, Rückmeldungen über ihre individuellen Stärken und Verbesserungspotenziale erhalten. Während im *kognitiv anregenden und Laborexperimente durchführenden Unterricht (Typ 2)* noch häufig Rückmeldungen erfolgen, werden in den beiden weniger kognitiv anregenden Mustern (*Typ 3 und 4*) deutlich seltener Rückmeldungen berichtet. Diese Unterschiede treffen sowohl für Gymnasien als auch für andere Schularten zu. Insgesamt bekommen Gymnasiastinnen und Gymnasiasten ihrer Meinung nach jedoch in allen Mustern weniger Rückmeldung als Schülerinnen und Schüler anderer Schularten.

Eine *Differenzierung* des Unterrichts nach den individuellen Bedürfnissen der Schülerinnen und Schüler ist ähnlich wie die wahrgenommene Unterstützung vor allem im *kognitiv anregenden und Experimente selbst entwickelnden Unterricht (Typ 1)* und im *kognitiv anregenden und Laborexperimente durchführenden Unterricht (Typ 2)* zu finden. In diesen beiden Unterrichtsmustern unterscheidet sich die Differenzierung nicht signifikant. Im Vergleich dazu erleben Jugendliche im weniger kognitiv anregenden Unterricht mit seltenen oder keinen Gelegenheiten zum Experimentieren (*Typ 3 und 4*) deutlich weniger Anpassungen des Unterrichts an die Bedürfnisse und den Wissensstand der Klasse. Betrachtet man die Zusammenhänge getrennt nach Schularten, zeigt sich ein interessanter Unterschied. Während die Ergebnisse der nicht gymnasialen Schularten die der Gesamtstichprobe widerspiegeln, gibt es an den Gymnasien zusätzlich einen signifikanten Unterschied in der Differenzierung zwischen den beiden kognitiv anregenden Unterrichtsmustern. Schülerinnen und Schüler, die im Unterricht selbst Experimente entwickeln, erfahren eine höhere Differenzierung im Naturwissenschaftsunterricht als Fünfzehnjährige, die eher angeleitete Laborexperimente durchführen.

Im Vergleich zur Unterstützung durch die Lehrkraft, den wahrgenommenen individuellen Rückmeldungen und der Differenzierung zeigen sich bei der *Disziplin im Klassenzimmer* für alle Unterrichtsmuster deutlich höhere mittlere Skalenwerte. Anders als

© Waxmann Verlag GmbH

bei den anderen Qualitätsmerkmalen weist hier der *kognitiv anregende und Laborexperimente durchführende Unterricht (Typ 2)* die höchsten Ausprägungen hinsichtlich der Disziplin im Klassenzimmer auf und unterscheidet sich signifikant von allen anderen Mustern. Während im wünschenswertesten *kognitiv anregenden und Experimente selbst entwickelnden Unterricht (Typ 1)* und dem am häufigsten vorkommenden *durchschnittlich kognitiv anregenden Unterricht mit wenigen Experimenten (Typ 3)* ein ähnliches Ausmaß an Disziplin herrscht, kennzeichnet sich der problematische *wenig kognitiv anregende und keine Experimente ermöglichende Unterricht* (*Typ 4)* durch signifikant geringere Ausprägungen in der Skala Disziplin. Untersucht man die Zusammenhänge getrennt für die Schularten, so berichten Schülerinnen und Schüler am Gymnasium in allen Unterrichtsmustern eine höhere Disziplin als Schülerinnen und Schüler nicht gymnasialer Schularten. Allerdings unterscheidet sich am Gymnasium die Disziplin im *kognitiv anregenden und Experimente selbst entwickelnden Unterricht (Typ 1)* nicht von der im *kognitiv anregenden und Laborexperimente durchführenden Unterrichts (Typ 2)*. Beide zeichnen sich durch die höchste berichtete Disziplin im Klassenzimmer aus.

Zusammenfassend zeigt sich in Deutschland, dass insbesondere in einem *kognitiv anregenden und Experimente selbst entwickelnden Unterricht (Typ 1)* die häufigste individuelle Rückmeldung, häufige Unterstützung und Differenzierung sowie eine ausgeprägte Disziplin berichtet wird. Am Gymnasium weist der *kognitiv anregende und Experimente selbst entwickelnde Unterricht (Typ 1)* sogar in allen Qualitätsmerkmalen die höchsten mittleren Skalenwerte auf. Auch der *kognitiv anregende und Laborexperimente durchführende Unterricht (Typ 2)* zeichnet sich insgesamt durch ein hohes Maß an wahrgenommener Unterstützung und Differenzierung sowie die höchste Disziplin im naturwissenschaftlichen Unterricht aus. Rückmeldungen zu ihren individuellen Stärken und Schwächen werden in diesem Unterricht jedoch bereits deutlich seltener berichtet. Im Gegensatz zu diesen beiden Unterrichtsmustern zeigt das am häufigsten vorkommende Muster des *durchschnittlich kognitiv anregenden Unterrichts mit wenigen Experimenten (Typ 3)* zwar eine hohe Disziplin im Klassenzimmer, aber wenig Rückmeldung und Unterstützung. In allen Qualitätsmerkmalen niedrige Werte erreicht das Muster des *wenig kognitiv anregenden und keine Experimente ermöglichende Unterrichts* (*Typ 4).* Damit weisen vor allem die Muster, in denen Schülerinnen und Schüler häufig eigene Ideen erklären dürfen und ihre Begründungen in Schlussfolgerungen aus Experimenten miteinbezogen werden („*Minds on*“) hohe Werte in den domänenübergreifenden Qualitätsmerkmalen auf. Auch der häufigere Einsatz von praktischen Experimenten („*Hands on*“), die durchaus auch selbstständig entwickelt worden sein können, muss nicht mit einer geringeren allgemeinen Unterrichtsqualität einhergehen.

© Waxmann Verlag GmbH

4.3.6 Zusammenhänge zwischen Mustern des forschend-endeckenden Unterrichts und mehrdimensionalen Bildungszielen am Gymnasium

Nachdem die Unterrichtsmuster identifiziert und hinsichtlich domänenübergreifender Qualitätsmerkmale betrachtet wurden, wird nachfolgend untersucht, in welchem Zusammenhang die Muster des forschend-entdeckenden Unterrichts mit naturwissenschaftlicher Kompetenz sowie Freude und Interesse an den Naturwissenschaften stehen. Diese Analysen beschränken sich auf das Gymnasium, da diese Schulart in allen Bundesländern vertreten und die Schülerschaft in Bezug auf die Kompetenz relativ homogen zusammengesetzt ist. Für die nicht gymnasialen Schularten wären die Ergebnisse aufgrund der Überführung in neue oder neu bezeichnete Schularten in verschiedenen Bundesländern und der weniger homogen zusammengesetzten Schülerschaft schwieriger zu interpretieren.

Abbildung 4.9 zeigt die mittlere Ausprägung der naturwissenschaftlichen Kompetenz für die Muster des forschend-entdeckenden Unterrichts an Gymnasien. Schülerinnen und Schüler, die ihren Unterricht als *Typ 1* (M = 590 Punkte), *Typ 2* (M = 592 Punkte) und *Typ 3* (M = 596 Punkte) beschreiben, erzielen eine ähnlich ausgeprägte naturwissenschaftliche Kompetenz. Lediglich Jugendliche, die das Muster des *wenig kognitiv anregenden und keine Experimente ermöglichenden Unterrichts* (*Typ 4)* berichten, erreichen eine signifikant geringere naturwissenschaftliche Kompetenz (M = 566 Punkte) als Schülerinnen und Schüler mit anderen Unterrichtsmustern. Muster des forschend-entdeckenden Unterrichts, die demnach das Erklären eigener Ideen sowie das Schlussfolgern aus Experimenten ermöglichen („*Minds-on*“) und regelmäßige Experimente durchführen („*Hands on*“) – selbst entwickelt oder im Labor angeleitet – gehen am Gymnasium mit einer ähnlich hohen Kompetenz einher.

Deutlichere Unterschiede zwischen den Mustern des forschend-entdeckenden Unterrichts offenbaren sich bei der Betrachtung der Freude und Interesse an Naturwissenschaften. Ähnlich den Befunden des Zusammenhangs mit domänenübergreifenden Qualitätsmerkmalen gehen ein *kognitiv anregender und Experimente selbst entwickelnder Unterricht* (*Typ 1, M* = 0.50) sowie ein *kognitiv anregender und Laborexperimente durchführender Unterricht* (*Typ 2, M* = 0.39) am Gymnasium mit der höchsten Freude an Naturwissenschaften einher und unterscheiden sich nicht signifikant. Schülerinnen und Schüler mit dem am häufigsten vorkommenden *durchschnittlich kognitiv anregenden Unterricht mit wenigen Experimenten (Typ 3)* zeigen signifikant geringere Freude und Interesse an den Naturwissenschaften (M = 0.00). Im Vergleich dazu sind nur bei Jugendlichen mit problematischem *wenig kognitiv anregenden und keine Experimente ermöglichenden Unterricht* (*Typ 4)* noch geringere Ausprägungen hinsichtlich der Freude an Naturwissenschaften vorzufinden (M = -0.31).

Betrachtet man diese Ergebnisse insgesamt im Hinblick auf die Förderung mehrdimensionaler Bildungsziele, so ist festzuhalten, dass am Gymnasium ein kognitiv anregender Unterricht, der entweder das Entwickeln eigener Experimente *(Typ 1)* oder die

© Waxmann Verlag GmbH

Durchführung von Laborexperimenten ermöglicht *(Typ 2)*, sowohl mit hoher naturwissenschaftlicher Kompetenz als auch mit einer ausgeprägten Freude an den Naturwissenschaften einhergeht. Umgekehrt erreichen Schülerinnen und Schüler mit einem *wenig kognitiv anregenden und keine Experimente ermöglichenden Unterricht* (*Typ 4)* nur niedrige Kompetenz- und Interessenswerte. Der am häufigsten vorkommende *durchschnittlich kognitiv anregende Unterricht mit wenigen Experimenten (Typ 3)* geht zwar am Gymnasium mit einem eher durchschnittlich hohen Kompetenzniveau einher, jedoch gleichzeitig mit geringerer Freude und weniger ausgeprägtem Interesse an den Naturwissenschaften.

Bei der Interpretation dieser Zusammenhänge muss berücksichtigt werden, dass zwei Erklärungsrichtungen möglich sind: Zum einen ist denkbar, dass bestimmte Unterrichtsmuster nur dann umsetzbar sind, wenn Schülerinnen und Schüler hohe Kompetenzen und gegebenenfalls auch Freude und Interesse mitbringen. Zum anderen ist denkbar, dass Unterrichtsmuster bestimmte Wirkungen in Bezug auf Kompetenz und Lernfreude entfalten. Es bedarf daher weiterer Untersuchungen, um die tatsächliche Effektivität von Unterrichtsmustern zu prüfen.

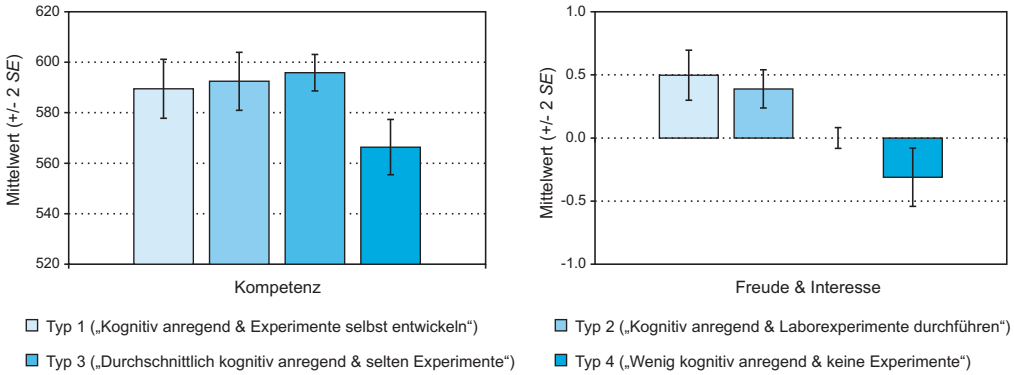


Abbildung 4.9: Vergleich der Unterrichtsmuster hinsichtlich naturwissenschaftlicher Kompetenz und Interesse

4.4 Zusammenfassung und Diskussion

In diesem Kapitel wurde der Frage nachgegangen, wie sich der Unterricht aus der Sicht der Fünfzehnjährigen in Deutschland im Vergleich zu ausgewählten OECD-Staaten charakterisieren lässt. Dabei gilt es, einige methodische Einschränkungen zu berücksichtigen, wie z. B., dass die Rekonstruktion des Unterrichts ausschließlich auf der individuellen Wahrnehmung von Schülerinnen und Schülern beruht, die sich auf unterschiedliche Schulklassen, Schularten und Schulen verteilen. Mithilfe der Schülerperspektive können zwar Aussagen darüber getroffen werden, wie sie ihren Unterricht wahrnehmen (Individualebene), streng genommen weiß man aber nicht, ob die Schülerinnen und Schüler einer Schulklasse in der Wahrnehmung ihres Unterrichts übereinstimmen und ob

© Waxmann Verlag GmbH

der Unterricht in den Klassenzimmern in Deutschland (Klassenebene) auch tatsächlich so durchgeführt wird. Trotz dieser auf das internationale Vorgehen zurückzuführenden Einschränkungen zeigen sich interessante Ergebnisse, die weitestgehend mit Befunden anderer Studien übereinstimmen, in denen objektivere videobasierte Verfahren eingesetzt wurden und die Zuordnung der Schülerinnen und Schüler zu ihren Klassen bekannt ist (Neumann et al., 2012; Roth et al., 2006; Seidel & Prenzel, 2006b).

Die Betrachtung der domänenübergreifenden Qualitätsmerkmale ergab, dass sich der naturwissenschaftliche Unterricht in Deutschland durch eine im internationalen Vergleich überdurchschnittliche Disziplin auszeichnet. Allerdings nehmen die Jugendlichen wenig individuelle Rückmeldungen zu ihren Stärken und Schwächen und Differenzierung im Hinblick auf den Umgang mit unterschiedlichen Lernvoraussetzungen und Bedürfnissen im Unterricht wahr. Dies spiegelt sich auch im Ausmaß der wahrgenommenen Unterstützung wider, die – ebenso wie in den Niederlanden – deutlich unterhalb des OECD-Durchschnitts liegt. Auch für individuelle Rückmeldungen berichteten die Fünfzehnjährigen im OECD-Vergleich in Deutschland, den Niederlanden, Finnland und der Schweiz über seltene Rückmeldungen zu Leistungsverbesserungen durch ihre Lehrkraft. Für Deutschland weisen die Befunde darauf hin, dass die Lehrpersonen zumindest im MINT-Bereich stärker dahingehend sensibilisiert werden könnten, im Unterricht den Fokus auf mehr individuelle Rückmeldungen, eine stärkere Differenzierung und häufigere Unterstützung zu legen. Diese Qualitätsmerkmale sind gerade für die Entwicklung motivational-affektiver Lernziele bedeutsam (z. B. Baumert et al., 2010; Decristan et al., 2015; Klieme & Rakoczy, 2003; Kunter & Voss, 2011; Schiepe-Tiska et al., 2016) und mit Blick auf die zunehmenden Anforderungen im Umgang mit Heterogenität im Klassenzimmer sinnvolle Ziele in der Aus- und Weiterbildung von Lehrkräften, um den Naturwissenschaftsunterricht in Deutschland konsequent weiterzuentwickeln. Dass ein solcher Unterricht nicht zulasten einer hohen Kompetenz gehen muss, zeigt Kanada, wo die Jugendlichen eine hoch ausgeprägte Unterstützung mit häufigen Rückmeldungen und einer Ausrichtung des Unterrichts an ihren Bedürfnissen erleben und gleichzeitig – neben positiv ausgeprägten motivational-affektiven Lernergebnissen – auch ein hohes Kompetenzniveau erreichen.

Die Ergebnisse zeigen, dass sich in Deutschland – zumindest im internationalen Vergleich – häufiger ein forschend-entdeckender Unterrichtsansatz finden lässt. Ein differenzierter Blick verdeutlicht, dass sozialen Lernaktivitäten wie dem Einbringen eigener Ideen und Diskutieren über naturwissenschaftliche Fragen oder Experimente (d. h. der Vorbereitung von Experimenten) substanzieller Raum gegeben wird. Dies deckt sich mit den Befunden von Börlin (2012), der auf etwa gleiche Zeitanteile im Unterricht für die Vorbereitung, Durchführung und Auswertung von Experimenten kommt. Mit Blick auf prozedurale Lernaktivitäten werden Experimente in geringerem Umfang von den Schülerinnen und Schülern selbst entwickelt. Auch dies passt zu den Befunden z. B. aus Videostudien, die zeigen, dass ein substanzieller Teil der Experimente im Naturwissenschaftsunterricht Demonstrationsexperimente der Lehrkraft sind (Neumann et al., 2012; Seidel & Prenzel, 2006b). Das Ziehen von Schlussfolgerungen als zen-

© Waxmann Verlag GmbH

trale epistemische Lernaktivität findet sich im internationalen Vergleich in Deutschland wieder häufiger. Das Muster forschend-entdeckenden Unterrichts in Deutschland ist also scheinbar durch eine gleichberechtigte Vorbereitung, Durchführung und Nachbereitung gekennzeichnet, wobei Schülerinnen und Schüler im Rahmen der Vorbereitung eigene Fragestellungen und Ideen einbringen dürfen, die Durchführung des Experiments selber jedoch durch die Lehrkraft vorbereitet und damit stark vorgegeben wird. In der Nachbereitung sind die Schülerinnen und Schüler jedoch wieder gefragt, sich mit eigenen Schlussfolgerungen einzubringen. Bei den durchgeführten Experimenten scheint es weniger darum zu gehen, naturwissenschaftliche Phänomene, die den Schülerinnen und Schülern aus dem Alltag bekannt sind, zu erklären bzw. Erklärungen zu selbigen zu erarbeiten. Stattdessen scheinen Experimente stärker der Erarbeitung fachlicher Inhalte zu dienen (vgl. Börlin, 2012).

Der Vergleich gymnasialer und nicht gymnasialer Schularten zeigt Unterschiede im forschend-entdeckenden Unterricht. So gibt es am Gymnasium einen deutlicheren Fokus auf die Diskussion naturwissenschaftlicher Fragestellungen und das Erklären von Ideen, gleichzeitig wird der experimentelle Prozess stärker auf der epistemologischen Ebene thematisiert – unter anderem werden Experimente häufiger zu naturwissenschaftlichen Prinzipien und Modellvorstellungen in Beziehung gesetzt. In nicht gymnasialen Schularten hingegen wird öfter der Bezug zum alltäglichen Leben hervorgehoben und insgesamt mehr experimentiert. Diese Befunde legen nahe, dass das Experiment an nicht gymnasialen Schularten als didaktisches Instrument zur Untersuchung und Erklärung von Phänomenen dient, während am Gymnasium die Theoriebildung im Vordergrund steht und das Experiment zur Prüfung theoretisch begründeter Hypothesen eingesetzt wird. Dafür spricht auch, dass am Gymnasium häufiger epistemische Aspekte des Experimentierens angesprochen werden. Diese Befunde lassen sich in Teilen sicher durch Unterschiede in der Lehramtsausbildung erklären. Sie ist für Lehrkräfte an Gymnasien stärker fach- und theoretisch orientiert, während Lehrkräfte für nicht gymnasiale Schularten oftmals eher praxis- und phänomenorientiert ausgebildet werden und der Fokus auf die Nutzung des Experiments zur Theoriebildung geringer ist.

Die für die Einzelmerkmale berichteten Ergebnisse spiegeln sich auch in den gefundenen Unterrichtsmustern des forschend-entdeckenden Unterrichts wider, die das Zusammenspiel der Lernaktivitäten darstellen. Ein guter Naturwissenschaftsunterricht in Deutschland zeichnet sich durch eine Kombination hoher kognitiver Anregung („*Minds-on*“) und regelmäßiger „*Hands-on*“-Aktivitäten, wie Experimente selbst entwickeln oder strukturierte Laborexperimente durchführen, aus (*Typ 1* und *Typ 2*). Ein solcher Unterricht, der soziale, prozedurale und epistemische Lernaktivitäten miteinander verknüpft und die Lebenswelt der Jugendlichen durch das Darstellen von Anwendungsbezügen berücksichtigt, geht am Gymnasium mit einer hohen naturwissenschaftlichen Kompetenz sowie Freude und Interesse an den Naturwissenschaften einher. In beiden Unterrichtsmustern scheint demnach eine gleichzeitige Förderung kognitiver und motivational-affektiver Lernergebnisse zu gelingen – wobei allerdings bei PISA offen bleibt, welche Erklärungsrichtung im Sinne von Ursache und Wirkung greift. Fast ein Drit-

© Waxmann Verlag GmbH

tel aller Schülerinnen und Schüler in Deutschland berichten einen solchen Unterricht. Unterschiede zeigen sich zwischen den Unterrichtsmustern in Bezug auf die domänenübergreifenden Qualitätsmerkmale. Während sich der Unterricht, in dem eher strukturierte Laborexperimente durchgeführt werden (*Typ 2*), durch eine höhere Disziplin auszeichnet, erleben die Jugendlichen in einem Unterricht, in dem sie selbst Experimente entwickeln und durchführen (*Typ 1*), mehr individuelle Rückmeldungen zu ihren Stärken und Schwächen.

Am häufigsten findet sich in Deutschland jedoch über alle Schularten hinweg ein Naturwissenschaftsunterricht, der durchschnittlich kognitiv anregend gestaltet ist und selten „*Hands-on*"-Gelegenheiten zum Experimentieren gibt (*Typ 3*). Die Hälfte der Schülerinnen und Schüler nicht gymnasialer Schularten sowie 60 Prozent der Gymnasiastinnen und Gymnasiasten erfahren nach ihrer Einschätzung einen solchen Unterricht, der sich zwar durch eine vergleichsweise positive Disziplin, aber deutlich weniger Unterstützung, Rückmeldung und Differenzierung auszeichnet. Dieser geht am Gymnasium mit einer hohen naturwissenschaftlichen Kompetenz einher, aber mit weniger Freude und Interesse an Naturwissenschaften. Einen problematischen Naturwissenschaftsunterricht (*Typ 4*) – kaum kognitiv anregend, keine Experimente, deutlich unterdurchschnittlich ausgeprägte domänenübergreifende Qualitätsmerkmale, am Gymnasium niedrig ausgeprägte Kompetenz und kaum Freude und Interesse an den Naturwissenschaften – erfahren immerhin noch 14 Prozent der Jugendlichen.

Betrachtet man die Veränderung der Wahrnehmung des forschend-entdeckenden Unterrichts seit PISA 2006 in Deutschland, so zeigen sich für die meisten Lernaktivitäten hohe Stabilitäten. Positiv entwickelt hat sich sowohl in Deutschland als auch in den meisten Vergleichsstaaten, dass Fünfzehnjährige häufiger ihre eigenen Meinungen und Ideen im naturwissenschaftlichen Unterricht äußern können. Schlüsse aus eigenen Experimenten zu ziehen wird hingegen in allen Vergleichsstaaten und in Deutschland – sowohl am Gymnasium als auch in nicht gymnasialen Schularten – weniger als vor neun Jahren im Unterricht verlangt. Hier zeichnet sich ein klarer Handlungsbedarf ab, denn gerade das Schlussfolgern als eine zentrale epistemische Aktivität hängt sowohl mit einer hohen naturwissenschaftlichen Kompetenz (Lavonen & Laaksonen, 2009) als auch günstigen motivational-affektiven Lernergebnissen (Kjærnsli & Lie, 2011) zusammen. Die Muster des Naturwissenschaftsunterrichts in Deutschland bei PISA 2015 ähneln denen, die bei PISA 2006 identifiziert wurden (Kobarg et al., 2011). Auch hier überwog in Deutschland ein Unterrichtsmuster mit nur wenigen Schülerexperimenten und einem Fokus auf dem Ziehen von Schlussfolgerungen und Erklären von Ideen. In der Zusammenschau mit den Ergebnissen zu den Einzelmerkmalen scheint sich demnach der naturwissenschaftliche Unterricht in Deutschland kaum verändert zu haben. Betrachtet man aus dieser Perspektive die im Vergleich zu PISA 2006 ebenfalls nicht veränderte naturwissenschaftliche Kompetenz (vgl. Kapitel 2), überrascht dieser Befund nicht. Die im Vergleich zu PISA 2006 geringer ausgeprägten motivational-affektiven Lernergebnisse (vgl. Kapitel 3) weisen darauf hin, dass mehr „*Hand on*"-Lernaktivitäten, die das eigenständige Arbeiten der Schülerinnen und Schüler ermöglichen und gleichzeitig struktu-

© Waxmann Verlag GmbH

riert durch die Lehrkraft angeleitet werden, im naturwissenschaftlichen Unterricht in Deutschland wünschenswert wären. Ein solcher Unterricht, der mehrdimensionale Bildungsziele verfolgt, wird dem Anspruch einer tragfähigen Grundbildung für alle Schülerinnen und Schüler gerecht und fördert neben fachlichen Kompetenzen auch die Entwicklung von Interessen und Lernmotivation.

Literatur

Abd-El-Khalick, F., BouJaoude, S., Duschl, R., Lederman, N., Mamlok-Naaman, R., Hofstein, A., Niaz, M., Treagust, D. & Tuan, H.-I. (2004). Inquiry in science education. International perspectives. *Science Education, 88*, 397–419. https://doi.org/10.1002/sce.10118

Aebli, H. (2011). *Zwölf Grundformen des Lehrens. Eine allgemeine Didaktik auf psychologischer Grundlage. Medien und Inhalte didaktischer Kommunikation, der Lernzyklus,* (14. Aufl.). Stuttgart: Klett-Cotta.

Aktionsrat Bildung. (2015). *Bildung. Mehr als Fachlichkeit*. Münster: Waxmann.

Alfieri, L., Brooks, P. J., Aldrich, N. J. & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*, 1–18. https://doi.org/10.1037/a0021017

Andersson-Bakken, E. & Klette, K. (2016). Teachers' use of questions and responses to students' contributions during whole class discussions. Comparing language arts and science classrooms. In K. Klette, O. K. Bergem & A. Roe (Hrsg.), *Teaching and learning in lower secondary schools in the era of PISA and TIMSS* (S. 63–84). Cham: Springer. https://doi.org/10.1007/978-3-319-17302-3_5

Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*, 133–180. https://doi.org/10.3102/0002831209345157

Bayer, S., Klieme, E. & Jude, N. (im Druck). Assessment and evaluation in educational contexts. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing context of learning world-wide – extended context assessment framework and documentation of questionnaire material*. New York: Springer.

Bayrhuber, H., Bögeholz, S., Elster, D., Hammann, M., Hößle, C., Lücken, M., Mayer, J., Nerdel, C., Neuhaus, B., Prechtl, H. & Sandmann, A. (2007). Biologie im Kontext – Ein Programm zur Kompetenzförderung durch Kontextorientierung im Biologieunterricht und zur Unterstützung von Lehrerprofessionalisierung. *Der mathematische und naturwissenschaftliche Unterricht, 60* (5), 282–286.

Berland, L. K. & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*, 765–793. https://doi.org/10.1002/sce.20402

Blanchard, M. R., Southerland, S. A., Osborne, J. W., Sampson, V. D., Annetta, L. A. & Granger, E. M. (2010). Is inquiry possible in light of accountability? A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Science Education*, *94*, 577–616. https://doi.org/10.1002/sce.20390

Börlin, J. (2012). *Das Experiment als Lerngelegenheit. Vom interkulturellen Vergleich des Physikunterrichts zu Merkmalen seiner Qualität*. Berlin: Logos-Verlag.

© Waxmann Verlag GmbH

Börlin, J. & Labudde, P. (2014). Practical work in physics instruction. An opportunity to learn?. In H. E. Fischer, P. Labbude, K. Neumann & J. Viiri (Hrsg.), *Quality of instruction in physics. Comparing Finland, Switzerland and Germany* (S. 111–128). Münster: Waxmann.

Cohen, D. K., Raudenbush, S. W. & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, *25*, 119–142. https://doi.org/10.3102/01623737025002119

Decristan, J., Klieme, E., Kunter, M., Hochweber, J., Büttner, G., Fauth, B., Hondrich, A. L., Rieser, S., Hertel, S. & Hardy, I. (2015). Embedded formative assessment and classroom process quality. How do they interact in promoting students' science understanding? *American Educational Research Journal, 52*, 1133–1159. https://doi.org/10.3102/0002831215596412

Demuth, R., Gräsel, C., Parchmann, I. & Ralle, B. (Hrsg.). (2008). *Chemie im Kontext: Von der Innovation zur nachhaltigen Verbreitung eines Unterrichtskonzepts*. Münster: Waxmann.

Donovan, M. S. & Bransford, J. D. (Hrsg.). (2005). *How students learn: History, mathematics, and science in the classroom*. Washington, DC: National Academy Press.

Fischer, H. E., Labudde, P., Neumann, K. & Viiri, J. (Hrsg.). (2014). *Quality of instruction in physics: Comparing Finland, Germany and Switzerland*. Münster: Waxmann.

Formann, A. K. (1984). *Die Latent-Class-Analyse. Einführung in Theorie und Anwendung*. Weinheim: Beltz.

Frey, A., Taskinen, P., Schütte, K., Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. & Pekrun, R. (Hrsg.). (2009). *PISA 2006 Skalenhandbuch. Dokumentation der Erhebungsinstrumente*. Münster: Waxmann.

Furtak, E. M., Seidel, T., Iverson, H. & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching. A meta-analysis. *Review of Educational Research*, *82*, 300–329. https://doi.org/10.3102/0034654312457206

Hazelkorn, E., Ryan, C., Beernaert, Y., Constantinou, C. P., Deca, L., Grangeat, M., Karikorpi, M., Lazoudis, A., Casulleras, R. P. & Welzel-Breuer, M. (2015). *Science education for responsible citizenship: Report to the European Commission of the expert group on science education. EUR: Vol. 26893*. Luxembourg: Publications Office. http://dx.doi.org/10.2777/12626

Henderson, J. B., MacPherson, A., Osborne, J. & Wild, A. (2015). Beyond construction. Five arguments for the role and value of critique in learning science. *International Journal of Science Education*, *37*, 1668–1697. https://doi.org/10.1080/09500693.2015.1043598

Hofstein, A. & Lunetta, V. N. (2004). The laboratory in science education. Foundations for the twenty-first century. *Science Education, 88*, 28–54. http://dx.doi.org/10.1002/sce.10106

Jiménez-Aleixandre, M. P. & Puig, B. (2012). Argumentation, evidence evaluation and critical thinking. In B. J. Fraser, K. Tobin & C. McRobbie (Hrsg.), *Second international handbook of science education* (S. 1001–1015). Dordrecht: Springer.

Johnson, T. P., Shavitt, S. & Holbrook, A. L. (2011). Survey response styles across cultures. In D. R. Matsumoto & F. J. R. d. van Vijver (Hrsg.), *Culture and psychology. Cross-cultural research methods in psychology* (S. 130–178). New York: Cambridge University Press.

Junge, C., von Arx, M. & Labudde, P. (2014). Classroom management. In H. E. Fischer, P. Labudde, K. Neumann & J. Viiri (Hrsg.), *Quality of Instruction in Physics. Comparing Finland, Germany and Switzerland* (S. 161–176). Münster: Waxmann.

Kirschner, P. A., Sweller, J. & Clark, R. E. (2006). Why minimal guidance during instruction does not work. An analysis of the failure of constructivist, discovery, problem-based,

© Waxmann Verlag GmbH

experiential, and inquiry-based teaching. *Educational Psychologist*, *41*, 75–86. https://doi.org/10.1207/s15326985ep4102_1

Kjærnsli, M. & Lie, S. (2011). Students' preference for science careers. International comparisons based on PISA 2006. *International Journal of Science Education, 33*, 121–144. https://doi.org/10.1080/09500693.2010.518642

Klahr, D. & Nigam, M. (2004). The equivalence of learning paths in early science instruction. Effects of direct instruction and discovery learning. *Psychological Science, 15*, 661–667. https://doi.org/10.1111/j.0956-7976.2004.00737.x

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). Heidelberg: Springer.

Klieme, E., Pauli, C. & Reusser, K. (2009). The Pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Hrsg.), *The power of video studies in investigating teaching and learning in the classroom* (S. 137–169). Münster: Waxmann.

Klieme, E. & Rakoczy, K. (2003). Unterrichtsqualität aus Schülerperspektive. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, U. Schiefele, W. Schneider, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 333–359). Opladen: Leske+Budrich.

Klieme, E., Schümer, G. & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I. Aufgabenkultur und Unterrichtsgestaltung. In Bundesministerium für Bildung und Forschung (BMBF) (Hrsg.), *TIMSS-Impulse für Schule und Unterricht. Forschungsbefunde, Reforminitiativen, Praxisberichte und Video-Dokumente* (S. 43–57). Bonn: BMBF Publik.

KMK (2005a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005a). *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. (2005b). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards 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. (2005c). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Physik für den mittleren Bildungsabschluss. Beschluss vom 16. Dezember 2004*. München: Wolters Kluwer.

Kobarg, M., Prenzel, M., Seidel, T., Walker, M., McCrae, B., Cresswell, J. & Wittwer, J. (2011). *An international comparison of science teaching and learning. Further results from PISA 2006*. Münster: Waxmann.

Kunter, M. & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environment Research*, *9*, 231–251. https://doi.org/10.1007/s10984-006-9015-7

Kunter, M. & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 115–132). Münster: Waxmann.

© Waxmann Verlag GmbH

Labudde, P. & Börlin, J. (2013). Inquiry-Based Learning: Versuch einer Einordnung zwischen Bildungsstandards, Forschungsfeldern und PROFILES. In S. Bernholt (Hrsg.), *Inquiry-based Learning – Forschendes Lernen, Gesellschaft für Didaktik der Chemie und Physik (GDCP)* (S. 183–185). Kiel: Institut für die Pädagogik der Naturwissenschaften und Mathematik.

Labudde, P. & Möller, K. (2015). Stichwort Naturwissenschaftlicher Unterricht. *Zeitschrift für Erziehungswissenschaft*, *15*, 11–36. https://doi.org/10.1007/s11618-012-0257-0

Lavonen, J. & Laaksonen, S. (2009). Context of teaching and learning school science in Finland. Reflections on PISA 2006 results. *Journal of Research in Science Teaching*, *46*, 922–944. https://doi.org/10.1002/tea.20339

Lazarsfeld, P. F. (1950). The logical and mathematical foundations of latent structure analysis. In S. A. Stouffler, L. Guttman, E. A. Suchman, P. F. Lazarsfeld, S. A. Star & J. A. Clausen (Hrsg.), *Studies in social psychology in World War II*. Princeton: Guilford Press.

Lazonder, A. W. & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, *86*, 681–718. https://doi.org/10.3102/0034654315627366

Linzer, D. & Lewis, J. (2014). *poLCA: Polytomous variable Latent Class Analysis* (Version 1.4.1). Zugriff am 03.11.2016. Verfügbar unter https://cran.r-project.org/web/packages/poLCA/index.html

Lüdtke, O., Robitzsch, A., Trautwein, U. & Kunter, M. (2009). Assessing the impact of learning environments: How to use student ratings of classroom or school characteristics in multilevel modeling. *Contemporary Educational Psychology, 34*, 120–131. https://doi.org/10.1016/j.cedpsych.2008.12.001_

Mikelskis-Seifert, S. & Duit, R. (2010). *Physik im Kontext – Konzepte, Ideen, Materialien für effizienten Physikunterricht*. Seelze: Friedrich Verlag.

Müller, K., Prenzel, M., Seidel, T., Schiepe-Tiska, A. & Kjærnsli, M. (im Druck). Science teaching and learning in schools: Theoretical and empirical foundations for investigating classroom-level processes. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing contexts of learning world-wide.* Berlin: Springer.

NRC (2012) = National Research Council (U.S.). (2012). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. Washington, D.C.: The National Academies Press.

Neumann, K., Kauertz, A. & Fischer, H. E. (2012). Quality of instruction in science education. In B. J. Fraser, K. Tobin & C. McRobbie (Hrsg.), *Second international handbook of science education* (S. 247–258). Amsterdam: Springer. https://doi.org/10.1007/978-1-4020-9041-7_18

Ødegaard, M. & Klette, K. (2012). Teaching activities and language use in science classrooms. In D. Jorde und J. Dillon (Hrsg.), *Science education research and practice in Europe*. (S. 181–202). Rotterdam: SensePublishers. https://doi.org/10.1007/978-94-6091-900-8_8

OECD. (2009). *PISA 2006 technical report*. Paris: OECD.

Osborne, J. (2010). Arguing to learn in science. The role of collaborative, critical discourse. *Science*, *328*, 463–466. https://dx.doi.org/10.1126/science.1183944

Osborne J. & Dillon J. (2008). *Science education in Europe: Critical reflections. A report to the Nuffield Foundation*. London: Nuffield Foundation.

Oser, F. K. & Baeriswyl, F. J. (2001). Choreographies of teaching. Bridging instruction to learning. In V. Richardson (Hrsg.), *Handbook of research on teaching. Fourth Edition* (S. 1031–1065). Washington, D.C.: American Educational Research Association.

© Waxmann Verlag GmbH

Pant, H. A., Stanat, P., Pöhlmann, C., Hecht, M., Jansen, M., Kampa, N., Lenski, A. E., Penk, C., Radmann, S., Roppelt, A., Schroeders, U., Siegle, T. & Ziemke, A. (2013). Der Blick in die Länder. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 159–247). Münster: Waxmann.

Parchmann, I., Gräsel, C., Baer, A., Nentwig, P., Demuth, R. & Ralle, B. (2006). "Chemie im Kontext": A symbiotic implementation of a context-based teaching and learning approach. *International Journal of Science Education*, *28*, 1041–1062. https://doi.org/10.1080/09500690600702512

Prenzel, M. & Lankes, E.-M. (2013). Was können Schülerinnen und Schüler über ihren Unterricht sagen? Ein Blick in die Schülerfragebogen von internationalen Vergleichsstudien. In N. McElvany & H. G. Holtappels (Hrsg.), *Empirische Bildungsforschung: Theorien, Methoden, Befunde und Perspektiven. Festschrift für Wilfried Bos* (S. 93–107). Münster: Waxmann.

R Core Team (2016). *R: A Language and Environment for Statistical Computing* (Version R version 3.3.1). Wien.

Roth, K. J., Druker, S. L., Garnier, H., Lemmens, M. Chen, C., Kawanaka, T., Rasmussen, D., Trubacova, S., Warvi, D., Okamoto, Y., Gonzales, P., Stigler, J. & Gallimore, R. (2006). *Teaching science in five countries: Results from the TIMSS 1999 video. Statistical analysis report. (NCES 2006-011).* Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Schiepe-Tiska, A., Heine, J.-H., Lüdtke, O., Seidel, T. & Prenzel, M. (2016). Mehrdimensionale Bildungsziele im Mathematikunterricht und ihr Zusammenhang mit den Basisdimensionen der Unterrichtsqualität. *Unterrichtswissenschaft*, *44* (3), 211–225.

Schiepe-Tiska, A., Roczen, N., Müller, K., Prenzel, M. & Osborne, J. (im Druck). Science-related outcomes: Attitudes, motivation, value beliefs, strategies. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.). *Assessing contexts of learning world-wide.* Berlin: Springer.

Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T.-Y. & Lee, Y.-H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching, 44*, 1436–1460. https://doi.org/10.1002/tea.20212

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464. https://doi.org/10.1214/aos/1176344136

Seidel, T. & Prenzel, M. (2006a). Teaching and learning of science. In ACER (Hrsg.), *PISA 2006 conceptual framework*, (S. 47–62). Camberwell: ACER. Zugriff am 03.11.2016. Verfügbar unter https://www.acer.edu.au/files/pisa2006_context_framework.pdf

Seidel, T. & Prenzel, M. (2006b). Stability of teaching patterns in physics instruction. Findings from a video study. *Learning and Instruction, 16*, S. 228–240. http://dx.doi.org/10.1016/j.learninstruc.2006.03.002

Seidel, T., Prenzel, M., Rimmele, R., Herweg, C., Kobarg, M., Schwindt, K. & Dalehefte, I. M. (2007). Science teaching and learning in German physics classrooms. Findings from the IPN video study. In M. Prenzel (Hrsg.), *Studies on the educational quality of schools. The final report on the DFG Priority Programme* (S. 79–99). Münster: Waxmann.

Seidel, T., Prenzel, M., Wittwer, J. & Schwindt, K. (2007). Unterricht in den Naturwissenschaften. 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. 147–179). Münster: Waxmann.

© Waxmann Verlag GmbH

Seidel, T. & Reiss, K. (2014). Lerngelegenheiten im Unterricht. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie* (S. 253–275). Weinheim: Beltz Psychologie Verlags Union.

Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade. The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, *77*, 454–499. http://dx.doi.org/10.3102/0034654307310317

Smith, G. & Matthews, P. (2000). Science, technology and society in transition year: A pilot study. *Irish Educational Studies, 19*, 107–119.

Taylor, J. A., Stuhlsatz, M. A. & Bybee, R. W. (2009). Windows into high-achieving science classrooms. In R. W. Bybee & B. McCrae (Hrsg.), *PISA Science 2006: Implications for science teachers and teaching* (S. 123–132). Arlington, VA: NSTA Press.

Tsai, C-C. (2000). The effects of STS-oriented instructions on female tenth graders' cognitive structure outcomes and the role of student scientific epistemological beliefs. *International Journal of Science Education*, *22*, 1099–1115. https://doi.org/10.1080/095006900429466

© Waxmann Verlag GmbH

© Waxmann Verlag GmbH

5 Schulische Rahmenbedingungen der Kompetenzentwicklung

Christine Sälzer, Manfred Prenzel, Anja Schiepe-Tiska & Marcus Hammann

Die schulischen Rahmenbedingungen in Deutschland weichen in vielen Aspekten nicht vom Durchschnitt der OECD-Staaten ab. Die Einschulung erfolgt mit sechs Jahren, die Klassen sind mit 25 Schülerinnen und Schülern durchschnittlich groß, die Zeitstunden, die für den naturwissenschaftlichen Unterricht zur Verfügung stehen, liegen mit 3.7 leicht über dem OECD-Durchschnitt von 3.5. Überdurchschnittlich häufig werden Instrumente zur Qualitätssicherung und -entwicklung eingesetzt, und Schulen in Deutschland engagieren sich stärker als in anderen OECD-Staaten dafür, auch freiwillig Evaluationen durchzuführen. Das subjektive Zugehörigkeitsgefühl der Schülerinnen und Schüler zu ihrer Schule ist überdurchschnittlich groß. Es ist allerdings bei Gymnasiastinnen und Gymnasiasten höher ausgeprägt als bei Fünfzehnjährigen an anderen Schularten. Zwischen den Schularten zeigen sich ansonsten Unterschiede etwa in der materiellen Ausstattung für den naturwissenschaftlichen Unterricht, die an Gymnasien besser bewertet wird als an nicht gymnasialen Schularten. Nach wie vor kritisch zu sehen ist die im Vergleich mit anderen OECD-Staaten hohe Zahl von Schülerinnen und Schülern, die mindestens eine Klasse wiederholt haben. In der Folge sind in Deutschland mehr Fünfzehnjährige als in anderen OECD-Staaten noch in Klasse 9. Ein weiteres auffälliges Ergebnis ist, dass etwas mehr als die Hälfte der Schülerinnen und Schüler in Deutschland angibt, weniger als zwei Stunden pro Woche für Hausaufgaben, Klausurvorbereitung oder zusätzlichen Unterricht in den Naturwissenschaften aufzuwenden. Im OECD-Durchschnitt ist es nur ein knappes Drittel, das ein so geringes Zeitbudget nennt.

Die fünfzehnjährigen Jugendlichen, die im Rahmen von PISA betrachtet werden, nähern sich in den meisten teilnehmenden Bildungssystemen dem Ende ihrer Pflichtschulzeit. Sie haben demnach durchschnittlich bereits etwa ein Jahrzehnt als Schülerinnen und Schüler erlebt und wurden von ihrem schulischen Umfeld geprägt. In Bezug auf die in PISA erfassten Domänen Naturwissenschaften, Mathematik und Lesen ist die Schule der Ort, an dem die entsprechenden fachlichen Kompetenzen vorwiegend aufgebaut und entwickelt werden. Für große Teile der mathematischen Kompetenz besitzt die Schule gewissermaßen das Vermittlungsmonopol (Baumert & Stanat, 2010), ähnlich verhält es sich mit der naturwissenschaftlichen Kompetenz. Die Lesekompetenz wird nicht nur im Deutschunterricht entwickelt, sondern in fast allen Fächern auf jeweils spezifische Weise

© Waxmann Verlag GmbH

gefördert und gefordert. Für das Lesen gilt auch, dass vielfältige Leseaktivitäten in der Freizeit zur Stärkung dieser Kompetenz beitragen (Cunningham & Stanovich, 1997). Dennoch werden wesentliche Aspekte der Lesekompetenz, insbesondere das systematische Lesen und Interpretieren von Texten, speziell im schulischen Kontext angeleitet und erworben.

Die Schule als gesellschaftliche Institution hat einen Bildungsauftrag zu erfüllen und ist an definierte rechtliche und administrative Rahmenbedingungen gebunden. Zugleich kann jede einzelne Schule als ein „Haus" oder Ort des Lernens und als Lebensraum mit bestimmten Merkmalen betrachtet werden (van Ackeren, Klemm & Kühn, 2015). Diese beiden Perspektiven auf die Schule als *Institution* und als *Einzelschule* weisen der Schule neben ihrer Qualifikationsfunktion mit Blick auf Anforderungen in Beruf, Privatleben und Gesellschaft auch eine Sozialisationsfunktion zu (Blömeke & Herzig, 2009; Hoffmann, 2008; Tippelt, 2013). PISA als internationale Bildungsvergleichsstudie ermöglicht entsprechend durch das Erfassen von Kompetenzen und Personenmerkmalen gegen Ende der Schulzeit auch, die Wirkungen von Schule hinsichtlich ihrer Qualifikations- und Sozialisationsfunktion einzuschätzen und Hinweise auf Unterschiede zwischen Schulen und gegebenenfalls zwischen nationalen Bildungssystemen zu geben.

International vergleichende Schulleistungsstudien dienen vor allem den Funktionen eines Monitorings und Benchmarkings (Seidel & Prenzel, 2008). Beides impliziert Vergleiche mit bestimmten Kriterien (*Monitoring*) beziehungsweise mit Strukturen, Prozessen und Ergebnissen anderer Bildungssysteme (*Benchmarking*). Während sich die Kapitel zur Beschreibung der Schülerkompetenzen (Kapitel 2, 6 und 7 in diesem Band) hauptsächlich dem Monitoring widmen, steht in diesem Kapitel mit den schulischen Rahmenbedingungen der Aspekt des Benchmarkings etwas stärker im Vordergrund. Hier ist es von besonderem Interesse, Merkmale des deutschen Schulsystems und deutscher Schulen zu identifizieren, die im Vergleich zu anderen Bildungssystemen auffallen und möglicherweise mit bestimmten Wirkungen im Hinblick auf den Lernerfolg und die Entwicklung der Schülerinnen und Schüler verbunden sind. Im Unterschied zur Auswertungsstrategie der OECD spielt dabei die statistische Aufklärung von Leistungsvarianz (die lediglich einen kleinen Teil der hier berichteten Ergebnisse ausmacht) durch Schulmerkmale nur eine relativ geringe Rolle (vgl. exemplarisch OECD, 2013; OECD, 2016a). Stärker im Vordergrund steht eine theoretisch geleitete Betrachtung von Indikatoren auf der Systemebene, die einen Bezug sowohl zu strukturellen Bedingungen als auch zu Einzelschulen in Deutschland haben. Diese werden mit Blick auf einige Besonderheiten des deutschen Schulsystems im Folgenden beschrieben und diskutiert.

Schulen als Institutionen unterliegen bestimmten organisatorischen, rechtlichen und administrativen Rahmenbedingungen. In Deutschland unterscheiden sich diese Rahmenbedingungen zum Teil auch zwischen den Ländern, und entsprechend unterschiedlich sind Gestaltungsmöglichkeiten und Entscheidungsspielräume der Schulen. Schulleitung, Fachkollegien und Lehrkräfte handeln im Rahmen von Vorgaben, die etwa die Schulstruktur oder die grundlegende Schulorganisation betreffen, aber auch Zielsetzun-

© Waxmann Verlag GmbH

gen, Lehrpläne, verbindliche Standards, Zuständigkeiten und Schulaufsicht sowie die Entscheidungsbefugnisse von Schulleitungen oder Mitwirkungsmöglichkeiten der Eltern.

Eingebettet in diese Rahmenbedingungen bringt die *Einzelschule* als Ort des Lernens Schülerinnen und Schüler in der Regel in Lerngruppen (meist Klassen) zusammen, die von Lehrkräften im Unterricht angeleitet und unterstützt werden. Die Lehrkräfte gestalten Lernumgebungen, schaffen Lerngelegenheiten und begleiten Lernprozesse (vgl. Helmke & Schrader, 2013). Daneben bieten zahlreiche Schulen auch außerunterrichtliche Lernangebote, welche von den Schülerinnen und Schülern für systematische Nacharbeit, interessengeleitete Vertiefung oder auch zur Freizeitgestaltung genutzt werden können. Insbesondere Ganztagsangebote an Schulen bieten zusätzliche Möglichkeiten für Freizeitaktivitäten und gemeinsame Tätigkeiten. Insofern umfasst jede einzelne Schule mehrere, unterschiedlich gestaltete Lernorte. Quasi unter einem Dach verbindet sie Gruppen von Schülerinnen und Schülern und ist somit ein sozialer und kultureller Lebensraum. Als organisatorisches Ganzes vereint eine Schule Lehrkräfte und Schülerschaft, stellt Ressourcen bereit, stimmt Unterricht über Jahrgangsstufen und Klassen hinweg ab, verpflichtet sich einem pädagogischen Programm, definiert Verantwortlichkeiten und gibt sich möglicherweise ein bestimmtes fächerorientiertes Profil (vgl. exemplarisch Hanßen, 2011).

Diese beiden Aspekte, die Schule als Institution und als Einzelschule, spielen in der Konzeption von PISA eine zentrale Rolle. Über die Beschreibung der Schülerleistungen hinaus verfolgt PISA das Ziel, Auskunft zu geben über Umgebungen, in denen die getesteten Kompetenzen entwickelt wurden. Bezogen auf diese Umgebungen werden Merkmale erhoben, die zwischen Staaten, aber auch innerhalb von Bildungssystemen deutlich variieren können.

Das PISA-Rahmenmodell

Merkmale, die im Rahmen von PISA zur Beschreibung struktureller Eigenschaften von Schulen und Bildungssystemen erfasst werden, können auf der Basis eines Modells ausgewählt werden. Dieses Modell wurde zuletzt im Framework für PISA 2015 (OECD, 2016b) dargestellt. Es unterscheidet zwischen Kontext-, Input-, Prozess- und Ergebnismerkmalen und setzt diese zueinander in Beziehung. Relevante *Kontextmerkmale* können sich auf vielfältige Aspekte beziehen: etwa die Differenzierung in der Sekundarstufe I, die Frage, zu welchen Zeitpunkten differenziert wird, welche Vorgaben zur Unterrichtszeit gemacht werden oder ob es verbindliche Regelungen gibt, die sich auf die Dauer des Schulbesuchs auswirken können (z. B. Klassenwiederholungen). Auf der *Inputebene* werden Merkmale beschrieben, die mit Lerngelegenheiten zusammenhängen und damit letztlich die Qualität der schulischen Bildungsprozesse oder deren Ergebnisse beeinflussen können. Dazu gehören vor allem Unterrichtsmerkmale (vgl. Kapitel 4 in diesem Band), aber beispielsweise auch die Ressourcen einer Schule oder die Zusammensetzung der Schülerschaft. Zur *Prozessebene* zählen unter anderem die Leistungsorientierung einer Schule, das Engagement der Lehrkräfte für Fortbildungen, das

© Waxmann Verlag GmbH

Kurs- oder Zusatzangebot oder auch das Schulklima und die Arbeit daran. Auf der *Ergebnisebene* relevant sind schließlich Indikatoren für Schulerfolg beziehungsweise Schulversagen oder aggregierte Wahrnehmungen des sozialen und emotionalen Befindens der Fünfzehnjährigen in ihrer Schule.

Die Auswahl der Indikatoren und Konstrukte, die zur Beschreibung diverser Aspekte von Bildungssystemen und Schulen im Rahmen von internationalen Vergleichsstudien herangezogen werden, beruht auf einer breiten interdisziplinären Theoriebasis. So gibt es sowohl erziehungswissenschaftliche als auch soziologische Arbeiten zur Funktion von Schule sowie zu institutionellen Bedingungen und sozialen Prozessen in Schulen, die auf Bedingungen institutioneller Bildungsprozesse hinweisen (z. B. Cicourel & Kitsuse, 2012; Dreeben & Barr, 1988; Fend, 2006; Meyer & Rowan, 2006; Neuenschwander, Gerber, Frank & Rottermann, 2012; Tippelt, 2013). Auf empirischer Ebene kann auf Forschung zur Schuleffektivität zurückgegriffen werden, welche die oben beschriebene grundlegende Unterscheidung von Kontext-, Input-, Prozess- und Ergebnisindikatoren nutzt (vgl. etwa Eickelmann, Gerrick & Koop, 2016; Klieme, 2013; Scheerens, 2004; Scheerens & Bosker, 1997). Studien zum Thema Schuleffektivität gehen so vor, dass Unterschiede in den Schülerkompetenzen auf Unterschiede in Merkmalen von Schulen zurückgeführt werden können. Wie hoch dieser Anteil erklärter Leistungs- bzw. Kompetenzvarianz ist, variiert zwischen den Studien, aber die Effektstärken fallen unter Kontrolle anderer Einflussfaktoren deutlich geringer aus als die Effektstärken für Unterrichtsmerkmale oder individuelle und familiäre Voraussetzungen der Schülerinnen und Schüler (vgl. etwa Hattie, 2008). Dies belegt, dass Lernprozesse viel unmittelbarer von Unterricht und seinen Merkmalen beeinflusst werden als von Charakteristika der Schule. Ausprägungen der Schule im Sinne einer „Schulkultur“ sind dennoch mit Werten und Einstellungen von Kindern und Jugendlichen verbunden, etwa mit ihrem Gefühl von sozialer Zugehörigkeit. Der Einfluss von Organisationsmerkmalen oder strukturellen Rahmenbedingungen ist eher über die Beobachtung von Veränderungen über mehrere Jahre hinweg zu erfassen, wie dies bereits in mehreren Ergänzungsstudien zu PISA untersucht wurde (Bischof, Hochweber, Hartig & Klieme, 2013; Klieme & Steinert 2008; Lehrl, Smidt, Grosse & Richter, 2013; Prenzel et al., 2006).

Da PISA international als Querschnittstudie angelegt ist und die teilnehmenden Schulen nur zu einem Zeitpunkt untersucht werden, müssen Gemeinsamkeiten oder Unterschiede zwischen Schulen sowohl auf nationaler als auch auf internationaler Ebene mit Vorsicht interpretiert werden. Kausalmodelle (Ursache-Wirkungs-Modelle) können auf der Basis der PISA-Daten nicht geprüft werden. Jedoch ermöglichen es die PISA-Daten teilweise, die Stichhaltigkeit von Hypothesen zu prüfen, die einen Zusammenhang zwischen verschiedenen Merkmalen von Bildungssystemen vermuten.

In der Berichterstattung zu PISA stehen die Ergebnisse zu den Schülerkompetenzen in den Naturwissenschaften, der Mathematik und im Lesen im Mittelpunkt. Durch Angaben über Unterschiede und Gemeinsamkeiten in Merkmalen von Schulen und Schulsystemen werden weitergehende Aussagen möglich. Die in PISA gewonnenen Daten geben so Auskunft über Rahmenbedingungen der Kompetenzentwicklung und

© Waxmann Verlag GmbH

erlauben es beispielsweise, der Frage nachzugehen, ob besonders leistungsstarke Staaten einige Schul- oder Strukturmerkmale gemeinsam haben. In diesem Sinne werden die PISA-Daten für ein *Benchmarking* genutzt. Vorsicht ist dabei jedoch geboten, wenn es um die Berücksichtigung möglicher Fehlerquellen für Interpretationen geht (z. B. Effekte nicht kontrollierter Faktoren wie des kulturellen Hintergrundes oder Wechselwirkungen von Merkmalen, vgl. dazu u. a. Vieluf, Kunter & van de Vijver, 2013). In Bezug auf die Funktion des *Monitorings* von Schulmerkmalen können PISA-Daten hilfreich bei der Aufdeckung von Problembereichen sein. In dreijährigem Abstand liefern die PISA-Erhebungen dazu Informationen über verzögerte Schullaufbahnen, die Häufigkeiten von Schülerabsenzen oder Veränderungen in Indikatoren für das Schulklima. Entsprechende Daten zeichnen gemeinsam mit den Ergebnissen der Kompetenztests ein mehrdimensionales Bild von Stärken oder Schwächen eines Bildungssystems.

Wie werden Schulmerkmale in PISA erhoben?

Schulische Rahmenbedingungen der Kompetenzentwicklung werden in PISA anhand von Daten aus mehreren Quellen erfasst. Neben einigen strukturellen Grunddaten über die Schulsysteme werden Fragebögen genutzt, um Informationen zu Schulmerkmalen und organisatorischen Rahmenbedingungen zu gewinnen. Der größte Teil der Informationen wird dem sogenannten *Schulfragebogen* entnommen, der sich an die Schulleitung richtet und Merkmale auf der Ebene der Einzelschule und ihrer institutionellen Gegebenheiten erfragt. Die Zuverlässigkeit der Auskünfte kann für Fragen, die konkrete Ereignisse oder Fakten an den Schulen betreffen, als hoch bewertet werden. Daneben werden Schulleiterinnen und Schulleiter im Schulfragebogen auch um Einschätzungen bestimmter Gegebenheiten an ihrer Schule gebeten, etwa Ressourcen oder Angebote. Auch der Schülerfragebogen liefert Einschätzungen der besuchten Schule, sodass verschiedene Perspektiven miteinander verknüpft werden können.

In PISA 2015 stand der in Deutschland bereits seit Langem eingesetzte Lehrerfragebogen erstmals auch als internationale Option zur Verfügung. Diese Option wurde von 18 Staaten genutzt, sodass ein internationaler Vergleich eingeschränkt möglich ist. Unter den OECD-Staaten haben sich die USA, Chile, Italien, die Tschechische Republik, Australien, Portugal, Spanien und Korea für den Einsatz des Lehrerfragebogens entschieden; darunter ist kein weiteres Land mit Deutsch als Unterrichtssprache. Für einen internationalen Vergleich im Sinne eines OECD-Benchmarkings eignet sich der Lehrerfragebogen in PISA 2015 daher nicht und kann ausschließlich mit Fokus auf einen Vergleich einzelner Staaten herangezogen werden. Für die hier im Mittelpunkt stehenden Vergleiche mit der Gruppe der OECD-Staaten steht ausschließlich der Fragebogen für die Schulleitungen als Datengrundlage zur Verfügung und damit die Perspektive der Schulleitungen selbst. Mit Bezug auf diese Perspektive lassen sich beispielsweise Profile des Führungshandelns nachzeichnen, die dann für weiterführende Analysen genutzt werden können. Im Vergleich der OECD-Staaten wird unter anderem ersichtlich, dass es große Unterschiede dahingehend gibt, wie Schulleitungen mit den Ergebnissen interner Evaluatio-

© Waxmann Verlag GmbH

nen umgehen. Konsequenzen solcher Ergebnisse können auf verschiedenen Ebenen greifen, etwa beim Lehrerkollegium, den Schülerinnen und Schülern, der Zusammenarbeit mit den Elternhäusern oder auch mit Blick auf die Umsetzung von Lehrplänen.

Die in diesem Kapitel berichteten Ergebnisse zu schulischen Rahmenbedingungen der Kompetenzentwicklung sollen einen ersten, weitgehend deskriptiven Überblick über Erkenntnisse aus PISA 2015 geben. Merkmale von Schulen und Bildungssystemen sind seit PISA 2000 regelmäßig berichtet worden, weil mit ihnen bestimmte Vorteile oder Schwierigkeiten für das Lernen und die persönliche Entwicklung von Schülerinnen und Schülern assoziiert werden. Hier schließt das Kapitel an und reiht sich ein in ein kontinuierliches internationales Bildungsmonitoring, das regelmäßig diese mutmaßlichen Erfolgsfaktoren oder Hemmnisse für Bildungsprozesse betrachtet.

In den folgenden Abschnitten befasst sich der erste Schritt mit der oben angesprochenen Aufteilung der Varianz der naturwissenschaftlichen Kompetenz in Komponenten *innerhalb* von Schulen und *zwischen* den Schulen. Diese Aufschlüsselung der Varianz kann Anhaltspunkte über Qualitätsunterschiede zwischen Schulen geben, aber auch zu unterschiedlichen Zusammensetzungen und Rekrutierungsstrategien von Schulen. Allerdings lässt es das internationale Design bei PISA nicht zu, die Varianz innerhalb der Schulen weiter in Varianzen auf Klassenebene zu zerlegen. Dies geht darauf zurück, dass innerhalb der Schulen eine altersbasierte Stichprobe fünfzehnjähriger Schülerinnen und Schüler gezogen wird, die unterschiedliche Klassen und Klassenstufen besuchen. Die durch PISA gewonnenen Daten reichen auch nicht aus, um spezifisch Schulentwicklungsprozesse auf der Ebene der Einzelschule zu begründen.

Anschließend werden in diesem Kapitel weitere Ergebnisse aus Analysen zu den beiden genannten Ebenen der schulischen Rahmenbedingungen für die Entwicklung von Kompetenzen dargestellt. Zunächst wird auf institutionelle und organisatorische Rahmenbedingungen eingegangen und dann im Anschluss auf Merkmale von Einzelschulen. Das Kapitel greift Indikatoren auf, die über mehrere PISA-Erhebungsrunden vorliegen und einige Besonderheiten des deutschen Schulsystems verdeutlichen. Diese betreffen zum Beispiel das Alter der Schülerinnen und Schüler bei der Differenzierung in verschiedene Schularten der Sekundarstufe, den Anteil der Klassenwiederholungen bis zum Alter von 15 Jahren oder die Anzahl der verfügbaren Schularten auf der Sekundarstufe (vgl. u. a. Baumert et al., 2001; Fuchs, 2003). In Bezug auf die Naturwissenschaften und die Schulebene gibt das Kapitel Auskunft über Ressourcen und gezielte Aktivitäten, um die Auseinandersetzung der Schülerinnen und Schüler mit naturwissenschaftlichen Themen zu fördern.

© Waxmann Verlag GmbH

5.1 Institutionelle und organisatorische Rahmenbedingungen

5.1.1 Leistungsvarianz und Gliederung des Sekundarschulwesens

Besonders aufschlussreich für den Vergleich unterschiedlicher Bildungssysteme ist die Betrachtung der Varianz *zwischen* den Schulen und der Varianz *innerhalb* der Schulen. Indem die Stichprobenziehung bei PISA in zwei grundlegenden Schritten erfolgt – zuerst die Ziehung der Schulen und innerhalb der gezogenen Schulen dann die teilnehmenden Schülerinnen und Schüler –, ist diese Zerlegung der Varianz in zwei Varianzanteile möglich. Interessant ist die Varianzaufteilung vor allem deshalb, weil sie angibt, ob beispielsweise die naturwissenschaftliche Kompetenz von Fünfzehnjährigen innerhalb der Schulen stark variiert oder ob die durchschnittliche naturwissenschaftliche Kompetenz der Jugendlichen auf verschiedenen Schulen in einem Bildungssystem besonders unterschiedlich ist. Varianz innerhalb einer Schule geht unter anderem auf Varianz zwischen Klassen zurück, die beispielsweise durch Unterricht in unterschiedlicher Qualität erklärt werden kann. Je höher der Varianzanteil *zwischen* Schulen ist, desto stärker unterscheiden sich die Schulen in ihrem durchschnittlichen Kompetenzniveau. Ein kleiner Varianzanteil *zwischen* Schulen bedeutet also, dass sich die Schulen in Bezug auf ihren Kompetenzmittelwert tendenziell ähnlich sind. Für Deutschland ist in diesem Zusammenhang zu beachten, dass auch die Aufteilung der Varianz in Teile innerhalb und zwischen Schulen keine Schlussfolgerungen auf der Ebene eines Bundesländervergleichs erlaubt; Gymnasien in einem Bundesland können auf der Basis der PISA-Daten nicht mit Gymnasien in einem anderen Bundesland verglichen werden (vgl. Kapitel 1 und 12). Die Aufschlüsselung der Varianzanteile kann für Deutschland insgesamt und auch im internationalen Vergleich jedoch in Verbindung mit weiteren Merkmalen wie Prozessen in Schule und Unterricht Aufschluss über Qualitätsunterschiede geben oder auch über unterschiedliche Zusammensetzungen von Schulen.

In allen bisherigen PISA-Erhebungsrunden wurde ersichtlich, dass der zwischen den Schulen liegende Anteil an der Leistungsvarianz von Bildungssystem zu Bildungssystem unterschiedlich groß ausfällt (OECD, 2013). In der zuletzt veröffentlichten Erhebungsrunde, PISA 2012, lagen 36 Prozent der Varianz bei Schülerkompetenzen zwischen den Schulen, und damit etwas höher als in früheren PISA-Runden (vgl. etwa OECD, 2009, dort wird bis einschließlich PISA 2009 ein Mittel von 30 Prozent Varianzanteil zwischen den Schulen angegeben). Bestätigt hat sich dabei immer wieder der Befund, dass sich beispielsweise Schulen in Finnland oder Schweden kaum voneinander unterscheiden (Varianzanteil zwischen Schulen kleiner als 20 Prozent), während dieser Anteil in früh differenzierenden Bildungssystemen wie Belgien, Deutschland oder Österreich bei über 50 Prozent lag (OECD, 2009).

Wie sich die Varianz der Schülerkompetenzen innerhalb von Bildungssystemen aufteilt und wie stark sich die Schulen hierbei voneinander unterscheiden, lässt sich anhand mehrerer Faktoren erklären. Ein naheliegender Grund für unterschiedliche Varianzaufteilungen ist die unterschiedliche Gliederung der Schulsysteme. Zwar differenzieren

© Waxmann Verlag GmbH

alle Staaten früher oder später nach Leistung und die Schülerinnen und Schüler können dann verschiedene Schularten der Sekundarstufe besuchen. Die in PISA untersuchte Gruppe der Fünfzehnjährigen wird allerdings in einem Teil der Teilnehmerstaaten noch gemeinsam in einer Schulart unterrichtet, in anderen Staaten verteilt sich die Kohorte bereits auf zwei oder mehr Schularten. In einem Schulsystem, das in dieser Altersgruppe nach Leistung differenziert (hat), ist die Varianz *zwischen* den Schulen zwangsläufig größer als in Schulsystemen, die für diese Alterskohorte (noch) nicht differenziert haben. Entsprechend ist für die Varianz innerhalb von Schulen in mehrgliedrigen Schulsystemen zu erwarten, dass sie geringer ist und sich demnach zwar die Schulen im mittleren Kompetenzniveau relativ deutlich voneinander unterscheiden, aber die Schülerinnen und Schüler innerhalb einer Schule leistungsmäßig recht homogen sind. Die beobachtbaren Unterschiede in der Varianzaufteilung sind also Konsequenzen institutioneller Rahmenbedingungen in den nationalen Schulsystemen. Dies bedeutet dann allerdings auch, dass das Verhältnis der beiden Varianzanteile *zwischen* und *innerhalb* der Schulen nicht einfach als Indikator für Chancengerechtigkeit zu verstehen ist. Vielmehr zeigt sich Chancengleichheit oder -gerechtigkeit daran, wie wenig Lernergebnisse und insbesondere die erworbenen Abschlüsse und Kompetenzen der Jugendlichen an ihre Herkunft gekoppelt sind (vgl. hierzu Kapitel 8 und 9). Auch über die Qualität von Schulen sagen die Varianzanteile für sich genommen noch nichts aus; hierzu müssen zusätzlich Unterrichtsprozesse (siehe Kapitel 4) sowie weitere pädagogische Prozesse auf der Ebene der Schule (siehe unten) berücksichtigt werden.

Die Daten aus PISA 2015 liefern mehrere Indikatoren, die für die Interpretation der Gesamtleistungsvarianz sowie deren Aufteilung innerhalb und zwischen Schulen relevant sind. Dazu zeigt Tabelle 5.1 zunächst für die einzelnen OECD-Staaten den Mittelwert für die naturwissenschaftliche Kompetenz, nach dem die Staaten auch sortiert sind. Rechts daneben ist als Indikator für die Gliederung der Bildungssysteme abzulesen, wie viele allgemeinbildende Schularten für Jugendliche im Alter von 15 Jahren angeboten werden und in welchem Alter die Ausdifferenzierung in mehrere Schularten üblicherweise erfolgt. Dabei handelt es sich also um organisatorische und institutionelle Rahmenbedingungen, innerhalb derer die Schulen eines Staates arbeiten müssen. Schließlich wird in der Tabelle für jeden Staat die Gesamtvarianz angegeben sowie die prozentuale Aufteilung der Varianzanteile *innerhalb der Schulen* und *zwischen den Schulen*.

Aus der Tabelle wird ersichtlich, dass die erste Differenzierung der Schulsysteme in den 35 OECD-Staaten im Alter von 10 bis 16 Jahren erfolgt. In 15 OECD-Staaten findet diese Leistungsdifferenzierung erst im Alter von 16 Jahren statt. 22 OECD-Staaten weisen ein gegliedertes Schulsystem auf, jedoch sortieren 11 dieser Staaten ihre Schülerinnen und Schüler nicht während der Pflichtschulzeit in unterschiedliche Schularten. Unter den 22 gegliederten Schulsystemen differenzieren zwei Staaten im Alter von 16 Jahren (Chile, Lettland) und acht im Alter von 15 Jahren (Frankreich, Griechenland, Irland, Israel, Japan, Korea, Mexiko und Portugal). Mit üblicherweise 10 Jahren findet die Differenzierung in unterschiedliche Schularten der Sekundarstufe am frühesten in Deutschland und Österreich statt. Ebenfalls relativ früh, nämlich im Alter von 11 oder

© Waxmann Verlag GmbH

Tabelle 5.1: Varianz der naturwissenschaftlichen Kompetenz im internationalen Vergleich

OECD-Staaten	Mittelwert naturwissenschaftlicher Kompetenz	Zahl der Schularten/ Programme für 15-Jährige	Alter der ersten Differenzierung	Varianz der Schülerkompetenz und prozentuale Aufteilung		
				Gesamtvarianz	Varianz innerhalb der Schulen	Varianz zwischen den Schulen
Japan	538	2	15	97	56	44
Estland	534	1	16	88	85	15
Finnland	531	1	16	103	90	10
Kanada	528	1	16	95	84	16
Korea	513	3	15	101	75	25
Neuseeland	513	1	16	121	82	18
Slowenien	513	3	14	101	52	48
Australien	510	1	16	117	78	22
Deutschland[1]	509	4	10	110	56	44
Niederlande	509	7	12	114	42	58
Vereinigtes Königreich	509	1	16	111	78	22
Schweiz	506	4	12	111	62	38
Irland	503	4	15	88	87	13
Belgien	502	4	12	112	55	45
Dänemark	502	1	16	91	86	14
Polen	501	1	16	92	85	15
Portugal	501	3	15	94	77	23
Norwegen	498	1	16	103	92	8
Vereinigte Staaten	496	1	16	109	81	19
Frankreich	495	3	15	116	50	50
Österreich	495	4	10	106	56	44
Schweden	493	1	16	117	83	17
Spanien	493	1	16	86	87	13
Tschechische Republik	493	6	11	101	56	44
Lettland	490	5	16	75	84	16
Luxemburg	483	4	13	112	67	33
Italien	481	4	14	93	57	43
Ungarn	477	3	11	104	46	54
Island	473	1	16	93	96	4
Israel	467	2	15	126	63	37
Slowakische Republik	461	5	11	109	56	44
Griechenland	455	2	15	94	65	35
Chile	447	3	16	83	62	38
Türkei	425	3	11	70	47	53
Mexiko	416	3	15	57	70	30
OECD-Durchschnitt	491	3	14	100	69	31

[1]In Berlin und Brandenburg erfolgt die erste Differenzierung im Alter von 12 Jahren.

© Waxmann Verlag GmbH

12 Jahren, differenzieren Belgien, die Niederlande, die Schweiz, die Slowakische Republik, die Tschechische Republik, die Türkei sowie Ungarn. Dieser Indikator der Differenzierung von Schülerinnen und Schülern auf der Basis ihrer Leistungen ist ein grobes Merkmal auf struktureller Ebene, das in den einzelnen Bildungssystemen ergänzt wird durch interne, kurz- oder längerfristige, fachspezifische oder fächerübergreifende Differenzierungsmaßnahmen wie flexible Einschulung, Gruppierung nach Fähigkeit, unterschiedliche Bildungsgänge oder Kurswahlsysteme. Die Anzahl der verfügbaren Schularten oder Programme und das Alter bei der ersten Differenzierung geben daher einen ersten Anhaltspunkt für institutionelle Rahmenbedingungen, der jedoch für sich genommen nur bedingt aussagekräftig ist. In Bezug auf die Leistungsvarianz der Schülerinnen und Schüler in den unterschiedlichen PISA-Teilnehmerstaaten ist dieser Indikator jedoch relevant, da er mit der Verteilung der Varianz zusammenhängt und bei der Interpretation zu berücksichtigen ist.

Vor diesem Hintergrund ist zunächst die Gesamtvarianz naturwissenschaftlicher Kompetenz von Interesse. Eine Reihe von Staaten weist eine Gesamtstreuung auf, die deutlich über dem OECD-Mittelwert liegt. In dieser Gruppe befinden sich einerseits Staaten, die relativ früh differenzieren, etwa die Schweiz (111), Deutschland (110), Österreich (106), die Niederlande (114) oder Belgien (112). Andererseits ist auch in Staaten, die ihre Schülerinnen und Schüler erst im Alter von 15 oder 16 Jahren differenzieren, eine beträchtliche Gesamtvarianz festzustellen. Dies gilt unter anderem für Schweden (117), Frankreich (116) oder Neuseeland (121). Eine breite Varianz der Leistungen hängt also ebenso wenig zwingend mit einem gegliederten Sekundarschulsystem zusammen wie das durchschnittliche Kompetenzniveau.

Betrachtet man hingegen die Aufteilung der Varianz in Bezug auf Unterschiede *innerhalb* und *zwischen* den Schulen, so findet man erwartungsgemäß in den OECD-Staaten mit mehreren Schularten der Sekundarstufe (drei und mehr) relativ viel Varianz *zwischen* den Schulen. Während dieser Varianzanteil im OECD-Durchschnitt 31 Prozent der Gesamtvarianz ausmacht, liegt er in mehrgliedrigen Systemen wie Deutschland (44 Prozent), Österreich (44 Prozent), der Schweiz (38 Prozent) oder Ungarn (54 Prozent) deutlich darüber. In Staaten mit einer einzigen Schulart der Sekundarstufe für Fünfzehnjährige liegt der Varianzanteil hingegen bei 4 bis 22 Prozent und damit immer unterhalb des OECD-Durchschnitts.

Die Angaben in Tabelle 5.1 sind besonders interessant, wenn man sie in Beziehung zu den jeweiligen Mittelwerten naturwissenschaftlicher Kompetenz setzt (vgl. Kapitel 2). Dies wird deutlich anhand der Sortierung der Staaten in absteigender Reihenfolge ihres durchschnittlichen Kompetenzniveaus. Ausgehend von der Gliederung des Bildungssystems lässt sich kein systematischer Zusammenhang mit dem Mittelwert naturwissenschaftlicher Kompetenz bei den Fünfzehnjährigen feststellen (beispielsweise über Korrelationen). Vielmehr zeigt sich, dass Staaten, die ebenso wie Deutschland drei oder mehr verschiedene Schularten der Sekundarstufe haben, durchaus unterschiedliche Mittelwerte erreichen. Signifikant über dem OECD-Durchschnitt liegt der Mittelwert in den Naturwissenschaften bei PISA 2015 wie in Deutschland auch in Korea, Slowenien, den

© Waxmann Verlag GmbH

Niederlanden, der Schweiz, Irland, Belgien sowie Portugal. In anderen Staaten mit mehrgliedrigem Bildungssystem erzielen die Schülerinnen und Schüler Mittelwerte der naturwissenschaftlichen Kompetenz, die signifikant *unter* dem OECD-Durchschnitt liegen. Zu dieser Gruppe gehören Luxemburg, Italien, Ungarn, die Slowakische Republik, Chile, die Türkei und Mexiko.

In Bezug auf das Alter der Differenzierung in unterschiedliche Schularten der Sekundarstufe lässt sich ebenfalls kein unmittelbarer Rückschluss auf das mittlere Kompetenzniveau von Schülerinnen und Schülern eines Staates ziehen. Zwar differenzieren alle OECD-Staaten, die im Bereich Naturwissenschaften besser abschneiden als Deutschland, ihre Schülerinnen und Schüler frühestens im Alter von 14 Jahren. Jedoch erfolgt in den ebenfalls leistungsstarken Staaten Niederlande, Schweiz und Belgien der Übertritt in die Sekundarstufe mit spätestens 12 Jahren relativ früh. Fünfzehnjährige können dort vier bis sieben unterschiedliche Schularten besuchen. Irland und Dänemark, die ebenfalls überdurchschnittliche Kompetenzen aufweisen, differenzieren mit 15 beziehungsweise 16 Jahren, wobei die PISA-Kohorte in Dänemark noch eine einzige Schulart besucht und in Irland bereits zwischen vier Schularten wählen kann. Die Schweiz und Luxemburg differenzieren ähnlich früh wie Deutschland und Österreich, dort geht jedoch ein geringerer Anteil an der Leistungsvarianz auf Unterschiede zwischen Schulen zurück. Das heißt, dass sich verschiedene Schulen in der Schweiz und in Luxemburg in Bezug auf die durchschnittliche naturwissenschaftliche Kompetenz weniger stark voneinander unterscheiden als in Deutschland oder Österreich.

Varianz innerhalb und zwischen Schulen in Deutschland

Die Sekundarschullandschaft in Deutschland lässt sich mittlerweile nicht mehr über die Unterscheidung von fünf Schularten abbilden, wie dies in den vergangenen PISA-Erhebungsrunden der Fall war, sondern adäquater als eine Unterteilung in Gymnasien und nicht gymnasiale Schularten. Unter den nicht gymnasialen Schularten sind alle diejenigen Schularten zusammengefasst, die entweder als Monobildungsgangschulen in Form von Haupt- oder Realschulen existieren oder aber als Integrierte Gesamtschulen beziehungsweise (kooperative) Schulen mit mehreren Bildungsgängen innerhalb einer Schule mehr als einen Zweig anbieten. Das mehrgliedrige Sekundarschulsystem in Deutschland bringt mit sich, dass der Anteil der Leistungsvarianz *zwischen* verschiedenen Schulen vergleichsweise hoch ist. In Deutschland teilt sich mit Blick auf alle Schulen die Varianz der naturwissenschaftlichen Kompetenz der Schülerinnen und Schüler in PISA 2015 jeweils in etwa zur Hälfte auf Unterschiede *innerhalb von Schulen* (56 Prozent) und *zwischen Schulen* (44 Prozent) auf. Im Vergleich zu PISA 2012 hat sich diese Aufteilung verschoben: Während sowohl in PISA 2012 als auch in PISA 2015 jeweils etwa die Hälfte der Varianz innerhalb und zwischen Schulen lag, so ist der etwas höhere Anteil in PISA 2015 nun *innerhalb von Schulen* zu finden (in PISA 2012 waren dies 47 Prozent) und der etwas niedrigere Anteil *zwischen Schulen* (in PISA 2012 waren es 53 Prozent). Die Leistungsvarianz zwischen Schulen setzt sich aus zwei Bestandteilen zusammen: Unterschie-

© Waxmann Verlag GmbH

den zwischen Schularten sowie Unterschieden zwischen Einzelschulen innerhalb der Schularten, gegebenenfalls auch zwischen Bildungsgängen innerhalb von Schulen mit mehreren Bildungsgängen (integriert oder als kooperative Form). Entsprechend sollte der Varianzanteil zwischen den Schulen innerhalb der einzelnen Schularten deutlich niedriger ausfallen als der Varianzanteil zwischen den Schulen für die Gesamtstichprobe. Abbildung 5.1 zeigt die Varianzanteile naturwissenschaftlicher Kompetenz zwischen den Schulen getrennt für Gymnasien und nicht gymnasiale Schularten in Deutschland.

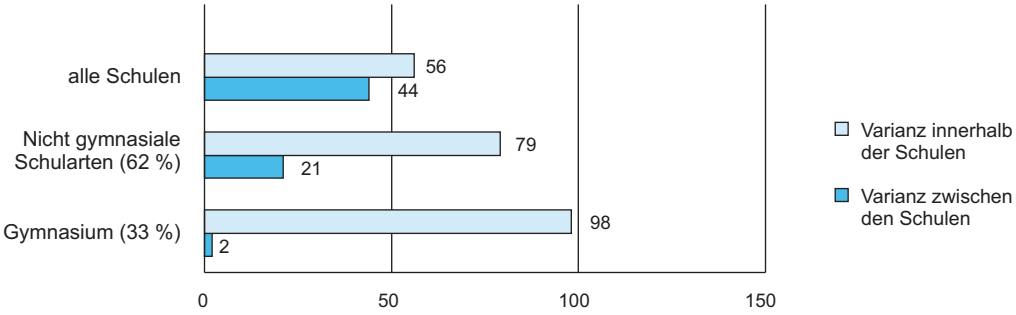


Abbildung 5.1: Varianzanteile der Schülerkompetenz (Naturwissenschaften) für alle Schulen sowie für Gymnasien und nicht gymnasiale Schularten (Bildungsbeteiligung in Klammern)

Erwartungsgemäß unterscheiden sich die Varianzanteile zwischen den Schulen in Abhängigkeit von der Schulart deutlich. Das Gymnasium als Schule mit nur einem Bildungsgang hat einen großen Varianzanteil *innerhalb der Schulen* (98 Prozent) und einen entsprechend geringen Varianzanteil (2 Prozent) *zwischen den Schulen*. Diese Schulart weist hinsichtlich der naturwissenschaftlichen Kompetenz in PISA 2015 eine sehr hohe Ähnlichkeit in der Leistungsverteilung von Schule zu Schule auf. Es sei angemerkt, dass sich die Sekundarschullandschaft in Deutschland in den letzten Jahren zunehmend zu einer Zweigliedrigkeit entwickelt hat: So existiert in acht Ländern neben dem Gymnasium und der Förderschule nur noch eine weitere Schulart, die zwei bis drei Bildungsgänge anbietet (Autorengruppe Bildungsberichterstattung, 2016). In einem Großteil der Länder wurden eigenständige Haupt- und Realschulen abgeschafft und durch kombinierte Schularten ersetzt, die mehrere Bildungsgänge und Abschlüsse anbieten. Erwartungsgemäß verteilt sich die Varianz bei den nicht gymnasialen Schularten auf 21 Prozent zwischen den Schulen und 79 Prozent innerhalb der Schulen. So ergibt sich ein Bild, dass für alle Schularten *zwischen* den Schulen in Deutschland eine deutlich geringere Leistungsvarianz zeigt als *innerhalb* der Schulen. Die zunehmende Zweigliedrigkeit der Sekundarschullandschaft in Deutschland wird in dieser vergleichsweise hohen Vari-

© Waxmann Verlag GmbH

anz innerhalb der Schulen sichtbar. Es liegt nahe, dass die nicht gymnasialen Schularten eine recht heterogene Schülerschaft zusammenführen, die zuvor auf unterschiedliche Schularten verteilt worden ist.

5.1.2 Alter und Klassenstufe der PISA-Kohorte

In bisherigen PISA-Erhebungsrunden wurde wiederholt deutlich, dass die fünfzehnjährigen Schülerinnen und Schüler in den verschiedenen Bildungssystemen bis zum Tag des PISA-Tests unterschiedliche Schulbiografien hinter sich haben (vgl. etwa Baumert et al., 2001). Insofern ist auch der Umgang mit der Lebenszeit der Fünfzehnjährigen ein Thema des internationalen Vergleichs. Die Verteilung der Fünfzehnjährigen auf verschiedene Klassenstufen wird bestimmt durch den Zeitpunkt der Einschulung, Wahlmöglichkeiten dieses Zeitpunkts oder auch Regelungen bezüglich verpflichtender oder freiwilliger Klassenwiederholung. Aufgrund der Möglichkeit der Klassenwiederholung oder einer vorgezogenen beziehungsweise verzögerten Einschulung sprechen wir im Folgenden von „Klassenstufen", um den Begriff vom Alter und Geburtsjahrgang der Jugendlichen zu entkoppeln.

Dieser Erhebungsansatz wirft bereits seit PISA 2000 immer wieder die Frage auf, ob die Tests in PISA eigentlich fair sein können, wenn doch die untersuchten Schülerinnen und Schüler in ihrem jeweiligen Bildungssystem unterschiedlich viel Unterrichtszeit erlebt haben. In Schulsystemen, die relativ früh einschulen oder auf Klassenwiederholungen verzichten, besuchen die Fünfzehnjährigen bereits höhere Klassenstufen und hätten so den Vorteil, bereits mehr Unterricht als andere erfahren zu haben und im Lehrstoff weiter fortgeschritten zu sein. Andererseits bieten flexible Einschulungszeitpunkte die Möglichkeit einer genaueren Passung von Unterrichtsangeboten zu Schülervoraussetzungen, sodass Lerngelegenheiten eventuell besser genutzt werden können. Ob und inwieweit diese beiden Überlegungen mit dem Kompetenzniveau oder der Leistungsvarianz der Schülerinnen und Schüler zusammenhängen, ist eine empirisch zu klärende Frage. Die Daten aus PISA ermöglichen zunächst einen Überblick über verschiedene Varianten des Umgangs mit Lebenszeit in unterschiedlichen Bildungssystemen.

Tabelle 5.2 zeigt, auf welche Klassenstufen die PISA-Kohorte der fünfzehnjährigen Schülerinnen und Schüler in den OECD-Staaten bei PISA 2015 verteilt ist. Angegeben wird hierbei die sogenannte Modalklasse, das heißt diejenige Klassenstufe, die im jeweiligen Staat von der Mehrheit der Fünfzehnjährigen besucht wird.

© Waxmann Verlag GmbH

Tabelle 5.2: Klassenstufe der PISA-Kohorte in den OECD-Staaten

		Prozentualer Anteil der Schülerinnen und Schüler					
	Modal-klasse	Klassenstufe unter der Modalklasse		Modalklasse		Klassenstufe über der Modalklasse	
OECD-Staaten		%	*(SE)*	%	*(SE)*	%	*(SE)*
Australien	10	11.2	(0.3)	74.6	(0.4)	14.0	(0.4)
Belgien	10	30.2	(0.7)	60.1	(0.9)	1.3	(0.1)
Chile	10	24.0	(0.7)	68.1	(1.0)	2.1	(0.2)
Dänemark	9	16.4	(0.6)	81.9	(0.7)	1.4	(0.5)
Deutschland	9	7.7	(0.4)	47.3	(0.8)	43.1	(0.8)
Estland	9	21.3	(0.6)	76.6	(0.6)	1.3	(0.3)
Finnland	9	13.6	(0.4)	85.7	(0.4)	0.0	(0.0)
Frankreich	10	23.1	(0.6)	72.5	(0.7)	3.2	(0.2)
Griechenland	10	3.8	(0.8)	95.3	(0.9)	0.0	(0.0)
Irland	9	1.8	(0.2)	60.6	(0.7)	26.5	(1.1)
Island	10	0.0	(0.0)	100.0	(0.0)	0.0	(0.0)
Israel	10	16.4	(0.9)	82.7	(0.9)	0.9	(0.3)
Italien	10	15.2	(0.6)	77.2	(0.7)	6.6	(0.3)
Japan	10	0.0	(0.0)	100.0	(0.0)	0.0	(0.0)
Kanada	10	10.8	(0.5)	87.5	(0.6)	0.8	(0.1)
Korea	10	9.1	(0.8)	90.4	(0.8)	0.5	(0.1)
Lettland	9	11.6	(0.5)	83.7	(0.7)	2.9	(0.3)
Luxemburg	9	7.9	(0.1)	50.9	(0.1)	40.3	(0.1)
Mexiko	10	31.9	(1.4)	60.3	(1.6)	0.5	(0.1)
Neuseeland	11	6.2	(0.3)	88.8	(0.5)	4.9	(0.5)
Niederlande	10	41.6	(0.6)	54.8	(0.6)	0.8	(0.2)
Norwegen	10	0.6	(0.1)	99.3	(0.2)	0.1	(0.1)
Österreich	10	20.8	(0.9)	71.2	(1.0)	5.9	(0.3)
Polen	9	4.9	(0.3)	93.8	(0.4)	0.6	(0.2)
Portugal	10	19.9	(0.8)	56.6	(1.3)	0.4	(0.1)
Schweden	9	3.1	(0.4)	94.9	(0.8)	1.8	(0.7)
Schweiz	9	11.8	(0.7)	60.9	(1.1)	25.7	(1.3)
Slowakische Republik	10	42.6	(1.3)	50.6	(1.2)	0.1	(0.0)
Slowenien	10	4.8	(0.3)	94.6	(0.4)	0.3	(0.1)
Spanien	10	23.4	(0.6)	67.9	(0.9)	0.1	(0.1)
Tschechische Republik	10	49.4	(1.2)	46.2	(1.2)	0.0	(0.0)
Türkei	10	20.7	(1.0)	72.9	(1.2)	3.0	(0.3)
Ungarn	9	8.5	(0.5)	75.8	(0.7)	14.0	(0.5)
Vereinigte Staaten	10	9.6	(0.7)	72.4	(0.9)	17.3	(0.6)
Vereinigtes Königreich	11	1.6	(0.3)	97.4	(0.4)	1.0	(0.3)
OECD-Durchschnitt	10	15.0	0.1	75.8	0.1	6.3	0.1

Modalklasse: Diejenige Klassenstufe, die in den Teilnehmerstaaten von der Mehrheit der Fünfzehnjährigen besucht wird.

© Waxmann Verlag GmbH

Wie Tabelle 5.2 verdeutlicht, ist die häufigste Modalklasse in den OECD-Staaten die 10. Klassenstufe. In den meisten OECD-Staaten (22) besuchen die fünfzehnjährigen Schülerinnen und Schüler also eine 10. Klasse. In weiteren 11 Staaten ist die 9. Klasse die am häufigsten besuchte Klasse, in zwei Staaten die 11. Klassenstufe (Neuseeland und Vereinigtes Königreich). In diesen Staaten befindet sich mit 89 Prozent (Neuseeland) beziehungsweise 97 Prozent (Vereinigtes Königreich) die PISA-Kohorte nahezu vollständig bereits in der 11. Klassenstufe. Dort beginnt das erste Pflichtschuljahr für Kinder im Alter von fünf Jahren. Insgesamt besuchen in den OECD-Staaten 76 Prozent der Fünfzehnjährigen die Modalklasse ihres Staates, 15 Prozent befinden sich eine Klassenstufe darunter und 6 Prozent eine Klassenstufe darüber.

Wie in den bisherigen PISA-Erhebungsrunden hat Deutschland hier einen vergleichsweise niedrigen Modalwert und eine breite Verteilung der Jugendlichen auf unterschiedliche Klassenstufen: 47 Prozent der Fünfzehnjährigen besuchten am PISA-Testtag im Frühjahr 2015 eine 9. Klasse, in der 10. Klassenstufe befanden sich 40 Prozent (etwa 1 Prozent besuchte eine höhere Klassenstufe als die zehnte), und 8 Prozent gehörten höchstens einer 8. Klassenstufe an. Allerdings wird ersichtlich, dass sich diese Verteilung während der letzten Jahre verändert hat und sich die Modalklasse auf lange Sicht hin zur 10. Klassenstufe verschieben wird: In PISA 2012 war der Anteil fünfzehnjähriger Schülerinnen und Schüler in der 9. Klassenstufe noch höher (52 Prozent; in PISA 2000: 61 Prozent), derjenige in der 10. Klassenstufe entsprechend geringer (PISA 2012: 38 Prozent; PISA 2000: 24 Prozent). Abbildung 5.2 zeigt, wie sich die prozentualen Anteile fünfzehnjähriger Schülerinnen und Schüler in den Klassenstufen 9 und 10 zwischen PISA 2000 und PISA 2015 verändert haben.

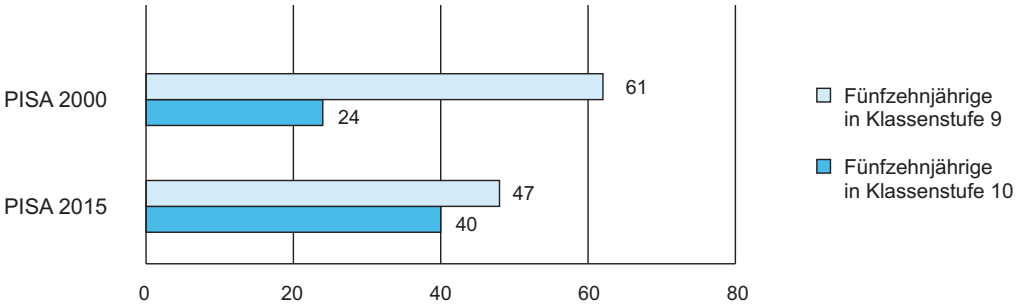


Abbildung 5.2: Prozentuale Anteile der Fünfzehnjährigen in Klassenstufe 9 und 10 (PISA 2000 und PISA 2015), inkl. berufliche Schulen und Förderschulen

© Waxmann Verlag GmbH

Dies spricht zunächst für eine Tendenz, dass sich der Anteil verzögerter Schullaufbahnen (etwa durch Klassenwiederholungen oder späte Einschulung) verringert hat. Der folgende Abschnitt geht vertiefend auf diesen Aspekt ein und zeigt, dass eine differenziertere Betrachtung des Alters bei der Einschulung sowie möglicher Klassenwiederholungen für die Beschreibung der Nutzung von Lebenszeit für schulische Bildung in Deutschland aufschlussreich ist.

Alter bei der Einschulung und Klassenwiederholung

Das Alter von Kindern bei ihrem Start in ihre Schullaufbahn variiert zwischen den OECD-Staaten beträchtlich (vgl. OECD, 2016c). Im OECD-Durchschnitt werden Kinder mit 6 Jahren eingeschult, für 46 Prozent der an PISA 2015 teilnehmenden Schülerinnen und Schüler trifft dies zu. Für jeweils etwa ein Viertel (26 bzw. 24 Prozent) der PISA-Kohorte begann die Pflichtschulzeit erst im Alter von 7 Jahren oder bereits vor dem 6. Geburtstag. In Irland und Portugal beginnt laut Auskunft der Schülerinnen und Schüler die Grundschule für die meisten Kinder bereits im Alter von 4 Jahren und in mehreren Staaten erst mit 7 Jahren (Estland, Finnland, Lettland, Polen, Schweden, Türkei, Ungarn). In einigen Staaten besteht die Möglichkeit einer flexiblen Einschulung, sodass die Grundschule im Alter von fünf Jahren oder auch früher anfangen kann. In Portugal können sogar bereits dreijährige Kinder eingeschult werden (dies traf auf 25 Prozent zu). Auch in Deutschland ist es seit einiger Zeit möglich, vor dem üblichen Alter von 6 Jahren eingeschult zu werden, während die Möglichkeiten einer verzögerten Einschulung weiter eingeschränkt wurden. Nach wie vor liegt das Durchschnittsalter der PISA-Kohorte 2015 bei Schulbeginn in Deutschland bei 6 Jahren und entspricht damit dem OECD-Durchschnitt. Knapp 69 Prozent der Fünfzehnjährigen waren im Alter von 6 Jahren in die Schule gekommen, 9 Prozent vor dem 6. Geburtstag und 23 Prozent im Alter von mindestens 7 Jahren. Auch wenn sich hier bereits andeutet, dass ein nicht unerheblicher Teil der Schulanfängerinnen und Schulanfänger in Deutschland schon zu Beginn der Schulzeit älter oder jünger als die durchschnittlichen 6 Jahre ist, so scheint die Verteilung der PISA-Kohorte auf unterschiedliche Klassenstufen noch deutlich breiter als im OECD-Durchschnitt zu sein. Im Vergleich zur ersten PISA-Erhebungsrunde im Jahr 2000 ist in PISA 2015 ferner der Anteil der verspätet eingeschulten Jugendlichen größer: Dieser betrug in PISA 2000 lediglich 12 Prozent (Stanat et al., 2002). Im Zusammenhang mit der Feststellung, dass fünfzehnjährige Schülerinnen und Schüler in Deutschland eine vergleichsweise niedrige Klassenstufe besuchen, wurden im Anschluss an PISA 2000 verschiedentlich Maßnahmen ergriffen, Verzögerungen bei der Einschulung zu verringern (vgl. etwa Konsortium Bildungsberichterstattung, 2006). Die Prioritäten scheinen 15 Jahre später jedoch etwas anders gelagert zu sein, denn obwohl mehr Jugendliche als in PISA 2000 angaben, erst im Alter von 7 Jahren eingeschult worden zu sein, hat sich der Anteil der Fünfzehnjährigen, die bereits die 10. Klassenstufe besuchen, deutlich erhöht (von 24 Prozent auf 40 Prozent). Ein möglicher Faktor, der sich hier niederschlägt, ist der Umgang mit Klassenwiederholungen, der in den verschiede-

© Waxmann Verlag GmbH

nen OECD-Staaten unterschiedlich gehandhabt wird und sich bei der Verteilung fünfzehnjähriger Schülerinnen und Schüler auf mehrere Klassenstufen manifestiert.

In einer Reihe von OECD-Staaten gibt es gar keine Klassenwiederholungen, in anderen Staaten sind Klassenwiederholungen ein mehr oder weniger häufig eingesetztes Instrument, und in einer dritten Gruppe besteht die Möglichkeit, auf individuellen Wunsch der Schülerinnen und Schüler ein Schuljahr zu wiederholen (vgl. etwa auch OECD, 2010). Üblicherweise besuchen Schülerinnen und Schüler im Verlauf ihrer Schulkarriere mit zunehmendem Alter höhere Klassenstufen, entsprechend steigt mit dem Alter auch die Klassenstufe an. Wenn während der Schullaufbahn eine oder mehrere Klassenstufen wiederholt werden, verzögert sich der schulische Werdegang.

In Tabelle 5.3 ist dargestellt, wie viele Schülerinnen und Schüler bei PISA 2015 in den OECD-Staaten angaben, bereits mindestens einmal ein Schuljahr wiederholt zu haben. Dabei werden Klassenwiederholungen in der Primarstufe von solchen in der Sekundarstufe I unterschieden.

Vor dem PISA-Testtag haben im OECD-Mittel 12 Prozent der Fünfzehnjährigen bereits mindestens einmal eine Klasse wiederholt. Davon entfallen 7 Prozent auf Klassenwiederholungen während der Primarstufe und knapp 6 Prozent während der Sekundarstufe.

In Norwegen, Japan und während der Grundschule in Slowenien existiert die Praxis der Klassenwiederholung nicht. In den nordischen Schulsystemen (Dänemark, Finnland, Norwegen und Schweden) erfolgen Klassenwiederholungen nur auf ausdrücklichen Wunsch der Schülerinnen und Schüler sowie ihrer Eltern. Entsprechend gering (maximal 4 Prozent) sind dort die Wiederholungsquoten bis zum Alter von fünfzehn Jahren. Auch im Vereinigten Königreich ist das erneute Absolvieren eines Schuljahres nur in Ausnahmefällen möglich. Bis vor wenigen Jahren bestand auch in Korea keine Möglichkeit für eine Klassenwiederholung, wobei in PISA 2012 und auch in PISA 2015 eine kleine Gruppe von Schülerinnen und Schülern angab, ein Schuljahr zweimal besucht zu haben. Mit Blick auf die in Kapitel 2 berichteten Leistungsergebnisse wird deutlich, dass Staaten, in denen Klassenwiederholungen nicht oder nur in Ausnahmefällen als Instrument der Förderung eingesetzt werden, durchaus sehr starke durchschnittliche Leistungen im PISA-Naturwissenschaftstest und teils auch eine relativ niedrige Leistungsvarianz vorzuweisen haben.

In Deutschland beläuft sich die Wiederholungsquote auf 18 Prozent und liegt damit über dem OECD-Durchschnitt. Sie ist ähnlich hoch wie in der Schweiz (20 Prozent), wobei die Wiederholungen in der Schweiz häufiger in der Primarstufe erfolgen als in der Sekundarstufe. Deutlich höher, nämlich zwischen 31 und 34 Prozent, ist die Klassenwiederholungsrate in Belgien, Luxemburg, Portugal und Spanien. Mit Ausnahme Belgiens befinden sich die Mittelwerte der naturwissenschaftlichen Kompetenz in diesen Staaten im Bereich des OECD-Mittelwerts und darunter.

Obwohl viel dafür spricht, dass die Wiederholung eines Schuljahres aufgrund mangelhafter Leistungen in der Regel nicht zu den gewünschten Effekten führt, ist dieses Thema nach wie vor Gegenstand öffentlicher wie auch wissenschaftlicher Diskussionen

© Waxmann Verlag GmbH

Tabelle 5.3: Prozentualer Anteil der Schülerinnen und Schüler, die mindestens eine Klasse wiederholt haben

	Prozentualer Anteil der Schülerinnen und Schüler, die mindestens eine Klasse wiederholt haben								Klassenwiederholungen in der bisherigen Schullaufbahn	
	Grundschule				Sekundarstufe I					
	einmal		zweimal und mehr		einmal		zweimal und mehr			
OECD-Staaten	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
Australien	6.0	(0.3)	0.3	(0.1)	1.0	(0.1)	0.2	(0.0)	7.1	(0.3)
Belgien	17.5	(0.6)	2.0	(0.2)	13.8	(0.6)	0.9	(0.2)	34.0	(0.8)
Chile	12.0	(0.5)	2.5	(0.5)	5.6	(0.4)	1.2	(0.3)	24.6	(0.9)
Dänemark	2.7	(0.2)	0.3	(0.1)	0.5	(0.1)	0.2	(0.1)	3.4	(0.3)
Deutschland	8.8	(0.6)	0.4	(0.1)	10.0	(0.6)	0.6	(0.1)	18.1	(0.8)
Estland	2.3	(0.3)	0.6	(0.1)	1.3	(0.2)	0.4	(0.1)	4.0	(0.4)
Finnland	2.3	(0.2)	0.3	(0.1)	0.5	(0.1)	0.2	(0.1)	3.0	(0.2)
Frankreich	12.3	(0.5)	0.4	(0.1)	10.1	(0.5)	0.6	(0.1)	22.1	(0.6)
Griechenland	1.2	(0.2)	0.8	(0.2)	2.8	(0.5)	1.3	(0.2)	5.0	(0.7)
Irland	6.3	(0.4)	0.1	(0.0)	0.9	(0.1)	0.0	(0.0)	7.2	(0.5)
Island	0.4	(0.1)	0.4	(0.1)	0.3	(0.1)	0.4	(0.1)	1.1	(0.2)
Israel	2.9	(0.3)	0.9	(0.1)	4.3	(0.3)	1.5	(0.2)	9.0	(0.6)
Italien	1.2	(0.2)	0.4	(0.1)	4.9	(0.4)	0.9	(0.2)	15.1	(0.6)
Japan	m	m	m	m	m	m	m	m	m	m
Kanada	3.1	(0.2)	0.3	(0.1)	2.4	(0.2)	0.4	(0.1)	5.7	(0.4)
Korea	2.6	(0.2)	1.7	(0.2)	2.6	(0.2)	1.5	(0.2)	4.7	(0.3)
Lettland	2.9	(0.3)	0.6	(0.2)	1.7	(0.2)	0.4	(0.1)	5.0	(0.4)
Luxemburg	14.6	(0.4)	1.6	(0.2)	17.5	(0.5)	1.2	(0.2)	30.9	(0.5)
Mexiko	10.9	(0.7)	1.3	(0.2)	3.9	(0.5)	0.3	(0.1)	15.8	(0.9)
Neuseeland	3.6	(0.3)	0.3	(0.1)	1.5	(0.2)	0.3	(0.1)	4.9	(0.3)
Niederlande	14.2	(0.5)	0.3	(0.1)	6.7	(0.4)	0.1	(0.1)	20.1	(0.5)
Norwegen	m	m	m	m	m	m	m	m	m	m
Österreich	5.9	(0.4)	0.1	(0.0)	4.2	(0.4)	0.3	(0.1)	15.2	(0.7)
Polen	1.8	(0.2)	0.4	(0.1)	3.3	(0.3)	0.2	(0.1)	5.3	(0.4)
Portugal	11.4	(0.7)	5.4	(0.4)	16.9	(0.8)	4.6	(0.4)	31.2	(1.2)
Schweden	2.7	(0.3)	0.3	(0.1)	1.6	(0.2)	0.3	(0.1)	4.0	(0.4)
Schweiz	13.2	(0.8)	0.6	(0.1)	7.6	(0.6)	0.4	(0.1)	20.0	(1.0)
Slowakische Republik	3.1	(0.3)	1.1	(0.2)	2.5	(0.3)	1.0	(0.2)	6.5	(0.5)
Slowenien	m	m	m	m	1.5	(0.3)	0.2	(0.1)	1.9	(0.3)
Spanien	12.0	(0.6)	0.7	(0.1)	24.4	(0.8)	2.2	(0.3)	31.3	(1.0)
Tschechische Republik	1.9	(0.2)	0.4	(0.1)	2.8	(0.3)	0.3	(0.1)	4.8	(0.4)
Türkei	3.0	(0.3)	0.3	(0.1)	1.2	(0.2)	0.3	(0.1)	10.9	(0.7)
Ungarn	4.2	(0.5)	0.7	(0.2)	4.5	(0.5)	0.5	(0.1)	9.5	(0.6)
Vereinigte Staaten	8.6	(0.6)	0.3	(0.1)	2.9	(0.3)	0.4	(0.1)	11.0	(0.8)
Vereinigtes Königreich	1.8	(0.2)	0.3	(0.1)	0.6	(0.1)	0.2	(0.1)	2.8	(0.3)
OECD-Durchschnitt	6.2	(0.1)	0.8	(0.0)	5.0	(0.1)	0.7	(0.0)	12.0	(0.3)

m: fehlende Angaben

© Waxmann Verlag GmbH

(vgl. etwa Boller, Möller & Palowski, 2013; Ehmke, Drechsel & Carstensen, 2008; Hong & Raudenbush, 2005; Krohne, Meier & Tillmann, 2011). Aus der Perspektive der unterrichtenden Lehrperson können Klassenwiederholungen insofern von Vorteil sein, als damit in Lerngruppen die Leistungshomogenität gestärkt beziehungsweise wiederhergestellt wird. Von Lehrkräften und zum Teil auch Eltern wird die ernsthafte Drohung mit Klassenwiederholung als relevanter Motivator für Lernanstrengungen betrachtet. Aus der Sicht der betroffenen Schülerinnen und Schüler überwiegen aufgrund des Entzugs von Lerngelegenheiten, des Verlusts relevanter sozialer Einbindung durch Peers oder auch durch die Minderung der Lernmotivation (vgl. Ehmke et al., 2008) die Nachteile. Das Verhältnis von aufgewendeten Ressourcen zu den Erträgen einer Klassenwiederholung stellt sich oftmals als unvorteilhaft heraus, da wenig empirische Evidenz vorliegt, dass die Wiederholung einer Klasse als wirksame pädagogische Maßnahme gelten kann (Alexander, Entwisle & Dauber, 2003; Klemm, 2009). Der derzeitige Stand einschlägiger Forschung zum Thema liefert keine überzeugende Evidenz, eine verpflichtende Klassenwiederholung als empfehlenswerte, wirksame pädagogische Maßnahme zu bezeichnen.

Betrachtet man die besuchte Klassenstufe der Fünfzehnjährigen sowie die Klassenwiederholungen in Deutschland differenziert nach Schuljahr, so zeigen sich Unterschiede in Bezug auf die bei der Stichprobenziehung noch vorhandene Differenzierung nach fünf Schularten. Tabelle 5.4 stellt dar, wie sich die Stichprobe der Schülerinnen und Schüler in PISA 2015 auf verschiedene Schularten und darin auf die Klassenstufen verteilt. Daneben sind auch die prozentualen Anteile der Fünfzehnjährigen, die bereits mindestens einmal ein Schuljahr wiederholt haben, getrennt nach Primar- und Sekundarstufe I abgebildet. In der Darstellung werden die fünf Schularten Hauptschule, Schule mit mehreren Bildungsgängen, Integrierte Gesamtschule, Realschule und Gymnasium differenziert, um auch allfällige Schulartwechsel zu erfassen, die teilweise mit Klassenwiederholungen verbunden sind.

Dem oberen Teil der Tabelle kann entnommen werden, dass sich die Fünfzehnjährigen in Deutschland an allen Schularten außer den Gymnasien mehrheitlich in der 9. Klassenstufe befinden. Zwischen den Schularten unterscheidet sich die Verteilung der PISA-Kohorte auf die unterschiedlichen Klassenstufen. An Hauptschulen etwa hat etwas mehr als ein Fünftel die 9. Klassenstufe noch nicht erreicht, etwa ebenso viele Schülerinnen und Schüler befinden sich in der 10. Klassenstufe. Deutlich geringer ist der Anteil der Fünfzehnjährigen, die höchstens in Klassenstufe 8 sind, an allen übrigen Schularten. Das Gymnasium ist die einzige Schulart, an der die Mehrheit der Fünfzehnjährigen bereits die 10. Klassenstufe besucht (53 Prozent). Insgesamt zeigt sich für alle Schularten, dass mittlerweile beinahe genauso viele Fünfzehnjährige in der 10. Klassenstufe sind (44 Prozent) wie in der 9. Klassenstufe (47 Prozent). Auch diese Anteile sprechen dafür, dass sich in Bezug auf die Verzögerung von Schullaufbahnen durch Klassenwiederholungen seit PISA 2000 einiges verändert hat und fünfzehnjährige Schülerinnen und Schüler inzwischen meist eine höhere Klassenstufe besuchen als noch vor 15 Jahren: In PISA 2000 besuchten auch an Gymnasien lediglich 30 Prozent der Fünfzehnjährigen

© Waxmann Verlag GmbH

Tabelle 5.4: Anteile der Fünfzehnjährigen in Prozent, Verteilung auf Klassenstufen und Klassenwiederholungen

	Hauptschule	Schule mit mehreren Bildungsgängen	Integrierte Gesamtschule	Realschule	Gymnasium	alle Schulen[1]
Klassenstufe						
Klasse 7 oder 8	21.1	9.4	6.7	9.8	2.4	7.9
Klasse 9	58.1	55.9	49.1	49.9	39.9	47.1
Klasse 10	20.8	34.5	43.9	40.2	53.3	43.5
Klasse 11		0.2	0.2		4.2	1.5
Wiederholer						
Grundschule	24.7	14.3	16.9	7.4	0.8	8.8
Sekundarstufe I	15.0	10.2	5.0	17.9	6.9	10.6

[1] Ohne berufliche Schulen und Förderschulen.

die 10. Klassenstufe, an den übrigen Schularten zwischen 5 Prozent (Hauptschule) und 25 Prozent (Integrierte Gesamtschule).

In Bezug auf Klassenwiederholungen bildet Tabelle 5.4 ab, dass sich in den Hauptschulen der größte prozentuale Anteil von Fünfzehnjährigen befindet, die in ihrer bisherigen Schullaufbahn bereits ein Schuljahr wiederholt haben. Im Gegensatz zu früheren PISA-Erhebungsrunden ist der Anteil der Wiederholer in der Sekundarstufe I in PISA 2015 jedoch nicht mehr an Hauptschulen, sondern an Realschulen am höchsten. Die Verteilung der Wiederholer auf unterschiedliche Schularten dürfte zumindest zum Teil auf Schulartwechsel zurückzuführen sein: Wer beispielsweise im Gymnasium nicht versetzt wird, wechselt unter Umständen zu diesem Zeitpunkt auf eine andere Schulart. Häufig finden die Klassenwiederholungen also nicht innerhalb einer Schule beziehungsweise innerhalb einer Schulart statt, sondern bringen einen Wechsel der Schulart mit sich und werden statistisch der aufnehmenden Schulart zugerechnet. Im Vergleich zur vorherigen PISA-Erhebungsrunde 2012 scheint diese Praxis allerdings weniger gängig zu sein: Während in PISA 2012 an Hauptschulen beinahe jeder dritte Fünfzehnjährige (30 Prozent) ein Schuljahr in der Grundschule wiederholt hatte und gut jeder vierte (27 Prozent) eines in der Sekundarstufe I, sind diese Anteile in PISA 2015 mit 25 beziehungsweise 15 Prozent erheblich geringer. In Deutschland ist die Klassenwiederholungsquote insgesamt seit PISA 2000 von damals 24 Prozent (Baumert et al., 2001) auf nun 18 Prozent bei PISA 2015 gesunken.

© Waxmann Verlag GmbH

5.1.3 Lernzeit für naturwissenschaftlichen Unterricht

Wie viel Zeit Schülerinnen und Schüler mit Lernen innerhalb und außerhalb der Schule verbringen, unterscheidet sich im Vergleich der OECD-Staaten teils erheblich. Tabelle 5.5 bildet zunächst ab, wie viele Zeitstunden pro Woche die fünfzehnjährigen Schülerinnen und Schüler mit *Unterricht* in naturwissenschaftlichen Fächern, Mathematik sowie ihrer Testsprache (in Deutschland: Deutschunterricht) verbringen. International schwankt die Dauer einer Schulstunde beziehungsweise einer Unterrichtslektion stark. Deshalb sind hier als Vergleichseinheit Zeitstunden (60 Minuten) angegeben. Die Zahlen in Tabelle 5.5 beruhen auf Angaben der Schülerinnen und Schüler.

In den OECD-Staaten werden pro Woche durchschnittlich etwa dreieinhalb Stunden Unterricht in den Fächern Naturwissenschaften, Mathematik oder in der Testsprache erteilt. In Deutschland liegt die Dauer der mit Unterricht in naturwissenschaftlichen Fächern verbrachten Zeit mit 3.7 Stunden pro Woche leicht über dem OECD-Mittelwert, während sie in der Mathematik und in Deutsch etwas darunter liegt. Im Vergleich zu PISA 2012 erhalten die Fünfzehnjährigen eine halbe Stunde pro Woche weniger Unterricht in den Naturwissenschaften (vgl. Sälzer, Prenzel & Klieme, 2013). Gegenüber PISA 2006, als die Naturwissenschaften erstmals die Schwerpunktdomäne bildeten, hat sich die Lernzeit um eine gute halbe Stunde erhöht (OECD, 2011).

Tabelle 5.5 bildet die OECD-Staaten in absteigender Reihenfolge der wöchentlichen Zeitstunden für naturwissenschaftlichen Unterricht ab. Bezieht man die in Tabelle 5.5 angegebenen Zeitspannen auf die in Kapitel 2, 6 und 7 berichteten Kompetenzmittelwerte für die einzelnen Staaten, so wird klar, dass die Sortierung der Staaten nicht ihrer Reihenfolge in der mittleren naturwissenschaftlichen Kompetenz entspricht. Korrelationsanalysen bestätigen, dass es keinen unmittelbar linearen Zusammenhang zwischen der Menge an Unterrichtszeit und der erreichten Kompetenz gibt. Dies ist nicht erstaunlich, da PISA die bis zum Testzeitpunkt kumulativ erworbenen Kompetenzen misst. Vielmehr müsste man also die Lernzeiten bis zum Testzeitpunkt zusammenrechnen (Baumert, Bos & Watermann, 2000). Tabelle 5.6 präsentiert den Anteil der Fünfzehnjährigen in den OECD-Staaten, die jeweils besonders wenig (< 2 Stunden) oder besonders viel Zeit (> 4 Stunden) pro Woche mit Lernen für naturwissenschaftliche Fächer und Themengebiete verbringen (vgl. dazu die Analysen zu PISA 2006 von Kobarg et al., 2011). Berücksichtigt werden hier sowohl der naturwissenschaftliche Unterricht als auch außerunterrichtliche Lernaktivitäten.

Ein Blick in Tabelle 5.6 zeigt, dass knapp 10 Prozent der Fünfzehnjährigen in der OECD weniger als 2 Stunden pro Woche mit Unterricht in naturwissenschaftlichen Fächern verbringen. In PISA 2006 waren dies deutlich mehr (33 Prozent, vgl. Kobarg et al., 2011). In Portugal und der Slowakischen Republik ist dieser Anteil besonders hoch (32 bzw. 21 Prozent), aber auch in leistungsstarken Staaten wie beispielsweise in Kanada befassen sich 17 Prozent der Schülerinnen und Schüler weniger als 2 Stunden wöchentlich mit naturwissenschaftlichem Unterricht. In Deutschland macht diese Gruppe 8 Prozent der Fünfzehnjährigen aus. Besonders viel Zeit, mehr als 4 Stunden pro Woche, ver-

© Waxmann Verlag GmbH

Tabelle 5.5: Anzahl der Zeitstunden pro Woche, die in OECD-Staaten mit Unterricht in den Naturwissenschaften, in Mathematik oder der Testsprache verbracht werden

	Normaler Naturwissenschaftsunterricht		Normaler Mathematikunterricht		Normaler Unterricht in der Testsprache	
	Lernzeit (Zeitstunden)		Lernzeit (Zeitstunden)		Lernzeit (Zeitstunden)	
OECD-Staaten	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
Chile	5.8	(0.2)	7.1	(0.2)	6.7	(0.2)
Kanada	4.9	(0.0)	5.0	(0.0)	5.2	(0.0)
Österreich	4.9	(0.0)	2.4	(0.0)	2.4	(0.0)
Vereinigtes Königreich	4.7	(0.0)	4.0	(0.1)	4.1	(0.0)
Niederlande	4.4	(0.0)	2.6	(0.0)	2.9	(0.0)
Lettland	4.3	(0.0)	3.8	(0.0)	2.7	(0.0)
Neuseeland	4.2	(0.0)	4.1	(0.0)	4.1	(0.0)
Tschechische Republik	4.2	(0.0)	3.1	(0.0)	3.1	(0.0)
Vereinigte Staaten	4.0	(0.0)	4.1	(0.0)	4.3	(0.0)
Mexiko	3.9	(0.0)	4.0	(0.0)	3.9	(0.1)
Griechenland	3.8	(0.0)	3.4	(0.0)	2.7	(0.0)
Portugal	3.8	(0.0)	4.4	(0.0)	4.0	(0.0)
Deutschland	3.7	(0.1)	3.2	(0.1)	3.3	(0.0)
Estland	3.7	(0.0)	3.5	(0.0)	3.2	(0.0)
Australien	3.5	(0.0)	3.7	(0.0)	3.9	(0.0)
Slowenien	3.5	(0.0)	2.7	(0.0)	2.9	(0.0)
Dänemark	3.4	(0.0)	4.2	(0.0)	5.5	(0.0)
Israel	3.4	(0.1)	4.0	(0.1)	3.2	(0.1)
Türkei	3.4	(0.0)	3.7	(0.0)	2.9	(0.0)
Spanien	3.3	(0.0)	3.6	(0.0)	3.4	(0.0)
Luxemburg	3.2	(0.0)	3.5	(0.0)	3.4	(0.0)
Slowakische Republik	3.1	(0.0)	3.1	(0.1)	3.2	(0.0)
Ungarn	3.1	(0.0)	2.6	(0.0)	2.7	(0.0)
Frankreich	3.0	(0.0)	3.6	(0.0)	3.8	(0.0)
Polen	3.0	(0.0)	3.5	(0.0)	3.8	(0.0)
Schweden	3.0	(0.0)	3.2	(0.0)	3.1	(0.0)
Belgien	2.9	(0.0)	3.4	(0.0)	3.5	(0.0)
Japan	2.9	(0.0)	4.0	(0.0)	3.7	(0.0)
Korea	2.9	(0.0)	3.5	(0.0)	3.3	(0.0)
Finnland	2.8	(0.0)	2.9	(0.0)	2.5	(0.0)
Italien	2.6	(0.0)	3.8	(0.0)	4.6	(0.0)
Schweiz	2.5	(0.0)	3.5	(0.0)	3.4	(0.0)
Irland	2.4	(0.0)	3.2	(0.0)	3.0	(0.0)
Norwegen	2.4	(0.0)	3.3	(0.0)	3.7	(0.0)
Island	2.3	(0.0)	3.9	(0.0)	4.0	(0.0)
OECD-Durchschnitt	3.5	(0.0)	3.7	(0.0)	3.6	(0.0)

© Waxmann Verlag GmbH

Tabelle 5.6: Lernzeit für Naturwissenschaften pro Woche in den OECD-Staaten (Zeitstunden)

	Normaler Naturwissenschaftsunterricht				Außerschulische Lernzeit (Hausaufgaben, Zusatzunterricht, Klausurvorbereitung)			
	Weniger als 2 Stunden		Mehr als 4 Stunden		Weniger als 2 Stunden		Mehr als 4 Stunden	
OECD-Staaten	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*
Australien	10.4	(0.2)	27.1	(0.1)	27.2	(0.5)	26.3	(0.1)
Belgien	19.1	(0.6)	30.1	(0.2)	33.6	(0.5)	17.8	(0.1)
Chile	3.3	(0.3)	65.2	(0.3)	27.0	(0.5)	35.4	(0.1)
Dänemark	1.9	(0.2)	33.1	(0.2)	18.2	(0.5)	40.9	(0.1)
Deutschland	8.4	(0.5)	48.9	(0.3)	53.7	(0.8)	10.3	(0.1)
Estland	10.1	(0.1)	59.5	(0.2)	27.5	(0.6)	30.0	(0.1)
Finnland	6.7	(0.6)	28.3	(0.4)	50.7	(0.7)	11.4	(0.1)
Frankreich	17.9	(0.6)	21.1	(0.1)	36.2	(0.6)	18.6	(0.1)
Griechenland	6.4	(0.3)	65.5	(0.2)	11.0	(0.4)	48.5	(0.2)
Irland	9.0	(0.6)	15.5	(0.2)	31.6	(0.6)	21.8	(0.1)
Island	4.6	(0.2)	7.2	(0.1)	44.4	(0.8)	10.6	(0.1)
Israel	10.6	(0.8)	38.5	(0.2)	41.5	(1.2)	27.3	(0.1)
Italien	5.2	(0.3)	8.5	(0.1)	18.5	(0.4)	35.0	(0.1)
Japan	3.4	(0.6)	16.4	(0.3)	55.4	(0.9)	12.8	(0.1)
Kanada	17.0	(0.4)	58.0	(0.1)	22.5	(0.4)	43.5	(0.1)
Korea	3.9	(0.4)	10.7	(0.2)	34.6	(0.7)	23.1	(0.1)
Lettland	1.5	(0.2)	76.8	(0.2)	30.9	(0.7)	29.8	(0.2)
Luxemburg	10.8	(0.2)	19.9	(0.1)	35.1	(0.5)	24.8	(0.1)
Mexiko	5.3	(0.6)	49.7	(0.2)	20.2	(0.5)	45.0	(0.2)
Neuseeland	6.1	(0.4)	31.3	(0.2)	26.5	(0.2)	30.9	(0.1)
Niederlande	16.9	(0.7)	46.9	(0.2)	37.8	(0.8)	17.0	(0.1)
Norwegen	2.0	(0.2)	2.9	(0.1)	29.5	(0.5)	21.8	(0.1)
Österreich	11.0	(0.6)	51.7	(0.2)	28.4	(0.7)	33.8	(0.1)
Polen	4.1	(0.3)	32.4	(0.4)	28.4	(0.6)	25.7	(0.1)
Portugal	31.8	(0.6)	26.7	(0.2)	23.1	(0.7)	37.8	(0.2)
Schweden	5.0	(0.5)	14.3	(0.3)	38.1	(0.6)	17.8	(0.1)
Schweiz	12.5	(0.8)	18.0	(0.1)	51.6	(0.9)	11.8	(0.1)
Slowakische Republik	21.4	(0.9)	40.2	(0.2)	33.5	(0.8)	31.1	(0.1)
Slowenien	2.6	(0.1)	52.7	(0.1)	34.4	(0.7)	23.3	(0.1)
Spanien	16.6	(0.4)	31.1	(0.2)	19.9	(0.5)	35.9	(0.1)
Tschechische Republik	2.7	(0.2)	61.1	(0.3)	38.2	(0.8)	21.6	(0.1)
Türkei	8.8	(0.4)	57.7	(0.4)	14.6	(0.5)	50.2	(0.2)
Ungarn	14.8	(0.8)	37.3	(0.2)	30.7	(0.7)	25.7	(0.1)
Vereinigte Staaten	19.5	(0.8)	59.5	(0.3)	21.9	(0.5)	49.6	(0.2)
Vereinigtes Königreich	1.9	(0.1)	58.8	(0.2)	20.6	(0.5)	36.1	(0.1)
OECD-Durchschnitt	9.5	(0.4)	37.2	(0.2)	31.3	(0.6)	28.1	(0.1)

© Waxmann Verlag GmbH

bringen im OECD-Mittel 37 Prozent der Fünfzehnjährigen (PISA 2006: 29 Prozent, vgl. Kobarg et al., 2011). In Deutschland beschäftigt sich beinahe die Hälfte der PISA-Kohorte mehr als 4 Stunden jede Woche mit Unterricht in Naturwissenschaften, was signifikant über dem OECD-Mittelwert liegt. Besonders groß ist diese Gruppe beispielsweise in Chile, Griechenland und Lettland, deren mittleres Kompetenzniveau in den Naturwissenschaften durchschnittlich bis unterdurchschnittlich ist. Aber auch in Estland und Kanada, sehr leistungsstarken Bildungssystemen, beläuft sich der Anteil der Schülerinnen und Schüler, die mehr als 4 Stunden pro Woche mit Naturwissenschaftsunterricht verbringen, auf rund 60 Prozent. Besonders klein ist diese Gruppe in Norwegen und Island.

In PISA 2006, als die Naturwissenschaften zum ersten Mal als Schwerpunktdomäne erfasst worden sind, stellte sich heraus, dass die Naturwissenschaften in den Stundentafeln der OECD-Staaten einen vergleichsweise geringen Stellenwert hatten (vgl. OECD, 2007; Senkbeil, Drechsel & Schöps, 2007). Dies zeigte sich daran, dass über alle OECD-Staaten hinweg 9 Prozent der Schülerinnen und Schüler im Schuljahr des PISA-Tests überhaupt keinen Unterricht in den naturwissenschaftlichen Fächern erhalten hatten. In PISA 2015 trifft dies immer noch auf 6 Prozent der Fünfzehnjährigen zu. Nach wie vor auffallend sind die enormen Unterschiede zwischen den Staaten: Neben Staaten, in denen praktisch niemand im PISA-Schuljahr keinen Unterricht in Naturwissenschaften hatte (z. B. Estland, Polen, Österreich, Finnland), gibt es OECD-Staaten, in denen dieser Anteil mit bis zu 61 Prozent der Fünfzehnjährigen durchaus bedeutsam ist (z. B. Slowenien). Teilweise erreichen die betreffenden Staaten, etwa Slowenien, in den Naturwissenschaften überdurchschnittliche Kompetenzmittelwerte, die sich nicht signifikant vom Mittelwert Fünfzehnjähriger in Deutschland unterscheiden. In Deutschland befindet sich der Anteil der Schülerinnen und Schüler, die laut eigenen Angaben im Schuljahr des PISA-Tests überhaupt keinen naturwissenschaftlichen Unterricht haben, mit 5 Prozent im Bereich des OECD-Durchschnitts und nahe am Wert von PISA 2006 (damals: 6 Prozent).

In Bezug auf die außerunterrichtliche Lernzeit sind über alle OECD-Staaten hinweg die Gruppen Fünfzehnjähriger, die sich besonders wenig (< 2 Stunden pro Woche) oder besonders viel (> 4 Stunden pro Woche) mit Naturwissenschaften befassen, mit jeweils etwa 30 Prozent annähernd gleich groß. In Deutschland fällt auf, dass diese Anteile deutlich auseinanderliegen: Während mehr als die Hälfte der Schülerinnen und Schüler (54 Prozent) angibt, weniger als 2 Stunden jede Woche neben dem Pflichtunterricht für außerschulische naturwissenschaftliche Lernzeit aufzuwenden, gilt dies nur für etwa ein Zehntel der Fünfzehnjährigen für mehr als 4 Stunden wöchentlich. Die Gruppe derjenigen, die sich besonders viel neben dem Unterricht mit naturwissenschaftlicher Lernzeit befassen, ist unter den OECD-Staaten in Deutschland am kleinsten. Dabei variiert die Anzahl der Stunden, in welchen sich Fünfzehnjährige pro Woche außerhalb des schulischen Unterrichts mit Lernen (Hausaufgaben, Klausurvorbereitung) in Naturwissenschaften beschäftigen, innerhalb der OECD-Staaten beträchtlich. So geben die Schülerinnen und Schüler in Deutschland im Mittel an, wöchentlich 2 Stunden außerhalb des

© Waxmann Verlag GmbH

regulären Unterrichts beispielsweise mit Hausaufgaben, Nachhilfe oder dem Lernen für naturwissenschaftliche Fächer zu verbringen. Diese Zeitdauer ist deutlich geringer als im OECD-Mittel (9 Stunden) und unter allen OECD-Staaten der niedrigste Wert. Wie genau die Zeit ausgestaltet ist, welche die Fünfzehnjährigen außerhalb der Schule mit Lernen in Naturwissenschaften verbringen, ist sehr unterschiedlich. Beispielsweise kann es sich hier um Stützkurse für leistungsschwache Schülerinnen und Schüler handeln oder auch um Talentförderung für besonders leistungsstarke Jugendliche.

Mit Blick auf die Lernzeit der Fünfzehnjährigen lässt sich zusammenfassen, dass es sowohl unter den in den Naturwissenschaften besonders leistungsstarken Bildungssystemen als auch unter den schwächer abschneidenden Staaten sowohl solche gibt, die relativ viele Stunden pro Woche für naturwissenschaftlichen Unterricht aufwenden, als auch solche, die eher wenig Zeit dafür vorsehen. Zur Klärung der Frage, inwieweit die Lernzeit mit der mittleren naturwissenschaftlichen Kompetenz zusammenhängt, bedarf es einer genaueren Betrachtung, *wie* zusätzliche Stunden in den Naturwissenschaften konkret genutzt werden. Etwaige Stützkurse für leistungsschwache Schülerinnen und Schüler oder Angebote zur Talentförderung für besonders leistungsstarke Jugendliche sind hierfür grundlegend unterschiedliche Angebote. Mit einer genaueren inhaltlichen Betrachtung, etwa des Unterrichts in Naturwissenschaften, kann auch der Zusammenhang zwischen der Häufigkeit und Dauer der Beschäftigung mit naturwissenschaftlichen Themen und der durchschnittlichen naturwissenschaftlichen Kompetenz genauer beurteilt werden.

5.2 Die Einzelschule: Gestaltung von Lernumwelten in Deutschland

Schulen befinden sich innerhalb unterschiedlicher Bildungssysteme in bestimmten organisatorischen und institutionellen Rahmenbedingungen. Innerhalb dieser Gegebenheiten werden in den Schulen Lernumwelten geschaffen und gestaltet. Unter den zahlreichen Merkmalen, die in Bezug auf einzelne Schulen im Rahmen der PISA-Studie erhoben werden, sind einige für die Sekundarschullandschaft in Deutschland besonders von Interesse, so beispielsweise, weil der aktuelle Stand der Forschung dazu nicht eindeutig ist, die Befunde scheinbar im Widerspruch zur Erfahrung stehen oder aber weil sich Trends als Entwicklungen über die Zeit beschreiben lassen. Dieses Teilkapitel befasst sich mit den Fragen, inwieweit sich die einzelnen Schulen in Deutschland von denen anderer Staaten unterscheiden, welche Merkmale sie jeweils aufweisen und ob Schulen in Deutschland in dieser Hinsicht Auffälligkeiten zeigen. In den nächsten Abschnitten werden die Themen *Klassengröße*, *Schulentwicklung und Qualitätssicherung*, *schulische Angebote und Ressourcen zur Förderung naturwissenschaftlicher Bildung*, *Schülerverhalten* sowie *Schulklima* behandelt.

© Waxmann Verlag GmbH

5.2.1 Klassengröße

Obwohl einschlägige Forschung zum Thema Klassengröße immer wieder betont, dass diese kaum mit der durchschnittlichen Leistung der Schülerinnen und Schüler zusammenhängt, ist die Frage nach der Größe einer Lerngruppe ein Aspekt, der im öffentlichen Diskurs immer wieder mit den Bedingungen für erfolgreiches Lernen in Verbindung gebracht wird (vgl. u. a. Arnhold, 2005; Hattie, 2008). In Bezug auf die Sekundarschullandschaft in Deutschland ist die Frage nach der Bedeutung der Klassengröße auch abhängig von der Schulart, der Schulstufe und letztlich der einzelnen Schule mit ihren jeweiligen Gegebenheiten. Jedenfalls kann die Klassengröße das Unterrichtsgeschehen und auch die Möglichkeiten auf das individuelle Lernen der Schülerinnen und Schüler einzugehen, in nicht unerheblichem Ausmaß beeinflussen. Intuitiv scheinen kleinere Klassen mehrere Vorteile zu bieten: Das Unterrichten fällt leichter, Lernprozesse können eher individuell gefördert und angeleitet werden, der Geräuschpegel ist potenziell niedriger als in großen Klassen. In großen Klassen scheint es naheliegend zu sein, dass das *Classroom-Management* herausfordernder ist und es häufiger Unterbrechungen oder Störungen des Unterrichts gibt. Allerdings werden diese verbreiteten Annahmen durch die derzeit vorliegende, einschlägige empirische Befundlage nicht konsistent gestützt (Arnhold, 2005; Hattie, 2008; Helmke & Jäger, 2002). Grundsätzlich scheint es zwar möglich zu sein, in kleinen Klassen lernförderlicheren Unterricht zu gestalten (Klieme, 2006), jedoch müssen die Lehrkräfte auch entsprechend die Vorteile kleiner Lerngruppen für ihren Unterricht nutzen (Im Brahm, 2006). Solange sich der Unterricht in großen oder kleinen Klassen nicht voneinander unterscheidet, werden die pädagogischen Möglichkeiten nicht ausgeschöpft und es ist nicht damit zu rechnen, dass der Lernerfolg in kleinen Klassen besonders hoch ist. So sprechen die Befunde auf der Basis von Daten aus *Large-Scale-Assessments* wie PISA nicht dafür, dass kleine Klassen systematisch einen höheren Lernerfolg mit sich bringen (Bromme, Prenzel & Jäger, 2014). Wo es jedoch gelingt, den Unterricht der Klassengröße anzupassen, werden durchaus substanzielle Erfolge erzielt (vgl. Helmke et al., 2008).

Auf internationaler Ebene ist die Vergleichbarkeit von Schulklassen oder Klassengrößen nur bedingt möglich. Während eine Schulklasse in Deutschland in aller Regel eine Gruppe von Schülerinnen und Schülern ist, die für das laufende oder mehrere Schuljahre einen gemeinsamen Stundenplan hat, von einer Klassenlehrkraft betreut und normalerweise von den gleichen Lehrkräften unterrichtet wird, so ist mit dem Begriff „Klasse“ in anderen Bildungssystemen teilweise etwas anderes gemeint. Wenn etwa in Staaten oder Schularten innerhalb einer Schule leistungs- oder interessensdifferenzierte Lerngruppen in einem bestimmten Fach zusammen unterrichtet werden („class of mathematics“), dann steht der spezifische Inhalt im Vordergrund und das Konzept entspricht nicht der Schulklasse in Deutschland. Ein Vergleich allein anhand der Zahl von Schülerinnen und Schülern, die in den PISA-Studien als „Klasse“ bezeichnet werden, ist daher stets mit Vorsicht zu interpretieren. Tabelle 5.7 gibt einen Überblick über die durchschnittlichen Klassengrößen in der Sekundarstufe I in allen OECD-Staaten.

© Waxmann Verlag GmbH

Die Angaben zur durchschnittlichen Klassengröße beziehen sich im internationalen Vergleich auf Unterricht in der Testsprache (in Deutschland also Deutschunterricht) und wurden wie in bisherigen PISA-Erhebungsrunden den Angaben der Schulleitungen entnommen. An Schularten, bei denen der Klassenverband fächerübergreifend erhalten bleibt, trifft diese Angabe auch auf den Unterricht in Naturwissenschaften zu; für Schularten mit unterschiedlichen Bildungsgängen oder Leistungsgruppen liefern die PISA-Daten keine Auskunft über die Größe der Lerngruppe in naturwissenschaftlichen Fächern.

Eine Schulklasse (bezogen auf den Unterricht in der Testsprache) besteht im OECD-Durchschnitt aus 26 Schülerinnen und Schülern. Der Mittelwert für Deutschland liegt mit durchschnittlich 25 Schülerinnen und Schülern nahe an diesem Durchschnitt. Entsprechend hat sich die Klassengröße in den OECD-Staaten gegenüber PISA 2006 nicht verändert (OECD, 2008), ist jedoch etwas höher als in der letzten Erhebungsrunde 2012 (24 Schüler, vgl. Sälzer et al., 2013). Wie bei zahlreichen anderen Merkmalen auch, ist die Streubreite der Klassengrößen zwischen den OECD-Staaten enorm. Ähnliche Klassengrößen wie Deutschland haben die europäischen Nachbarstaaten Niederlande, Österreich, Polen, Slowenien und Tschechische Republik. Auch einige in den Naturwissenschaften besonders leistungsstarke Bildungssysteme wie Estland oder Kanada unterrichten in ähnlichen Klassengrößen wie Deutschland. Deutlich andere Konstellationen sind jedoch ebenfalls in europäischen Nachbarstaaten zu finden; etwa in der Schweiz, wo die durchschnittlichen Klassengrößen im Unterricht in der Testsprache bei 20 Schülerinnen und Schülern liegen oder in Frankreich mit 29 Jugendlichen pro Klasse.

5.2.2 Schulentwicklung und Qualitätssicherung

Unter anderem angestoßen durch die Befunde aus den ersten PISA-Erhebungsrunden ist die Frage, wie sich die Qualität von Schulen und Bildungssystemen messen und verbessern lässt, ebenso Teil der bildungspolitischen Agenda wie Aspekte der Schulentwicklung. Auf der Basis von Angaben der Schulleitungen erfasst PISA 2015 den Einsatz von Instrumenten zur Qualitätssicherung und Schulentwicklung. Tabelle 5.8 gibt einen Überblick darüber, wie häufig diese Instrumente an Schulen der OECD-Staaten zum Tragen kamen.

Anders als im OECD-Durchschnitt hat sich das Bild für Deutschland im Vergleich zu PISA 2012, als die in Tabelle 5.8 berichteten Instrumente zur Qualitätssicherung und -entwicklung in Schulen ebenfalls im nationalen Berichtsband dargestellt worden waren, deutlich verändert. Während vor drei Jahren lediglich eines der genannten Instrumente ähnlich häufig eingesetzt wurde wie im Mittel der OECD-Staaten, so trifft dies in PISA 2015 bereits auf vier dieser sechs ausgewählten Instrumente zu. Nahezu alle Schulen in den OECD-Staaten geben an, eine interne Evaluation beziehungsweise Selbstevaluation zu nutzen (im Mittel 93 Prozent). In Deutschland ist dieser Anteil geringfügig niedriger (89 Prozent). Dabei wenden in Deutschland 62 Prozent der Schulen dieses Instrument der Qualitätssicherung auf freiwilliger Basis an und nicht, weil sie dazu verpflichtet

© Waxmann Verlag GmbH

Tabelle 5.7: Durchschnittliche Klassengröße in den OECD-Staaten (Sekundarstufe I)

OECD-Staaten	Durchschnittliche Klassengröße in der Sekundarstufe I	
	M	*(SE)*
Japan	36	(0.3)
Estland	25	(0.4)
Finnland	19	(0.2)
Kanada	26	(0.2)
Korea	31	(0.3)
Neuseeland	25	(0.2)
Slowenien	26	(0.0)
Australien	25	(0.1)
Deutschland	25	(0.3)
Niederlande	26	(0.3)
Vereinigtes Königreich	24	(0.3)
Schweiz	20	(0.5)
Irland	25	(0.3)
Belgien	20	(0.3)
Dänemark	22	(0.3)
Polen	24	(0.5)
Portugal	26	(0.3)
Norwegen	24	(0.3)
Vereinigte Staaten	26	(0.4)
Frankreich	29	(0.3)
Österreich	24	(0.5)
Schweden	23	(0.3)
Spanien	27	(0.5)
Tschechische Republik	24	(0.2)
Lettland	21	(0.3)
Luxemburg	21	(0.0)
Italien	23	(0.5)
Ungarn	28	(0.6)
Island	20	(0.0)
Israel	30	(0.4)
Slowakische Republik	22	(0.3)
Griechenland	24	(0.5)
Chile	34	(0.4)
Türkei	47	(1.0)
Mexiko	39	(0.6)
OECD-Durchschnitt	26	(0.1)

Die Staaten sind in absteigender Reihenfolge der Kompetenzmittelwerte für Naturwissenschaften sortiert.

© Waxmann Verlag GmbH

Tabelle 5.8: Instrumente der Qualitätssicherung und -entwicklung in Schulen

	Prozentualer Anteil von Schülerinnen und Schülern, deren Schulleitung angab, Instrumente zur Qualitätssicherung und -entwicklung einzusetzen											
	Systematische Erfassung von Daten, z. B. Anwesenheit, Prüfungsergebnisse		Interne Evaluation/ Selbstevaluation		Externe Evaluation		Mentorat für Lehrkräfte		Expertengespräche zur Schulentwicklung		Umsetzung standardisierter Vereinbarungen für den Unterricht in Naturwissenschaften	
OECD-Staaten	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*	*%*	*(SE)*
Australien	99.3	(2.1)	98.7	(2.1)	81.4	(1.6)	97.7	(1.7)	77.6	(1.8)	82.3	(2.1)
Belgien	88.6	(3.6)	84.7	(2.7)	85.7	(2.4)	82.1	(1.9)	47.4	(1.7)	55.4	(2.9)
Chile	92.1	(3.6)	93.8	(3.1)	76.5	(3.4)	56.7	(2.4)	42.4	(3.2)	57.5	(3.0)
Dänemark	89.3	(3.2)	84.2	(3.6)	69.5	(2.7)	65.7	(2.6)	38.0	(2.7)	61.3	(2.8)
Deutschland	87.0	(3.0)	88.5	(3.0)	72.4	(3.4)	40.4	(3.4)	33.1	(2.6)	66.2	(3.8)
Estland	96.1	(2.0)	99.7	(2.7)	90.7	(1.6)	97.7	(1.2)	51.2	(2.0)	86.2	(2.0)
Finnland	88.7	(3.4)	95.1	(4.3)	56.6	(3.1)	65.9	(2.8)	10.0	(1.5)	62.0	(3.7)
Frankreich	79.3	(3.6)	77.7	(3.3)	56.7	(2.3)	72.4	(3.4)	16.5	(2.0)	55.3	(3.2)
Griechenland	80.7	(3.5)	80.7	(3.6)	20.8	(2.2)	88.1	(3.4)	86.7	(2.8)	84.9	(3.5)
Irland	94.1	(4.8)	100.0	(0.0)	95.2	(2.9)	82.9	(2.7)	75.6	(3.3)	81.2	(3.2)
Island	98.3	(0.3)	99.9	(0.2)	93.1	(0.2)	22.0	(0.2)	39.0	(0.2)	52.3	(0.3)
Israel	99.5	(3.3)	95.2	(4.4)	87.6	(3.2)	96.6	(3.3)	60.3	(3.4)	82.7	(4.2)
Italien	80.2	(3.2)	95.5	(3.2)	38.9	(3.0)	30.1	(2.5)	12.6	(1.6)	44.0	(2.3)
Japan	80.8	(3.5)	98.2	(2.5)	75.5	(2.8)	82.6	(2.8)	15.4	(1.9)	45.1	(2.8)
Kanada	86.1	(3.0)	85.8	(3.0)	63.7	(3.5)	88.1	(2.6)	68.5	(2.6)	72.0	(2.9)
Korea	97.7	(3.9)	99.5	(3.7)	86.3	(3.4)	95.0	(2.7)	73.1	(3.5)	84.2	(3.5)
Lettland	99.7	(3.2)	100.0	(0.0)	95.9	(1.5)	79.9	(2.4)	39.4	(2.2)	76.6	(2.9)
Luxemburg	78.0	(0.1)	74.7	(0.1)	96.2	(0.1)	81.2	(0.1)	43.5	(0.1)	58.6	(0.1)
Mexiko	95.3	(3.3)	86.0	(3.3)	73.9	(2.3)	62.0	(3.1)	57.9	(3.0)	65.6	(3.2)
Neuseeland	97.6	(4.5)	99.3	(4.3)	96.7	(1.8)	97.3	(2.6)	78.2	(2.6)	77.1	(3.5)
Niederlande	88.5	(4.6)	91.8	(3.8)	85.6	(4.7)	89.4	(3.1)	57.5	(3.8)	37.0	(3.2)
Norwegen	85.5	(3.6)	98.5	(4.0)	63.9	(2.7)	92.1	(2.6)	77.8	(3.4)	18.8	(1.9)
Österreich	82.7	(3.4)	89.3	(3.3)	40.6	(2.8)	75.0	(2.8)	61.9	(3.2)	55.5	(3.3)
Polen	98.0	(3.9)	100.0	(3.8)	92.1	(1.5)	94.8	(1.8)	47.6	(2.6)	57.0	(2.8)
Portugal	86.6	(3.7)	99.7	(3.8)	97.5	(1.6)	84.0	(3.4)	36.0	(2.7)	50.9	(2.9)
Schweden	92.5	(4.0)	98.2	(4.0)	68.3	(2.6)	79.5	(2.8)	31.6	(2.4)	35.1	(2.8)
Schweiz	71.1	(4.0)	85.3	(3.7)	68.7	(2.3)	76.2	(2.9)	26.6	(2.6)	44.5	(3.3)
Slowakische Republik	99.6	(2.9)	97.0	(3.2)	62.4	(2.7)	99.3	(2.3)	62.5	(2.3)	83.3	(3.1)
Slowenien	99.8	(0.6)	98.2	(0.5)	46.7	(0.5)	82.3	(0.6)	32.0	(0.4)	65.0	(0.5)
Spanien	90.1	(3.3)	87.7	(3.1)	73.9	(2.7)	40.7	(2.8)	27.2	(1.9)	39.2	(2.3)
Tschechische Republik	96.2	(3.2)	96.7	(2.2)	61.2	(2.8)	95.6	(0.7)	27.8	(1.5)	87.2	(3.1)
Türkei	96.1	(3.3)	93.6	(3.5)	78.8	(3.1)	65.7	(3.1)	48.8	(3.0)	75.2	(4.3)
Ungarn	100.0	(0.0)	90.4	(3.4)	74.7	(3.0)	81.6	(3.6)	19.1	(1.9)	49.2	(3.1)
Vereinigte Staaten	97.4	(3.8)	97.7	(3.4)	84.9	(3.3)	95.8	(4.2)	66.8	(3.8)	85.9	(4.2)
Vereinigtes Königreich	99.9	(3.6)	100.0	(0.0)	96.9	(3.3)	98.3	(2.1)	83.9	(3.2)	83.7	(3.0)
OECD-Durchschnitt	91.2	(0.4)	93.2	(0.3)	74.6	(0.5)	78.2	(0.4)	47.8	(0.5)	63.3	(0.5)

© Waxmann Verlag GmbH

wären (OECD-Durchschnitt: 49 Prozent). Externe Evaluation als Werkzeug der Schulentwicklung kommt sowohl im OECD-Durchschnitt (75 Prozent der Schulen) als auch in Deutschland (72 Prozent) seltener zum Einsatz als die interne Evaluation. In Deutschland ebenfalls deutlich häufiger verwendet als in der letzten PISA-Erhebungsrunde wird die Umsetzung standardisierter Vereinbarungen für den Unterricht (55 Prozent für Mathematik in PISA 2012 im Vergleich zu 66 Prozent für Naturwissenschaften in PISA 2015). Hinter diesem fachspezifischen Indikator verbirgt sich eine eigene Strategie der Schule für eine interne Qualitätsentwicklung des Unterrichts. Der Fragebogen, der in PISA den Schulleitungen vorgelegt wird, spezifiziert hier beispielsweise ein Schulcurriculum, den Austausch von Unterrichtsmaterialien im Fachkollegium oder auch begleitende Weiterbildungsmaßnahmen. Solche Maßnahmen wurden in Deutschland unter anderem im Zuge der Implementation von Bildungsstandards vorgeschlagen (vgl. etwa Klieme et al., 2007). Deutschland nutzt diese Strategie zur schulinternen Qualitätsentwicklung deutlich häufiger als die deutschsprachigen Nachbarländer Schweiz, Österreich und Luxemburg.

Nur etwa halb so verbreitet wie im Durchschnitt der OECD-Staaten ist in Deutschland das Mentorat für Lehrkräfte. Während 78 Prozent der Schulleitungen in allen OECD-Staaten dieses Instrument an ihrer Schule nutzen, geben dies in Deutschland lediglich 40 Prozent an (in PISA 2012: 33 Prozent). Obwohl das Mentorat als Erleichterung des Berufseinstiegs für junge Lehrkräfte damit 2015 bereits häufiger zum Zug kommt als 2012, ist der Anteil der Schulen, die mit diesem Instrument arbeiten, im internationalen Vergleich gering. Dass die Begleitung junger Lehrkräfte durch erfahrene Kolleginnen und Kollegen in Deutschland bei Weitem nicht flächendeckend angeboten wird, hängt möglicherweise unter anderem mit der Gliederung der Lehrerbildung in Deutschland in eine erste und eine zweite Phase zusammen, sodass Mentorate häufig bereits zur Referendariatszeit gehören und der Mehrwert eines Mentorats zum Berufseinstieg gegebenenfalls nicht auf der Hand liegt.

Die deutlichste Veränderung in Bezug auf den Einsatz von Instrumenten der Schulentwicklung in Deutschland ist bei den Expertengesprächen zu verzeichnen. Gemeint ist hier in PISA die Begleitung einer Schule durch externe Beraterinnen und Berater über mindestens ein halbes Jahr hinweg. Während sich die Verbreitung dieser Expertengespräche im OECD-Durchschnitt seit PISA 2012 (43 Prozent) kaum verändert hat (2015: 48 Prozent), ist der Anteil der Schulen in Deutschland, die sich von externen Experten für mindestens 6 Monate beraten und begleiten lassen, merklich angestiegen: von 19 Prozent in PISA 2012 auf 33 Prozent in PISA 2015. Damit hat sich die Nutzungshäufigkeit dieses Instruments der Schulentwicklung unseren Nachbarstaaten angenähert.

In Deutschland unterscheiden sich die Gymnasien von den übrigen Schularten kaum in Bezug darauf, wie häufig sie die genannten Instrumente zur Qualitätssicherung und Schulentwicklung nutzen. Einzig beim kollegialen Mentoring für Junglehrkräfte zeigen sich Unterschiede: Nicht gymnasiale Schularten (45 Prozent) setzen dieses Instrument deutlich häufiger ein als Gymnasien (32 Prozent).

© Waxmann Verlag GmbH

Insgesamt lässt sich festhalten, dass sich Deutschland im internationalen Vergleich mittlerweile an die Gepflogenheiten zur Nutzungspraxis von Instrumenten zur Qualitätsentwicklung und -sicherung von Schulen angepasst hat. Das Engagement der Schulen, auch freiwillig Evaluationen durchzuführen, ist dabei überdurchschnittlich groß. Ein Aspekt der Qualitätsentwicklung von Schulen ist sicherlich auch das Führungshandeln der Schulleitungen, insbesondere aus der Sicht des Lehrerkollegiums. An Schulen der Sekundarstufe in Deutschland zeichnet sich aus dieser Perspektive ein konstruktives, unterstützendes Bild, das sich zwischen Gymnasien und nicht gymnasialen Schularten nicht nennenswert unterscheidet: Die Lehrkräfte beurteilen das Führungsverhalten ihrer Schulleitung überwiegend positiv. Die Mehrheit fühlt sich von ihrer Schulleitung als fachlich kompetent behandelt und in Entscheidungsfindungen einbezogen. Von einer Konsensorientierung der Schulleitung bei der Festlegung von Prioritäten und Zielen berichtet ebenfalls ein Großteil der Lehrkräfte.

Für Deutschland wird ersichtlich, dass Schulleitungen nach internen Evaluationen meist auf mehreren Ebenen tätig werden. Am häufigsten arbeiten die Schulleitungen nach eigenen Angaben an Verbesserungen des Lehrens und Lernens (76 Prozent), aber auch Maßnahmen zur Lehrerfortbildung, zur Verbesserung der Leistungen der Schülerschaft und ihrer lehrplanübergreifenden Kompetenzen (jeweils etwa 60 Prozent) sind den Schulleitungen wichtig. Im internationalen Vergleich geht es den Schulleitungen in Deutschland dabei etwas seltener als im Durchschnitt um die fachbezogenen Leistungen und etwas überdurchschnittlich häufig um die fächerübergreifenden Kompetenzen (z. B. Lernstrategien, Zeitmanagement). Maßnahmen zur Erhöhung der Chancengerechtigkeit stehen im Vergleich zum OECD-Durchschnitt weniger im Fokus; hier gibt die Hälfte aller Schulleitungen an, dass dieses Thema nicht akut sei, da die Ergebnisse der Evaluation hier zufriedenstellend gewesen seien.

5.2.3 Schulische Angebote, Ressourcen und Lerngelegenheiten für die Naturwissenschaften

Die naturwissenschaftliche Kompetenz fünfzehnjähriger Schülerinnen und Schüler wird in PISA 2015 als Schwerpunkt erhoben. Von besonderem Interesse bei der Betrachtung der Einzelschule ist daher, inwieweit diese den Jugendlichen über den regulären Unterricht hinaus Gelegenheiten bietet, sich mit Naturwissenschaften auseinanderzusetzen und naturwissenschaftliche Kompetenz zu entwickeln.

Neben dem üblichen Curriculum haben Schülerinnen und Schüler der OECD-Staaten oftmals die Möglichkeit, an ihrer Schule Zeit mit Zusatzunterricht oder (eigenständigem oder angeleitetem) Lernen in den Naturwissenschaften zu verbringen. In welchem Umfang, in welcher Form und zu welchen Themen solcher Zusatzunterricht angeboten wird, unterscheidet sich von Schule zu Schule. Angebote zur fachlichen Beschäftigung mit Naturwissenschaften außerhalb des regulären Unterrichts wie etwa Arbeitsgemeinschaften (AGs) oder Wettbewerbe sind in den OECD-Staaten recht unterschiedlich weit

© Waxmann Verlag GmbH

verbreitet. Im Mittel besuchen knapp 40 Prozent der Fünfzehnjährigen eine Schule, an der AGs mit naturwissenschaftlichem Schwerpunkt angeboten werden, und 67 Prozent berichten, dass sie an oder mit ihrer Schule an Wettbewerben zu naturwissenschaftlichen Themen teilnehmen können. Unter den OECD-Ländern gibt es mehrere, in denen naturwissenschaftliche AGs unüblich sind, insbesondere Österreich (5 Prozent), Schweden (7 Prozent) oder auch Finnland (12 Prozent). Hingegen sind sie etwa in Polen nahezu selbstverständlich (etwa 80 Prozent). In Deutschland steht knapp der Hälfte der Fünfzehnjährigen (48 Prozent) eine AG zu naturwissenschaftlichen Themen an ihrer Schule offen. Sehr niedrig ist der Anteil von Schülerinnen und Schülern, deren Schule eine AG zu naturwissenschaftlichen Themen anbietet, in der Schweiz (24 Prozent) sowie in Österreich (31 Prozent), während er in Luxemburg (81 Prozent), Finnland (86 Prozent), Estland (95 Prozent) und in Polen (95 Prozent) sehr hoch ist. In Deutschland liegt der Anteil mit 59 Prozent im Bereich des OECD-Durchschnitts. Dabei ist zu beachten, dass sowohl in der Gruppe der Staaten mit einem relativ geringen Angebot für Wettbewerbe in Naturwissenschaften als auch in derjenigen mit einem hohen Angebot die Mittelwerte der naturwissenschaftlichen Kompetenz unterschiedlich hoch ausgeprägt sind.

Innerhalb Deutschlands unterscheiden sich die Gymnasien hier erwartungsgemäß markant von den nicht gymnasialen Schularten: Während 73 Prozent der Gymnasien AGs mit naturwissenschaftlichem Bezug anbieten, ist dies lediglich an 37 Prozent der nicht gymnasialen Schularten der Fall. Noch deutlicher ist der Unterschied mit Blick auf die Teilnahme an naturwissenschaftlichen Wettbewerben: So gut wie alle Gymnasien in Deutschland bieten diese Möglichkeit an (97 Prozent), jedoch nur 40 Prozent der nicht gymnasialen Schularten.

Unterricht und weiterführende Angebote in naturwissenschaftlichen Fächern finden meist in speziell ausgestatteten Fachräumen statt. Labore für Biologie und Chemie, Vorbereitungs- und Geräteräume für Physik oder Kartenräume für Erdkunde gehören an vielen Schulen zur Ausstattung. Im internationalen Vergleich schildern die Schulleitungen in PISA 2015 jedoch auch eine stark unterschiedliche Verfügbarkeit materieller Ressourcen, die sie an ihrer Schule für den Unterricht in naturwissenschaftlichen Fächern zur Verfügung haben. Insgesamt schätzen die Schulleitungen in Deutschland diese Ressourcen überwiegend positiv ein. Besonders stark vom OECD-Durchschnitt weichen die Schulen in Deutschland bei zwei dieser Ressourcen ab: Zum einen geben signifikant mehr Schulleitungen in Deutschland als im OECD-Mittel an, dass zusätzliche Mittel oft für die Verbesserung des naturwissenschaftlichen Unterrichts verwendet werden. Zum anderen hat so gut wie keine Schule in Deutschland die Möglichkeit beziehungsweise scheint es unüblich zu sein, auf Laborantinnen oder Laboranten zurückzugreifen, die den naturwissenschaftlichen Unterricht unterstützen können. Im OECD-Durchschnitt kann dies immerhin jede dritte Schule (34 Prozent) tun.

In Deutschland beurteilen die Schulleitungen an Gymnasien die materielle Ausstattung für den fachspezifischen naturwissenschaftlichen Unterricht an ihrer Schule insgesamt positiver als die Schulleitung an anderen Schularten: So verfügen 61 Prozent der Gymnasien über vergleichsweise gut ausgestattete Laborräume, während dies an nicht

© Waxmann Verlag GmbH

gymnasialen Schularten lediglich auf 41 Prozent zutrifft. Auch in Bezug auf die Unterrichtsmaterialien zeigt sich dieser Unterschied. Während 88 Prozent der Gymnasien über ausreichend Material für die regelmäßige Nutzung durch alle Klassen verfügen, sind es an nicht gymnasialen Schularten nur 72 Prozent. Hingegen unterscheiden sich Gymnasien und nicht gymnasiale Schularten kaum hinsichtlich der Einschätzung der Qualifikation von Lehrkräften für naturwissenschaftliche Fächer an der eigenen Schule. In allen Schularten stimmt etwa die Hälfte der Schulleitungen zu, dass die Lehrkräfte für naturwissenschaftliche Fächer zu den am besten ausgebildeten Kolleginnen und Kollegen zählen.

5.2.4 Schülerverhalten und Schulklima

Im institutionellen Rahmen verfügen Schulen über Gestaltungsspielräume, wie sie ihren Bildungsauftrag umsetzen und sich als Lern- und Lebensort profilieren (Fend, 1988; Fend, 2006). Das Schulklima als Kultur des Umgehens miteinander wird auch im Rahmen solcher Spielräume geprägt. Dabei hängt ein Schulklima, das als angenehm empfunden wird, mit einer ganzen Reihe erstrebenswerter psychosozialer Entwicklungen und häufig auch mit Lernerfolg zusammen (vgl. etwa Grob, 2007; Jäger, 2012; van Ackeren et al., 2015). Beeinträchtigungen des Schulklimas sind daher eine besondere Herausforderung für Schulen, zumal sie nicht nur das Miteinander und die Identifikation mit der Schule, sondern auch das Lernen gefährden. Dazu gehören unter anderem unentschuldigtes Fernbleiben vom Unterricht oder das Schikanieren und Einschüchtern von Mitschülerinnen und Mitschülern.

Beeinträchtigungen des Lernens

Der Fragebogen für die Schulleitungen in PISA 2015 griff auch das Thema auf, wie oft Schülerinnen und Schüler unentschuldigt im Unterricht fehlen (während des ganzen Schultages oder in einzelnen Stunden), zu spät zum Unterricht erscheinen, mangelnden Respekt gegenüber Lehrkräften zeigen, den Unterricht stören, Alkohol bzw. Drogen konsumieren oder Klassenkameradinnen und -kameraden schikanieren und einschüchtern. Die Befunde zu diesen Aspekten sind mit einer gewissen Vorsicht zu interpretieren, da Effekte der sozialen Erwünschtheit bei der Beantwortung solcher Fragen nicht auszuschließen sind. Ähnlich wie in der letzten PISA-Erhebungsrunde 2012 bereiten aus der Sicht der Schulleiterinnen und Schulleiter in den PISA-Teilnehmerstaaten Schulschwänzen und Unterrichtsstörungen größere Probleme für das Lernen als der Konsum von Alkohol und Drogen sowie Einschüchterung oder mangelnder Respekt gegenüber Lehrkräften.

Insgesamt lässt sich über die OECD-Staaten sagen, dass die überwiegende Mehrheit der Schulleitungen das Lernen an ihrer Schule „nicht“ oder „nur wenig“ durch die genannten Faktoren beeinträchtigt sieht. Jeweils ungefähr 90 Prozent der Schülerin-

© Waxmann Verlag GmbH

nen und Schüler in den OECD-Staaten besuchen eine Schule, deren Schulleitung keine oder sehr wenige Probleme durch Alkohol- und Drogenkonsum oder das Schikanieren von Mitschülerinnen und Mitschülern angibt. Es gibt jedoch auch einige Staaten, deren Schulleitungen in größerem Umfang (20 Prozent und mehr „bis zu einem gewissen Ausmaß“ oder „sehr“) davon berichten, dass das Lernen an ihrer Schule unter dem Schikanieren und Einschüchtern von Mitschülerinnen und Mitschülern leidet. Darunter befindet sich auch Deutschland (20 Prozent). An Gymnasien ist dieses Problem ähnlich stark verbreitet (20 Prozent) wie an nicht gymnasialen Schularten (23 Prozent). Bemerkenswert ist, dass dies eine Reihe von Staaten betrifft, die im internationalen Leistungsvergleich sehr erfolgreich sind, wie Finnland (23 Prozent) oder die Niederlande (35 Prozent). Durch mangelnden Respekt gegenüber Lehrpersonen wird das Lernen aus Sicht der Schulleitungen an etwa 20 Prozent der Schulen in den OECD-Staaten beeinträchtigt.

Schulversäumnisse bei Jugendlichen und Lehrkräften

Als spürbar stärker hinderlich für das Lernen an ihrer Schule empfinden die Schulleitungen in den OECD-Staaten das unerlaubte Fernbleiben von einzelnen Schulstunden (36 Prozent) oder ganzer Schultage (34 Prozent). Für die Schulleitungen in Deutschland ist das unerlaubte Fernbleiben vom Unterricht deutlich seltener ein Problem (21 Prozent), wenn auch etwas häufiger als noch 2012 (17 Prozent). Auch in Bezug auf diese Beeinträchtigung findet sich in einigen sehr leistungsstarken Bildungssystemen eine auffallend hohe Quote an Schulen, deren Leitungen das Lernen und Schulklima gefährdet sehen. Dies mag daran liegen, dass in leistungsstärkeren Schulsystemen bestimmte Grundvoraussetzungen für Unterricht stärkere Aufmerksamkeit erfahren. Auffallend hohe Quoten von Schulleitungen, die das Lernen und Schulklima an ihrer Schule durch Absentismus gefährdet sehen, finden sich in Finnland (44 Prozent), Österreich (49 Prozent), Slowenien (53 Prozent) und Kanada (56 Prozent). Diese Verteilung der Einschätzung durch die Schulleitung ergibt ein sehr ähnliches Bild wie in PISA 2012, wobei der Anteil der Schulen, an denen die Schulleitung das Lernen durch Schülerabsenzen beeinträchtigt sieht, insgesamt leicht gestiegen ist.

Insgesamt scheint das unerlaubte Versäumen von Unterricht in der Einschätzung der Schulleitungen zwar kein flächendeckendes Problem zu sein, gemessen an anderen Beeinträchtigungen ist es jedoch ein sehr relevantes Thema. Meist wird in Bezug auf Schulversäumnisse die Aufmerksamkeit auf die Schülerinnen und Schüler gerichtet, die – im Alter von fünfzehn Jahren noch schulpflichtig – einzelne Stunden oder auch ganze Tage den Unterricht schwänzen. Aus der Perspektive der Schulleitungen rückt hier jedoch auch eine zweite Personengruppe in das Blickfeld, ohne die ebenfalls kein Unterricht stattfindet – die Lehrkräfte. Eine wichtige Funktion der Schulpflicht ist das Schaffen von Voraussetzungen, dass alle Kinder und Jugendlichen in einem Bildungssystem bestimmte Lernbedingungen vorfinden und damit die Chance haben, als relevant betrachtete Bildungsziele zu erreichen. Schülerinnen und Schüler, die dem Unterricht fernbleiben, verzichten auf die Gelegenheit, sich mit Themen des Lehrplans und

© Waxmann Verlag GmbH

mit ihrer schulischen Umwelt auseinanderzusetzen und unter Anleitung von Lehrkräften zu lernen. Lehrkräfte, die nicht zum Dienst erscheinen, stellen die Schulleitung vor eine schwierige Aufgabe. Ohne sie kann der planmäßige Unterricht nicht stattfinden, eine Vertretung muss unter Umständen sehr rasch organisiert werden, und nicht immer kann der Unterricht dann im vorgesehenen Fach erfolgen.

Schulversäumnisse in einzelnen Fächern oder auch während ganzer Schultage gehören zu den Ereignissen, die von den Schulleitungen in allen OECD-Staaten als besonders problematisch für gelingendes Lernen und ein konstruktives Schulklima betrachtet werden. Tabelle 5.9 gibt einen Überblick über die Einschätzung der Schulleitungen in den OECD-Staaten, inwieweit an ihrer Schule das Lernen und das Schulklima durch absente Schülerinnen und Schüler sowie Lehrkräfte beeinträchtigt ist. Durch Schülerabsenzen beeinträchtigt sieht etwa jede dritte Schulleitung in den OECD-Staaten das Lernen an ihrer Schule. In Deutschland ist dieser Anteil mit 21 Prozent der Schulleitungen wesentlich geringer. Anders sieht es bei Beeinträchtigungen durch Absenzen von Lehrkräften aus, die aus der Perspektive der Schulleitungen in Deutschland häufiger ein Problem darstellen (40 Prozent) als im Mittel der OECD-Staaten (20 Prozent).

Dass sich das absichtliche Versäumen von Unterricht tatsächlich negativ auf den Kompetenzerwerb auswirken kann, legen vertiefende Analysen der Daten aus PISA 2012 für Deutschland nahe. Damals wurde in Deutschland in einer nationalen Ergänzung des Schülerfragebogens gezielt nach fachspezifischem Schwänzen von Unterricht gefragt, sodass Unterrichtsversäumnisse in den „PISA-Fächern“ Mathematik, Naturwissenschaften (Biologie, Chemie, Physik) sowie in Deutsch mit der durchschnittlichen Kompetenz in den entsprechenden PISA-Domänen in Verbindung gebracht werden konnten (Sälzer & Heine, 2016). Die Befunde deuten darauf hin, dass unter Kontrolle mehrerer Hintergrundmerkmale das Schwänzen insbesondere im Bereich Mathematik und Naturwissenschaften mit Einbußen von bis zu 40 Punkten gegenüber Mitschülerinnen und Mitschülern, die lückenlos am Unterricht teilnehmen, einhergehen kann.

Mit Blick auf das Lehrerkollegium an der eigenen Schule sehen die Schulleitungen – über alle OECD-Staaten hinweg – an etwa jeder fünften Schule das Lernen durch Absentismus der Lehrkräfte gefährdet. Auch in diesem Punkt finden sich bemerkenswerte Unterschiede im Vergleich der OECD-Staaten, wobei sich kein Zusammenhang mit der Leistungsstärke der jeweiligen PISA-Kohorte finden lässt. Sehr niedrig ist der Anteil der Schulen, deren Leitung das Lernen durch abwesende Lehrkräfte beeinträchtigt sieht, in leistungsstarken Bildungssystemen wie Estland, Finnland, Polen und der Schweiz, aber auch in leistungsschwächeren Staaten wie Griechenland, Lettland oder Österreich. In Deutschland geben vier von zehn Schulleiterinnen und Schulleitern an, dass Absentismus bei Lehrkräften an ihrer Schule das Schulklima beeinträchtigt (40 Prozent), wobei es keinen Unterschied zwischen Gymnasien und nicht gymnasialen Schularten gibt. Unter den OECD-Staaten ist dieser Wert in Deutschland insgesamt am höchsten. Ähnlich hoch ist der Anteil von Schulleitungen, die das Lernen an ihrer Schule durch Absenzen von Lehrkräften beeinträchtigt sehen mit mehr als 30 Prozent in Belgien, Chile, Israel, den Niederlanden und Norwegen.

© Waxmann Verlag GmbH

Tabelle 5.9: Beeinträchtigungen des Lernens durch Absenzen von Schülerinnen und Schülern und Lehrkräften

	Beeinträchtigung des Lernens durch Schülerabsenzen				Beeinträchtigung des Lernens durch Lehrerabsenzen			
	Nicht oder wenig		Bis zu einem gewissen Grad oder sehr		Nicht oder wenig		Bis zu einem gewissen Grad oder sehr	
OECD-Staaten	**%**	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
Australien	72.5	(1.8)	27.5	(1.0)	82.8	(1.7)	17.1	(1.0)
Belgien	70.5	(2.7)	29.5	(2.1)	64.3	(2.6)	35.6	(2.1)
Chile	81.7	(3.3)	18.3	(2.0)	64.7	(3.1)	35.3	(3.3)
Dänemark	64.1	(2.8)	35.8	(2.0)	72.7	(3.2)	25.9	(3.1)
Deutschland	77.0	(2.9)	21.3	(2.9)	59.9	(3.0)	40.1	(2.8)
Estland	63.5	(2.4)	36.6	(1.8)	89.9	(2.7)	10.1	(1.2)
Finnland	56.3	(2.8)	43.8	(2.7)	84.1	(3.0)	15.9	(3.1)
Frankreich	53.8	(2.7)	46.2	(2.7)	77.7	(3.1)	21.9	(3.0)
Griechenland	73.6	(2.9)	26.4	(3.3)	92.7	(4.0)	5.6	(1.8)
Irland	48.6	(2.9)	51.4	(3.0)	88.6	(3.5)	10.6	(2.3)
Island	81.9	(0.3)	17.5	(0.2)	86.6	(0.3)	13.4	(0.2)
Israel	51.0	(2.9)	48.9	(3.3)	66.7	(3.3)	33.3	(2.8)
Italien	64.2	(2.7)	35.9	(2.2)	86.6	(3.9)	12.6	(2.5)
Japan	86.2	(2.9)	11.8	(2.3)	91.1	(3.0)	8.9	(2.0)
Kanada	44.5	(2.0)	55.6	(2.3)	90.4	(2.6)	9.2	(1.8)
Korea	76.1	(3.5)	23.9	(2.4)	97.9	(2.5)	0.0	(0.0)
Lettland	55.7	(2.0)	44.2	(1.8)	95.4	(2.8)	4.0	(1.2)
Luxemburg	41.6	(0.1)	46.5	(0.1)	86.3	(0.0)	0.0	(0.0)
Mexiko	52.4	(2.3)	47.6	(3.2)	85.9	(2.9)	14.2	(1.6)
Neuseeland	58.9	(2.5)	41.1	(2.6)	94.3	(3.8)	5.7	(1.9)
Niederlande	71.6	(3.5)	24.8	(4.2)	59.1	(4.2)	36.9	(4.9)
Norwegen	80.1	(3.2)	19.3	(3.3)	64.8	(3.1)	34.1	(3.6)
Österreich	51.4	(2.5)	48.6	(2.8)	83.3	(3.3)	14.9	(2.4)
Polen	73.4	(3.3)	26.5	(3.4)	90.0	(3.7)	10.1	(2.1)
Portugal	59.0	(2.8)	41.1	(2.9)	88.8	(3.6)	10.5	(2.4)
Schweden	72.6	(3.0)	25.8	(3.4)	81.4	(3.3)	18.0	(2.8)
Schweiz	72.5	(3.1)	25.7	-(3.3)	91.1	(3.7)	8.9	(2.4)
Slowakische Republik	68.3	(2.9)	31.8	(2.0)	93.5	(3.5)	6.1	(1.6)
Slowenien	47.1	(0.4)	52.9	(0.4)	82.5	(0.4)	17.1	(0.4)
Spanien	73.4	(2.5)	26.7	(2.2)	96.3	(3.7)	3.7	(1.4)
Tschechische Republik	76.4	(2.5)	23.6	(2.8)	87.5	(3.0)	12.2	(2.2)
Türkei	47.3	(4.0)	50.6	(3.2)	93.1	(3.1)	6.9	(1.9)
Ungarn	77.0	(3.3)	23.1	(2.1)	93.4	(3.5)	6.5	(1.8)
Vereinigte Staaten	54.1	(2.7)	46.0	(2.5)	83.0	(3.9)	16.5	(3.3)
Vereinigtes Königreich	90.2	(3.4)	9.2	(2.1)	75.9	(3.0)	22.7	(3.5)
OECD-Durchschnitt	65.4	(2.6)	35.7	(2.2)	81.3	(3.0)	19.7	(2.1)

© Waxmann Verlag GmbH

Gefühl der Zugehörigkeit zur eigenen Schule

Das Gefühl der Zugehörigkeit zur eigenen Schule, in PISA *Sense of Belonging* genannt, ist ein Merkmal, das im Rahmen der PISA-Studien regelmäßig erfasst wird. Die Schülerinnen und Schüler in Deutschland hatten bereits bei PISA 2012 ein überdurchschnittlich hohes Gefühl der Zugehörigkeit zur eigenen Schule berichtet. Die in PISA 2015 untersuchte Kohorte bestätigt dieses vergleichsweise hohe subjektive Zugehörigkeitsgefühl zur eigenen Schule. Für Deutschland besonders interessant ist – auch im Hinblick auf die sich verändernde Sekundarschullandschaft –, wie sich das Gefühl der Zugehörigkeit zwischen den verschiedenen Schularten (Gymnasium und nicht gymnasiale Schularten) unterscheidet. In der Tat findet sich eine signifikante Differenz zwischen dem Gefühl der Zugehörigkeit zur eigenen Schule an Gymnasien und nicht gymnasialen Schularten ($M_{Gymnasium}$ = .38; SE_{Gym} = .03; $M_{nicht\ gymnasiale\ Schulart}$ = .20; SE_{ngym} = .03). Demnach fühlen sich Schülerinnen und Schüler an Gymnasien subjektiv stärker mit ihrer Schule verbunden als Schülerinnen und Schüler anderer Schularten.

5.2.5 Zusammenfassung und Diskussion

Die schulischen Rahmenbedingungen der Kompetenzentwicklung wurden in diesem Kapitel aus zwei Perspektiven betrachtet. Schule wurde dabei einerseits als Institution innerhalb gegebener Rahmenbedingungen und andererseits als Einzelschule mit bestimmten Merkmalen und Gestaltungsspielräumen gesehen. Im ersten Teil des Kapitels wurden Ergebnisse aus PISA 2015 in Bezug auf die Gliederung verschiedener Bildungssysteme in der OECD berichtet und dabei festgestellt, dass die Unterteilung in ein- und mehrgliedrige Bildungssysteme allein noch keinen zuverlässigen Hinweis auf die naturwissenschaftliche Kompetenz fünfzehnjähriger Schülerinnen und Schüler gibt. Viel aufschlussreicher ist für die Abbildung der Schulstruktur und möglicher Wirkungen bei mehrgliedrigen Systemen wie in Deutschland ein Vergleich von Schulen innerhalb einer Schulart. So lässt sich beispielsweise untersuchen, inwieweit sich Gymnasien hinsichtlich einzelner Merkmale voneinander unterscheiden.

Aus der Perspektive auf Schulen als Institutionen weichen Schulen in Deutschland nur teilweise von den Mittelwerten der OECD-Staaten ab. So wurde beispielsweise in Bezug auf die Altersstruktur und die Klassenstufe der in PISA untersuchten Schülerinnen und Schüler deutlich, dass Fünfzehnjährige in Deutschland noch immer meist auf niedrigere Klassenstufen verteilt sind als in zahlreichen anderen Schulsystemen der OECD. Dazu trägt unter anderem der relativ hohe Anteil an verzögerten Schulkarrieren bei, die sich überwiegend aus vergleichsweise häufigen Klassenwiederholungen ergeben. Sehr variabel sind im internationalen Vergleich auch die Zeitanteile, die in der Sekundarstufe für den Unterricht in Naturwissenschaften vorgesehen sind. Hier liegt die reine Stundenzahl für den naturwissenschaftlichen Unterricht in Deutschland allerdings nur

© Waxmann Verlag GmbH

minimal über dem OECD-Mittelwert, für Mathematik und Deutsch nur minimal darunter.

Die Frage, inwieweit die Gliederung eines Schulsystems mit dem mittleren Kompetenzniveau und auch mit sozialen Disparitäten zusammenhängen, ist in der öffentlichen Diskussion seit PISA 2000 ein vielfach aufgegriffenes Thema. In der Tat zeigt eine gemeinsame Betrachtung der bisherigen PISA-Erhebungsrunden, dass soziale Unterschiede in Bildungsbeteiligung und Bildungserfolg in stark und vor allem starr gegliederten Bildungssystemen stärker ausgeprägt sind als in eingliedrigen Systemen. Eingliedrige Bildungssysteme bedeuten allerdings nicht automatisch eine niedrigere Kopplung sozialer Hintergrundmerkmale mit der Leistung. Ein vertiefender Blick in einzelne PISA-Erhebungsrunden und die Entwicklung über die letzten 15 Jahre zeigt allerdings, dass die Strukturmerkmale der Bildungssysteme weder die durchschnittliche Kompetenz der Schülerinnen und Schüler noch die Entwicklung dieses Leistungsniveaus über die Zeit erklären können.

Der Blick auf die Einzelschule als Ort gemeinsamen Lernens und Lebens brachte wenig Auffälligkeiten hervor. Die den Schulen zur Verfügung stehenden Ressourcen scheinen aus der Perspektive der Schulleitungen weitgehend akzeptabel zu sein, die Klassengrößen gemessen am OECD-Mittelwert handhabbar. Unter den besonderen Herausforderungen für Schulen in Form von Verhaltensweisen, die das Lernen beeinträchtigen können, schätzen die Schulleitungen in Deutschland besonders Unterrichtsversäumnisse bei Schülerinnen und Schülern, aber auch bei Lehrkräften als schwerwiegend an ihrer Schule ein. Solche Perspektiven und Aspekte sind ein wichtiger Teil der Erkenntnisse aus PISA, weil sie ebenso wie etwa das Gefühl der Zugehörigkeit zur Schule empirische Anhaltspunkte dafür liefern, wie es um das Verhalten und die psychosoziale Entwicklung in Schulen steht. Aussagen dieser Art geben Auskunft über Qualitätsmerkmale und erzieherische Ergebnisse des schulischen Alltags, die anschlussfähig sind für vertiefende Studien.

Insgesamt setzt dieses Bild die Befunde der letzten PISA-Erhebung fort: Die Klassen sind mit 25 Schülerinnen und Schülern durchschnittlich groß, das Lernen wird nicht übermäßig beeinträchtigt und das subjektive Zugehörigkeitsgefühl der Schülerinnen und Schüler ist nach wie vor überdurchschnittlich groß. Bei den Absenzen fällt auf, dass die Schülerinnen und Schüler vergleichsweise selten schwänzen. Problematisch erscheinen eher spontane Abwesenheitsmeldungen von Lehrkräften, die es nach Auskunft der Schulleitungen an mehr als doppelt so vielen Schulen wie im OECD-Mittel gibt. Mit Blick auf weiterführende Forschung oder bildungspolitische Maßnahmen geben die Befunde aus PISA 2015 also wenig Anlass zur Sorge, bieten jedoch einige Anknüpfungspunkte, denen nachgegangen werden kann.

© Waxmann Verlag GmbH

Literatur

Alexander, K., Entwisle, D. & Dauber, S. (2003). *On the success of failure: a reassessment of the effects of retention in the early grades*. Cambridge: Cambridge University Press.

Arnhold, G. (2005). *Kleine Klassen – große Klasse? Eine empirische Studie zur Bedeutung der Klassengröße für Schule und Unterricht*. Bad Heilbrunn: Klinkhardt.

Autorengruppe Bildungsberichterstattung. (2016). *Bildung 2016 in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: Bertelsmann.

Baumert, J., Bos, W. & Watermann, R. (2000). Mathematische und naturwissenschaftliche Grundbildung im internationalen Vergleich. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *TIMSS/III. Dritte internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn* (S. 135–197). Opladen: Leske + Budrich.

Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W. & Weiß, M. (Hrsg.). (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.

Baumert, J. & Stanat, P. (2010). Internationale Schulleistungsvergleiche. In D. H. Rost (Hrsg.), *Handwörterbuch pädagogische Psychologie* (4. Aufl., S. 324–335). Weinheim: Beltz.

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.), *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung.* Zeitschrift für Pädagogik, 59. Beiheft (S. 172–199). Weinheim: Beltz.

Blömeke, S. & Herzig, B. (2009). Schule als gestaltete und zu gestaltende Institution – ein systematischer Überblick über aktuelle und historische Schultheorien. In S. Blömeke, T. Bohl, L. Haag, G. Lang-Wojtasik & W. Sacher (Hrsg.), *Handbuch Schule. Theorie – Organisation – Entwicklung* (S. 15–28). Bad Heilbrunn: Klinkhardt.

Boller, S., Möller, M. & Palowski, M. (2013). Wiederholen in der gymnasialen Oberstufe – wissenschaftliche Befunde und pädagogische Unterstützungsmöglichkeiten. In D. Bosse, F. Eberle & B. Schneider-Taylor (Hrsg.), *Standardisierung in der gymnasialen Oberstufe* (S. 175–188). Wiesbaden: Springer VS.

Bromme, R., Prenzel, M. & Jäger, M. (2014). Empirische Bildungsforschung und evidenzbasierte Bildungspolitik. Eine Analyse von Anforderungen an die Darstellung, Interpretation und Rezeption empirischer Befunde. *Zeitschrift für Erziehungswissenschaft, 17* (Sonderheft 27), 3–54.

Cicourel, A. & Kitsuse, J. I. (2012). Die soziale Organisation der Schule und abweichende jugendliche Karrieren. In U. Bauer, U. H. Bittlingmayer & A. Scherr (Hrsg.), *Handbuch Bildungs- und Erziehungssoziologie* (S. 137–149). Wiesbaden: Springer VS.

Cunningham, A. E. & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33,* 934–945.

Dreeben, R. & Barr, R. (1988). Classroom composition and the design of instruction. *Sociology of Education, 61,* 129–142.

Ehmke, T., Drechsel, B. & Carstensen, C. H. (2008). Klassenwiederholen in PISA-I-Plus: Was lernen Sitzenbleiber in Mathematik dazu? *Zeitschrift für Erziehungswissenschaft, 11* (3), 368–387. http://dx.doi.org/doi:10.1007/s11618-008-0033-3

Eickelmann, B., Gerick, J. & Koop, C. (2016). ICT use in mathematics lessons and the mathematics achievement of secondary school students by international comparison. Which

© Waxmann Verlag GmbH

role do school level factors play? *Education and Information Technologies.* http://dx.doi.org/10.1007/s10639-016-9498-5

Fend, H. (1988). Schulqualität. Die Wiederentdeckung der Schule als pädagogische Gestaltungsebene. *Neue Sammlung, 28* (4), 537–547.

Fend, H. (2006). *Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Fuchs, H.-W. (2003). Auf dem Weg zu einem Weltcurriculum? Zum Grundbildungskonzept von PISA und der Aufgabenzuweisung an die Schule. *Zeitschrift für Pädagogik, 49* (2), 161–179.

Grob, U. (2007). Schulklima und politische Sozialisation. Hat das perzipierte soziale Klima einen nachhaltigen Einfluss auf die Entwicklung von politischem Interesse und Toleranz? *Zeitschrift für Pädagogik, 53* (6), 774–799.

Hanßen, K. (2011). *Rechtliche Regelungen zu Tätigkeitsfeldern von Schulleiterinnen und Schulleitern bei erweiterter Eigenverantwortung von Schulen. Eine Untersuchung der Rechtslage in den Ländern Bayern, Hessen und Nordrhein-Westfalen.* Berlin. Zugriff am 17.10.2016. Verfügbar unter http://www.dipf.de/de/forschung/projekte/pdf/steubis/projekt-sharp-pdf/Endfassung_BY_HE_NW.pdf/view.

Hattie, J. (2008). *Visible learning. A synthesis of meta-analyses relating to achievement*. London: Routledge.

Helmke, A., Helmke, T., Schrader, F.-W., Wagner, W., Klieme, E., Nold, G. & Schröder, K. (2008). Wirksamkeit des Englischunterrichts. In E. Klieme, W. Eichler, A. Helmke, R. Lehmann, G. Nold, H.-G. Rolff, K. Schröder, G. Thomé & H. Willenberg (Hrsg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 382–397). Weinheim: Beltz.

Helmke, A. & Jäger, R. (2002). *Das Projekt MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext*. Landau: Verlag Empirische Pädagogik.

Helmke, A. & Schrader, F.-W. (2013). Angebots-Nutzungs-Modell. In M. A. Wirtz (Hrsg.), *Dorsch – Lexikon der Psychologie* (16. Aufl., S. 147–148). Bern: Huber.

Hoffmann, H. (2008). Sozialisationsinstanz Schule: Zwischen Erziehungsauftrag und Wissensvermittlung. In M. K. W. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion. Inhaltsfelder, Forschungsperspektiven und methodische Zugänge* (2., vollst. überarb. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Hong, G. & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis, 27* (3), 205–224.

Im Brahm, G. (2006). Klassengröße: Eine wichtige Variable von Schule und Unterricht? *bildungsforschung, 3* (1), 1–23.

Jäger, D. J. (2012). Schulklima, Selbstwirksamkeit und Arbeitszufriedenheit aus Sicht der Lehrpersonen und Schüler/-innen in Hessen und Bremen. In K. Maag Merki (Hrsg.), *Zentralabitur. Die längsschnittliche Analyse der Wirkungen der Einführung zentraler Abiturprüfungen in Deutschland* (S. 65–93). Wiesbaden: VS Verlag für Sozialwissenschaften.

Klemm, K. (2009). *Klassenwiederholungen – teuer und unwirksam.* Bertelsmann Stiftung.

Klieme, E. (2006). *Zusammenfassung zentraler Ergebnisse der DESI-Studie*, DIPF. Zugriff am 17.10.2016. Verfügbar unter http://www.dipf.de/de/forschung/projekte/pdf/biqua/DESI_Ausgewaehlte_Ergebnisse.pdf/view.

© Waxmann Verlag GmbH

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.

Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M., Reiss, K., Riquarts, K., Rost, J., Tenorth, H.-E. & Vollmer, H. J. (2007). *Zur Entwicklung nationaler Bildungsstandards. Expertise*. Bonn: BMBF.

Klieme, E. & Steinert, B. (2008). Schulentwicklung im Längsschnitt. Ein Forschungsprogramm und erste explorative Analysen. In M. Prenzel & J. Baumert (Hrsg.), *Vertiefende Analysen zu PISA 2006* (S. 221–238). Wiesbaden: VS Verlag für Sozialwissenschaften.

Kobarg, M., Prenzel, M., Seidel, T., Walker, M., McCrae, B., Cresswell, J. & Wittwer, J. (2011). *An international comparison of science teaching and learning. Further results from PISA 2006.* Münster: Waxmann.

Konsortium Bildungsberichterstattung. (2006). *Bildung in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: Bertelsmann.

Krohne, J. A., Meier, U. & Tillmann, K.-J. (2011). Sitzenbleiben, Geschlecht und Migration – Klassenwiederholungen im Spiegel der PISA-Daten. *Zeitschrift für Pädagogik, 50* (3), 373–391.

Lehrl, S., Smidt, W., Grosse, C. & Richter, D. (2013). Patterns of literacy and numeracy activities in preschool and their relation to structural characteristics and children's home activities. *Research Papers in Education, 29* (5), 577–597. http://dx.doi.org/10.1080/02671522.2013.792865

Meyer, H.-D. & Rowan, B. (2006). *The new institutionalism in education*. Albany: State University of New York Press.

Neuenschwander, M. P., Gerber, M., Frank, N. & Rottermann, B. (2012). *Schule und Beruf. Wege in die Erwerbstätigkeit.* Wiesbaden: VS Verlag für Sozialwissenschaften.

OECD. (2007). *Education at a glance. OECD indicators – 2007 Edition*. Paris: OECD.

OECD. (2008). *Education at a glance 2008. OECD indicators.* Paris: OECD.

OECD. (2009). *PISA data analysis manual* (2. Aufl.). Paris: OECD.

OECD. (2010). *What makes a school successful?* (*PISA 2009 Results*). Paris: OECD.

OECD. (2011). *Quality time for students. Learning in and out of school*. Paris: OECD.

OECD. (2013). *PISA 2012 results: What makes schools successful? Resources, policies and practices. Volume IV*. Paris: OECD.

OECD. (2016a). *PISA 2015 results (Volume I): Excellence and equity in education.* Paris: OECD.

OECD. (2016b). *PISA 2015 assessment and analytical framework. Science, reading, mathematic and financial literacy*. Paris: OECD.

OECD. (2016c). *PISA 2015 results (Volume II): Policies and practices for successful schools.* Paris: OECD.

Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, D., Neubrand, M., Pekrun, R., Rost, J. & Schiefele, U. (2006). *PISA 2003. Untersuchungen zur Kompetenzentwicklung im Verlauf eines Schuljahres*. Münster: Waxmann.

Sälzer, C. & Heine, J.-H. (2016). Students' skipping behavior on truancy items and (school) subjects and its relation to test performance in PISA 2012. *International Journal of Educational Development, 46,* 103–113.

© Waxmann Verlag GmbH

Sälzer, C., Prenzel, M. & Klieme, E. (2013). Schulische Rahmenbedingungen der Kompetenzentwicklung. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012: Fortschritte und Herausforderungen in Deutschland* (S. 155–187). Münster: Waxmann.

Scheerens, J. (2004). *Review of school and instructional effectiveness research. Paper commissioned for the EFA Global Monitoring Report 2005, The Quality Imperative*. Paris: Unesco.

Scheerens, J. & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.

Seidel, T. & Prenzel, M. (2008). Large scale assessment. In J. Hartig, E. Klieme & D. Leutner (Hrsg.), *Assessment of competencies in educational contexts* (S. 279–304). Göttingen: Hogrefe & Huber.

Senkbeil, M., Drechsel, B. & Schöps, K. (2007). Schulische Rahmenbedingungen und Lerngelegenheiten für die Naturwissenschaften. 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. 181–202). Münster: Waxmann.

Stanat, P., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Schümer, G., Tillmann, K.-J. & Weiß, M. (2002), *PISA 2000: Die Studie im Überblick. Grundlagen, Methoden und Ergebnisse*. Zugriff am 03.11.2016. Verfügbar unter https://www.mpib-berlin.mpg.de/Pisa/PISA_im_Ueberblick.pdf

Tippelt, R. (2013). Wandel pädagogischer Institutionen. In I. Gogolin, H. Kuper, H.-H. Krüger & J. Baumert (Hrsg.), *Stichwort: Zeitschrift für Erziehungswissenschaft* (S. 183–198). Wiesbaden: Springer VS.

van Ackeren, I., Klemm, K. & Kühn, S. M. (2015). Die qualitätsorientierte Perspektive: Wie stellt sich die Qualität der deutschen Schule im Spiegel der Leistungsstudien dar? In I. van Ackeren, K. Klemm & S. M. Kühn (Hrsg.), *Entstehung, Struktur und Steuerung des deutschen Schulsystems. Eine Einführung* (3., überarb. u. aktual. Aufl., S. 121–160). Wiesbaden: Springer VS.

Vieluf, S., Kunter, M. & van de Vijver, F. J. R. (2013). Teacher self-efficacy in cross-national perspective. *Teaching and Teacher Education, 35,* 92–103.

© Waxmann Verlag GmbH

6 Mathematische Kompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven

Sabine Hammer, Kristina Reiss, Matthias C. Lehner,
Jörg-Henrik Heine, Christine Sälzer & Aiso Heinze

Schülerinnen und Schüler in Deutschland erreichen bei PISA 2015 in der Mathematik durchschnittlich 506 Punkte und liegen damit 16 Punkte über dem OECD-Durchschnitt. Im Vergleich zur letzten PISA-Erhebung 2012 hat sich dieser Wert nicht signifikant verändert, wohingegen sich die Leistungsstreuung reduziert hat. Im internationalen Vergleich befindet sich Deutschland weiterhin im oberen Drittel der OECD-Staaten, auch wenn der Anschluss an die Spitzengruppe nicht gelingt. Nach wie vor gibt es deutliche Leistungsunterschiede in Mathematik zwischen Mädchen und Jungen. Hier verfügen die Jungen über eine höhere mathematische Kompetenz als Mädchen, insbesondere überwiegt der Anteil der Jungen im oberen Leistungsbereich. Im unteren Leistungsbereich finden sich dagegen mehr Mädchen als Jungen. Deutschland gehört dabei zu den OECD-Staaten, in denen die Diskrepanz der mathematischen Kompetenz zwischen den Geschlechtern am größten ist. Der Anteil leistungsschwacher Schülerinnen und Schüler – also derjenigen, die Kompetenzstufe II nicht erreichen und nur über sehr grundlegende mathematische Kenntnisse verfügen – konnte im Vergleich zur letzten PISA Erhebung nicht reduziert werden, ist aber deutlich geringer als zu Beginn der Studien im Jahr 2000. An den Gymnasien ist der Anteil besonders leistungsstarker Fünfzehnjähriger seit der letzten Erhebung zurückgegangen. An den nicht gymnasialen Schularten zeigt sich diese Entwicklung nicht.

PISA hat im Sinne eines Bildungsmonitorings das Ziel, das durchschnittliche Kompetenzniveau fünfzehnjähriger Schülerinnen und Schüler in der Mathematik (neben den Naturwissenschaften und dem Lesen) zu beschreiben. Verbunden damit können die Ergebnisse der Studie aber auch aufzeigen, welche konkreten Aufgabenstellungen Schülerinnen und Schüler auf unterschiedlichen Kompetenzniveaus erfolgreich bearbeiten. In der Folge ist es möglich, Stärken und Schwächen, Herausforderungen und Probleme des Bildungssystems zu identifizieren und Hinweise auf Erklärungen und Handlungsoptionen zu geben. Der internationale Kontext ist geeignet, das von Jugendlichen in Deutschland erreichte Kompetenzniveau in der Mathematik einzuordnen und über den Vergleich Ideen dafür zu gewinnen, wie man die Bildungsqualität verbessern kann. Ein wesentlicher Aspekt ist, ob gleiche Bildungschancen für alle Schülerinnen und Schüler in Deutschland im Bereich

© Waxmann Verlag GmbH

Mathematik realisiert werden können. In diesem Kapitel werden zunächst die Ergebnisse für Deutschland im Detail dargestellt. Bei der Frage, inwieweit *alle* Jugendlichen entsprechend ihrem Bedarf gefördert werden, wird besonders auf mögliche Disparitäten zwischen Mädchen und Jungen, unterschiedlichen Schularten sowie leistungsschwachen und leistungsstarken Schülerinnen und Schülern eingegangen.

Mathematische Kompetenz in PISA

Sowohl in alltäglichen als auch in professionellen Kontexten werden zunehmend Anforderungen gestellt, deren Bewältigung mathematische Kompetenz voraussetzt. Diese Kompetenz hat wenig zu tun mit einer schematischen Anwendung mathematischer Regeln. Sie ist vielmehr mit einer spezifischen Sichtweise auf Problemstellungen und Situationen verbunden, die einer Bearbeitung beziehungsweise Bewertung mittels mathematischer Werkzeuge zugänglich gemacht werden. Die Rahmenkonzeption von PISA (OECD, 2013b; OECD, 2016) zeichnet genauso wie die in Deutschland gültigen Bildungsstandards (z. B. KMK, 2003) ein differenziertes Bild mathematischer Kompetenz und zeigt auf, dass Problemlösen, Argumentieren, Modellieren, Kommunizieren und das Verwenden vielfältiger Darstellungen die Fundamente mathematischen Arbeitens sind. Darüber hinaus gilt es, mit Mathematik symbolisch, formal und technisch umzugehen sowie Hilfsmittel (z. B. Computer) geeignet einzusetzen. In PISA geht es entsprechend um die Fähigkeit, mathematisches Wissen und mathematische Prozesse funktional auf unterschiedliche Problemstellungen anzuwenden. Mathematische Kompetenz in diesem Sinn ist eine der Voraussetzungen dafür, dass junge Menschen den Herausforderungen im persönlichen, beruflichen und gesellschaftlichen Leben gewachsen sind (OECD, 2013b). Entsprechend wird im Rahmen von PISA der Begriff der mathematischen Grundbildung verwendet. Eine mathematische Grundbildung in diesem Sinn kann als Voraussetzung für eine selbstbestimmte gesellschaftliche Teilhabe von Bürgerinnen und Bürgern an einer modernen Gesellschaft angesehen werden. Sie ist darüber hinaus essenziell für die individuelle Entwicklung, für einen erfolgreichen Einstieg in das Erwerbsleben und eine wichtige Voraussetzung für kontinuierliches Lernen auch nach der Pflichtschulzeit.

Ausgewählte Maßnahmen zur Förderung mathematischer Kompetenz

Ergebnisse der ersten international vergleichenden Schulleistungsstudien wie PISA und TIMSS zeigten, dass Schülerinnen und Schüler in Deutschland im internationalen Vergleich deutliche Defizite im Bereich Mathematik – und hier insbesondere beim Lösen komplexerer Modellierungsaufgaben – hatten (Baumert et al., 1997; Klieme, Neubrand & Lüdtke, 2001). In der Konsequenz wurden zahlreiche Maßnahmen initiiert, die zu einer Verbesserung des Unterrichts und damit zu einer besseren Unterstützung der Kompetenzentwicklung von Kindern und Jugendlichen beitragen sollten. Exemplarisch seien die folgenden Aktivitäten genannt:

© Waxmann Verlag GmbH

- Mit der Einführung von Bildungsstandards wurde ein bundesweit abgestimmter, länderübergreifender Bezugsrahmen für die Bildungsqualität an Schulen festgelegt. In den Bildungsstandards werden fachbezogene Kompetenzen bezüglich verschiedener Anforderungsbereiche beschrieben, die Schülerinnen und Schüler bis zu einem bestimmten Zeitpunkt erreicht haben sollen. Zur Veranschaulichung dienen konkrete Aufgabenstellungen. In diesem Zusammenhang wurde 2004 das Institut zur Qualitätsentwicklung im Bildungswesen (IQB) gegründet, das mit der Operationalisierung und Überprüfung der in den Bildungsstandards formulierten Leistungserwartungen betraut ist.
- Es wurden umfassende Bemühungen unternommen, Materialien und Aufgabenformate zur Verbesserung der Unterrichtsqualität zu entwickeln. Umfangreiche Programme wie *SINUS* und *SINUS-Transfer* hatten zum Ziel, auf Schulebene in den mathematisch-naturwissenschaftlichen Fächern Prozesse der Qualitätssicherung und der Verbesserung des Lehrens und Lernens anzustoßen und weiterzuentwickeln (z. B. Prenzel, Friedrich & Stadler, 2009). Ein Baustein war beispielsweise die Entwicklung einer kognitiv anregenden und reichhaltigen Aufgabenkultur, die durch einen reflektierten Umgang mit Aufgaben gekennzeichnet ist.
- Die Förderung leistungsschwacher – teilweise auch leistungsstarker – Schülerinnen und Schüler wurde in den Bundesländern durch vielfältige Ansätze und Maßnahmen unterstützt. Länderübergreifend wurde im Rahmen einer *Förderstrategie* von der Kultusministerkonferenz (KMK, 2010) zunächst ein besonderer Fokus auf die Gruppe der leistungsschwachen Schülerinnen und Schüler gelegt. Inzwischen ist eine entsprechende Strategie auch für die leistungsstärkeren Schülerinnen und Schüler formuliert worden (KMK, 2015).

Auch wenn mit PISA durch die Anlage als Querschnittsstudie keine Ursache-Wirkungs-Zusammenhänge untersucht werden können, deutet sich doch an, dass diese Maßnahmen Wirkung gezeigt haben. Sie dürften insbesondere dazu beigetragen haben, dass sich leistungsschwächere Schülerinnen und Schüler zwischen den Messzeitpunkten 2000 und 2012 kontinuierlich in ihrer Leistung verbessert haben. Insgesamt zeichnete sich von 2003 bis 2012 eine stetige Verbesserung der mathematischen Kompetenz von Jugendlichen in Deutschland ab (Sälzer, Reiss, Schiepe-Tiska, Prenzel & Heinze, 2013).

6.1 Mathematische Kompetenz in PISA 2015

Im Folgenden wird zunächst die Rahmenkonzeption mathematischer Kompetenz dargestellt und erläutert, wie diese im Mathematiktest in PISA 2015 umgesetzt wurde. Da Mathematik 2015 eine Nebendomäne in der PISA-Studie war, wird die mathematische Kompetenz nicht differenziert nach mathematischen Inhaltsbereichen erfasst und analysiert, sondern ausschließlich eine Gesamtskala berichtet. Es werden anschließend die Ergebnisse von PISA 2015 vor allem im Hinblick auf drei Fragestellungen berichtet:

© Waxmann Verlag GmbH

- Wie kann die mathematische Kompetenz von Fünfzehnjährigen in Deutschland im internationalen Vergleich eingeordnet werden?
- Hat sich die durchschnittliche mathematische Kompetenz der fünfzehnjährigen Schülerinnen und Schüler in Deutschland seit der PISA-Studie 2012 verändert?
- Gibt es bezogen auf die Mathematikleistung Disparitäten zwischen Mädchen und Jungen beziehungsweise zwischen verschiedenen Schularten?

Es wird dabei besonders die Situation leistungsschwächerer und leistungsstärkerer Schülerinnen und Schüler herausgearbeitet. Abschließend wird darauf eingegangen, wie sich die Ergebnisse einordnen lassen und welche Hinweise sich für die Förderung mathematischer Kompetenz daraus ergeben.

Mathematische Grundbildung oder *Mathematical Literacy* wird von der in PISA eingesetzten internationalen Expertengruppe definiert 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, 2016, S. 65). Unter Grundbildung wird demnach insbesondere die Fähigkeit gefasst, mathematische Begriffe, Aussagen und Prozesse zur Bearbeitung von Problemen, die in einem realen Kontext situiert sind, anwenden zu können. Außerdem wird in der Definition die Bedeutsamkeit einer aktiven Auseinandersetzung mit der Mathematik herausgestellt. Entsprechend richten sich die bei PISA eingesetzten Aufgaben vor allem auf die alltagspraktische Anwendung mathematischer Kenntnisse in unterschiedlichen realen Kontexten. Rein innermathematische Themen, also solche, bei denen explizit eine mathematische Problemstellung ohne realen Kontext formuliert wird, sind hingegen eher nicht Gegenstand von PISA-Mathematikaufgaben.

Die Rahmenkonzeption der Domäne Mathematik

Der aktuellen PISA-Erhebung liegt die theoretische Rahmenkonzeption von PISA 2012 zugrunde. Für PISA 2012 war diese Konzeption der Domäne Mathematik grundlegend überprüft und in Teilen erweitert worden, da Mathematik nach PISA 2003 bereits zum zweiten Mal als Hauptdomäne erfasst wurde. Diese Rahmenkonzeption dient dazu, die vielfältigen Facetten mathematischer Grundbildung und deren Breite abzubilden, zu strukturieren und schlussendlich anhand von Aufgaben zu erfassen, um ein differenziertes Bild über die Ausprägung der mathematischen Kompetenz fünfzehnjähriger Schülerinnen und Schüler zu erhalten.

Die Rahmenkonzeption mathematischer Grundbildung beschreibt mit Inhalten, Prozessen und Kontexten wesentliche Aspekte (OECD, 2016) von PISA-Aufgaben:

© Waxmann Verlag GmbH

- Grundlegende mathematische *Inhalt*sbereiche
- *Prozesse*, die elementare mathematische Aktivitäten präzisieren und die ihnen zugrunde liegenden *fundamentalen mathematischen Fähigkeiten*, welche zum Lösen mathematischer Aufgaben erforderlich sind
- *Kontexte*, in denen die Aufgaben eingebettet sind

Wie in PISA 2012 behandeln die in der aktuellen PISA-Erhebung verwendeten Aufgaben vier mathematische Inhaltsbereiche (OECD, 2016):

- *Veränderung und Beziehungen* greift Situationen auf, in denen sich – etwa im Verlauf der Zeit – Objekte gegenseitig beeinflussen oder Veränderungen in Systemen mit mehreren Elementen auftreten. Damit verbundene Problemstellungen treten beispielsweise beim Wachstum von Organismen, in der Musik oder beim Klima auf. Zentral für diesen Bereich sind die Inhalte aus den mathematischen Teilgebieten Algebra und Funktionen.
- *Raum und Form* behandelt Phänomene, die in unserer Umwelt physikalisch oder visuell präsent sind. Der Bereich der ebenen Geometrie liefert viele Themen für diesen Inhaltsbereich. Jedoch geht Raum und Form darüber hinaus und umfasst neben der ebenen und räumlichen Geometrie auch räumliches Vorstellungsvermögen sowie das damit verbundene Messen.
- *Quantität* bezieht sich auf alle Arten von Quantifizierung, wie das Beschreiben von Situationen unter Verwendung von Zahlen, das Größenverständnis und das Identifizieren von Zahlenmustern. Beispiele aus diesem Bereich sind (ebenfalls) das Messen, das Zählen, der Umgang mit Größen und Einheiten sowie mit Anteilen.
- *Unsicherheit und Daten* betrifft mathematische Situationen, die im Zusammenhang mit statistischen Daten und dem Zufall stehen. Beispiele hierfür sind Umfrageergebnisse, Wettervorhersagen oder wissenschaftliche Prognosen.

Diese vier Inhaltsbereiche werden als sogenannte *übergreifende Ideen* aufgefasst. Sie decken grundlegende Bereiche der Mathematik sowie die zugrunde liegenden Phänomene ab und spiegeln die Inhalte der Schulcurricula wider. Auch wenn PISA mit dem zugrunde gelegten *Literacy*-Begriff keine curriculare Validität anstrebt, existieren doch Überschneidungen mit Lehrplänen und den Bildungsstandards (vgl. Hartig & Frey, 2012). Wie der Grundbildungsbegriff in PISA werden auch Lehrpläne entlang der Frage entwickelt, was Jugendliche zu einem bestimmten Zeitpunkt oder in einer bestimmten Phase ihrer Bildungsbiografie wissen und beherrschen sollten. Dennoch ist es gewollt, dass Schülerinnen und Schüler im PISA-Test auch Aufgabenstellungen vorfinden, die für sie neu sind und Kreativität und Transferfähigkeit erfordern. Der Bereich *Raum und Form* ist beispielsweise auch eine Leitidee in den Bildungsstandards für den Mittleren Schulabschluss sowie die Sekundarstufe II in Deutschland (KMK, 2003; 2012). Diese Leitidee beinhaltet viele Aspekte aus dem Stoffgebiet Geometrie des Schulcurriculums – allerdings nicht das Messen.

© Waxmann Verlag GmbH

Neben diesem inhaltsorientierten Aspekt ist die zweite Dimension der theoretischen Rahmenkonzeption mathematischer Grundbildung bei PISA anforderungsorientiert. Die Beschreibung der *Prozesse* konkretisiert grundlegende mathematische Aktivitäten bei der Lösung von Aufgaben (OECD, 2016):

- Situationen mathematisch *formulieren*
- Mathematische Konzepte, Fakten, Prozeduren und Schlussfolgerungen *anwenden*
- Mathematische Ergebnisse *interpretieren*, *anwenden* und *bewerten*

Der Prozess des *Formulierens* bezieht sich auf die Fähigkeit, in einer Situation oder einem Problem eine mathematische Struktur zu erkennen und diese in eine geeignete mathematische Darstellungsform zu übertragen. Dazu zählen zum Beispiel das Vereinfachen eines Problems, um es für eine mathematische Untersuchung zugänglich zu machen, oder auch das Erkennen von Regelmäßigkeiten, Beziehungen und Mustern in Situationen. Die *Anwendung* mathematischer Fakten, Prozeduren und Schlussfolgerungen bezieht sich auf verschiedene Arbeitsweisen, um zur Lösung eines mathematischen Problems zu gelangen. *Interpretieren* umfasst das Reflektieren über eine mathematische Lösung oder Schlussfolgerung, deren Deutung bezüglich des Kontexts und die Bewertung, inwieweit ein Ergebnis sinnvoll ist (OECD, 2016).

Den drei Prozessen liegen *fundamentale mathematische Fähigkeiten* („*capabilities*“) zugrunde. Diese sind zur Lösung mathematischer Aufgaben notwendig und wurden von zentralen mathematischen Teilkompetenzen abgeleitet (Niss, 2003; Niss & Højgaard, 2011). In der Rahmenkonzeption von PISA werden diese fundamentalen mathematischen Fähigkeiten *Kommunizieren, Mathematisieren, Repräsentieren, Argumentieren, Problemlösestrategien entwickeln, mit Mathematik symbolisch formal und technisch umgehen* und *mathematische Hilfsmittel verwenden* detailliert beschrieben (OECD, 2016). Diese Fähigkeiten sind für jeden der drei oben beschriebenen Prozesse in unterschiedlichem Maße erforderlich. Abhängig von einer spezifischen Aufgabe und deren Komplexität werden von Schülerinnen und Schülern weniger oder mehr Fähigkeiten zur Lösung benötigt.

Die dritte Dimension zur Strukturierung der mathematischen Grundbildung sind die Kontexte (PISA 2003: *Situationen*), in denen mathematische Fragestellungen und Probleme verortet sind. In PISA werden Aufgaben in unterschiedliche Kontexte eingeordnet, konkret in *persönliche, berufliche, gesellschaftliche* und *wissenschaftliche* Kontexte (OECD, 2016). Dies bietet die Möglichkeit, einen breiten Interessensbereich abzudecken und eine Vielfalt von Situationen anzusprechen, in denen sich Jugendliche im 21. Jahrhundert befinden.

Der Mathematiktest in PISA 2015

Wie in der Rahmenkonzeption beschrieben, erfordern die in PISA eingesetzten Aufgaben eine funktionale Anwendung mathematischer Kenntnisse und Fähigkeiten auf kon-

© Waxmann Verlag GmbH

krete Probleme aus unterschiedlichen realen Kontexten. Da Mathematik bei PISA 2015 als Nebendomäne untersucht wurde, sind dies Mathematikaufgaben aus bereits vergangenen PISA-Erhebungen, sogenannte Anker- oder *Link*-Aufgaben. In PISA 2015 wurden diese Link-Aufgaben allerdings von dem ursprünglich papierbasierten Format auf ein geeignetes Format zur computerbasierten Testung übertragen (vgl. Kapitel 12). Insgesamt 45 Mathematikaufgaben mit 81 Teilaufgaben (*Items*) dienten der Messung mathematischer Kompetenz.[1] Die Aufgaben decken die vier Inhaltsbereiche, die drei Prozesse, die vier Kontextkategorien und darüber hinaus unterschiedliche Schwierigkeitsgrade ab (siehe Tabelle 6.1). Die sprachlichen Anforderungen der Aufgaben werden durch im Wesentlichen direkt formulierte Aufgabeninstruktionen möglichst gering gehalten, sodass der Einfluss der Lesekompetenz der Schülerinnen und Schüler auf die Aufgabenbearbeitung reduziert wird. Die Kontexte der Aufgaben sind so gewählt, dass sie zum einen relevant für die Jugendlichen sind und zum anderen einen Einstieg in die Aufgaben und deren Verstehen unabhängig vom jeweiligen Kulturkreis ermöglichen.

Tabelle 6.1 Verteilung der Items des Mathematiktests auf die Inhalte, Prozesse und Kontexte

	Anzahl der Items
Inhalte	
Veränderung und Beziehungen	20
Raum und Form	19
Quantität	21
Unsicherheit und Daten	21
Prozesse	
Formulieren	23
Anwendung	35
Interpretieren	23
Kontexte	
persönlich	13
beruflich	20
gesellschaftlich	28
wissenschaftlich	20

Die Kompetenzstufen und ein Aufgabenbeispiel

Durch die Anwendung einer IRT-Skalierung lassen sich die Schwierigkeiten der einzelnen Aufgaben und die Mathematikkompetenz der Schülerinnen und Schüler auf einer gemeinsamen Skala abbilden (vgl. Kapitel 12). Diese Skala wurde bei PISA 2003 für die

1 Eine detaillierte Übersicht zu den eingesetzten Aufgaben und Teilaufgaben findet sich in Anhang B.

© Waxmann Verlag GmbH

Tabelle 6.2: Stufen mathematischer Kompetenz

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**	

© Waxmann Verlag GmbH

Mathematik erstmals mit einem Mittelwert von $M = 500$ und einer Standardabweichung von $SD = 100$ normiert (vgl. OECD, 2003; Blum et al., 2004; Sälzer, 2016). Aufgrund der gemeinsamen Skala von Personenfähigkeit und Aufgabenschwierigkeit kann die mathematische Kompetenz der Schülerinnen und Schüler in PISA anforderungsbezogen und inhaltlich interpretiert werden: Dazu dienen sechs Niveaus beziehungsweise Kompetenzstufen. Jede dieser Stufen wurde bereits bei PISA 2003 durch die Anforderungen der Aufgaben, die gemäß ihrer Aufgabenschwierigkeit in dieser Stufe einzuordnen sind, charakterisiert (OECD, 2003; OECD, 2016). Der seit 2003 unveränderte Anforderungskatalog der sechs Kompetenzstufen ist in Tabelle 6.2 dargestellt. Jede der Kompetenzstufen I bis V umfasst dabei 62 Punkte auf der Kompetenzskala, Stufe VI hat lediglich eine untere Grenze.

Ein besonderes Augenmerk wird in diesem Berichtsband auf Jugendliche der obersten und untersten Kompetenzstufen gelegt. Schülerinnen und Schüler auf den Stufen V oder VI verfügen über eine hohe bis sehr hohe mathematische Kompetenz. Hingegen beherrschen Schülerinnen und Schüler auf Stufe I oder darunter allenfalls sehr grundlegende mathematische Anwendungen. Entsprechend haben diese Jugendlichen Nachholbedarf in Bezug darauf, was gemäß der Grundbildungskonzeption in PISA als notwendige Bedingung für anschlussfähiges, anwendbares Wissen und Können gilt. Dieser Gruppe am unteren Ende der Kompetenzskala drohen erschwerte Bedingungen beim Übergang von der Pflichtschulzeit in eine Ausbildung oder anderweitig weiterführendes Lernen und damit letztlich in ein eigenständiges Erwerbsleben.

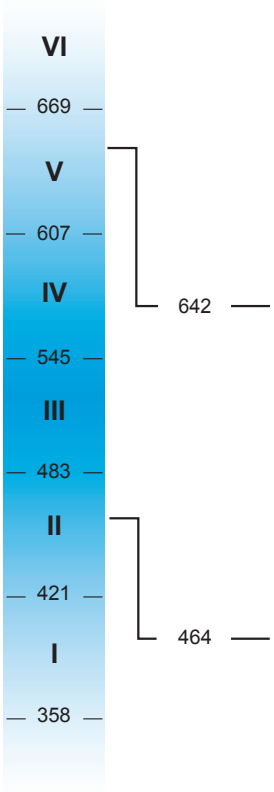
Um die Kompetenzstufen zu konkretisieren und verständlicher zu machen, wird im Folgenden die Beispielaufgabe *Bergsteigen am Mount Fuji* (OECD, 2014; 2013a) aus dem PISA-Mathematiktest 2012 vorgestellt (Abbildung 6.1). Die Aufgabe ist in einen Alltagskontext eingebettet und verdeutlicht exemplarisch den Anwendungsbezug des Tests. Zur Lösung der ersten Teilaufgabe soll die mittlere Zahl der Personen bestimmt werden, die pro Tag den Mount Fuji besteigen. Dazu muss im ersten Schritt aus den Daten die Anzahl der Tage entnommen werden, an denen die Besteigung möglich ist. Mit diesem Wert kann in einem zweiten Schritt die gesuchte mittlere Zahl der Besucher pro Tag ermittelt werden. Der Fokus dieser Teilaufgabe liegt also auf dem Bestimmen der Anzahl der Bergsteiger, weswegen diese Teilaufgabe dem Inhaltsbereich *Quantität* zugeordnet wurde. In den deutschen Bildungsstandards für den Mittleren Schulabschluss (KMK, 2003) finden sich die inhaltlichen Anforderungen dieser Aufgabe in der Leitidee *Zahl* wieder. Auf der Prozessebene liegt das Hauptaugenmerk darauf, eine vorliegende Situation mathematisch zu formulieren (OECD, 2013b; 2014), was in den Bildungsstandards in der mathematischen Kernkompetenz *Modellieren* wiederzufinden ist. Die Teilaufgabe hat ein geschlossenes Format und ist mit einer Schwierigkeit von 464 Punkten eine eher einfache Aufgabe. Sie ist in die Kompetenzstufe II eingeordnet.

Bei der zweiten Teilaufgabe müssen Schülerinnen und Schüler zunächst für den Aufstieg auf den Berg sowie für den Rückweg die Zeit bestimmen, die für eine 9 Kilometer lange Strecke bei der angegebenen Geschwindigkeit benötigt wird. Für den Aufstieg

© Waxmann Verlag GmbH

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

Anmerkung: Die Werte an den Verbindungslinien zwischen den Beispielen und der Kompetenzsäule geben das Kompetenzniveau an, mit dem eine Schülerin oder ein Schüler die Aufgabe mit einer Wahrscheinlichkeit von 62% lösen würde (vgl. Frey et al., 2008).

Abbildung 6.1: Beispielaufgabe „Bergsteigen am Mount Fuji“

sind 6 Stunden zu planen, für den Abstieg 3 Stunden. Dies führt zu dem Ergebnis, dass der Protagonist Toshi spätestens um 11 Uhr aufbrechen sollte. Es wird hier vor allem ein grundlegendes Verständnis von Geschwindigkeit benötigt, um das Problem zu lösen. Wie auch bei der ersten Teilaufgabe steht der Prozess des mathematischen Formulierens der gegebenen Problemsituation im Vordergrund. Auf der Inhaltsebene wird der Bereich *Veränderung und Beziehungen* thematisiert. Auch die Anforderungen dieser Aufgabe sind in den Bildungsstandards in der mathematischen Kompetenz *Modellieren* enthalten. Die zweite Teilaufgabe hat eine Schwierigkeit von 642 Punkten und wird entsprechend der Kompetenzstufe V zugeordnet. Zwar mag das offene Antwortformat einen Beitrag zur höheren Schwierigkeit leisten, entscheidend ist aber vor allem die kognitive Kom-

© Waxmann Verlag GmbH

plexität dieser Teilaufgabe. Insbesondere der Anspruch des Mathematisierungsprozesses ist verhältnismäßig hoch, zudem sind im Vergleich zur ersten Teilaufgabe mehr Schritte nötig, um das vorliegende Problem zu lösen.

6.2 Mathematische Kompetenz im internationalen Vergleich

Dieser Abschnitt ist der Frage gewidmet, wie sich die mathematische Kompetenz der Schülerinnen und Schüler aus Deutschland im internationalen Vergleich insbesondere der OECD-Staaten einordnen lässt. Einen ersten Eindruck liefern die Testmittelwerte und Leistungsstreuungen in den teilnehmenden Staaten. Eine inhaltliche Interpretation der mathematischen Kompetenz der Schülerinnen und Schüler aus den einzelnen Staaten kann der Verteilung auf die beschriebenen Kompetenzstufen entnommen werden. Schließlich werden Unterschiede zwischen Mädchen und Jungen im internationalen Vergleich betrachtet.

Mathematische Kompetenz: Mittelwerte im Vergleich

Bei PISA 2015 liegt die durchschnittliche mathematische Kompetenz über alle OECD-Staaten hinweg bei 490 Punkten, die Standardabweichung beträgt 89 Punkte. In PISA 2003 waren Mittelwert (*M*) und Standardabweichung (*SD*) auf 500 beziehungsweise 100 festgesetzt worden, was sich in den nachfolgenden Erhebungsrunden – teils bedingt durch eine minimal andere Zusammensetzung in der Gruppe der OECD-Staaten – leicht verändert hat. Fünfzehnjährige in Deutschland erreichen im Mathematiktest 506 Punkte (*SD* = 89) und liegen damit auch in PISA 2015 signifikant über dem OECD-Durchschnitt. In Abbildung 6.2 werden die wichtigsten Befunde des internationalen Vergleichs zusammengefasst und Mittelwerte sowie Perzentilbänder der mathematischen Kompetenz in den OECD-Teilnehmerstaaten dargestellt. Die Mittelwerte verdeutlichen die durchschnittlich erreichte mathematische Kompetenz, wohingegen die Perzentilbänder graphisch Informationen zur Streuung und Verteilung der Leistung liefern. Die OECD-Teilnehmerstaaten werden dabei in drei Gruppen unterteilt: solche, die statistisch über oder unter dem OECD-Durchschnitt liegen und solche, die um den OECD-Durchschnitt liegen und sich nicht signifikant von diesem unterscheiden.

In 17 Staaten erreichen Schülerinnen und Schüler mittlere Kompetenzwerte, die signifikant höher als der OECD-Durchschnitt sind. Auch Deutschland gehört zu dieser Gruppe von Staaten mit einer überdurchschnittlichen mathematischen Kompetenz. An der Spitze dieser Gruppe stehen Japan mit 532 Punkten, Korea mit 524 Punkten, die Schweiz mit 521 Punkten, Estland mit 520 Punkten und Kanada mit 516 Punkten. Die erreichten Kompetenzwerte dieser fünf Staaten liegen bis zu knapp einer halben Standardabweichung über dem OECD-Durchschnitt und befinden sich auch signifikant über dem mittleren Kompetenzwert von Deutschland. Deutschlands relative Position innerhalb der OECD-Staaten ist damit im Wesentlichen unverändert. Auch in PISA 2012 lag

© Waxmann Verlag GmbH

die mittlere mathematische Kompetenz der Fünfzehnjährigen in Deutschland unter derjenigen fünf anderer OECD-Staaten. Die Schweiz ist der einzige Nachbarstaat Deutschlands, der zu dieser Gruppe gehört. Andere Nachbarstaaten wie die Niederlande, Dänemark, Belgien und Polen weisen einen zu Deutschland vergleichbaren Leistungsstand auf, wohingegen bei PISA 2015 Österreich, Frankreich, die Tschechische Republik und Luxemburg signifikant niedrigere Werte im Mathematiktest erzielen.

Sieben Staaten unterscheiden sich nicht signifikant vom OECD-Durchschnitt. Dies sind die europäischen Staaten Schweden (494 Punkte), Frankreich (493 Punkte), das Vereinigte Königreich (492 Punkte), die Tschechische Republik (492 Punkte), Portugal (492 Punkte), Italien (490 Punkte) und Island (488 Punkte). Das Kompetenzniveau der Staaten, die signifikant unter dem OECD-Durchschnitt liegen, reicht von Spanien mit 486 Punkten bis Mexiko mit 408 Punkten. Die mathematische Kompetenz der Schülerinnen und Schüler in Mexiko liegt damit fast eineinhalb Kompetenzstufen unter dem OECD-Durchschnitt.

Die höchsten Mittelwerte der mathematischen Kompetenz zeigen sich in einigen OECD-Partnerstaaten (vgl. Tabelle A5 im Anhang). Hier werden nochmals deutlich höhere Leistungen erreicht als im leistungsstärksten OECD-Staat Japan mit 532 Punkten. Den höchsten durchschnittlichen Wert im PISA-Mathematiktest erzielen Schülerinnen und Schüler in Singapur mit 564 Punkten. Auch Hongkong (China) mit 548 Punkten, Macao (China) mit 544 Punkten und Chinesisch Taipeh mit 542 Punkten weisen herausragende mittlere Leistungen auf. In diesen Bildungssystemen entspricht der Unterschied der mathematischen Kompetenz zum OECD-Mittelwert etwa einer Kompetenzstufe, und ihre Schülerinnen und Schüler erreichen – auch im Vergleich zu Deutschland – ein erkennbar höheres Niveau an mathematischer Kompetenz.

Die Streuung der mathematischen Kompetenz

Die Streuung der mathematischen Kompetenz um den Mittelwert herum wird in Form der Standardabweichung abgebildet. Sie kann als Maß für die Homogenität der Stichprobe gedeutet werden. Eine große Streuung bedeutet, dass es einen großen Unterschied in der Leistung zwischen schwachen und starken Schülerinnen und Schülern gibt und die Kompetenzmittelwerte der Jugendlichen entsprechend breit um den Gesamtmittelwert streuen. Die Bandbreite der erzielten Leistungen wird außerdem in den Perzentilbändern deutlich, die den Abstand zwischen den jeweils 5 Prozent leistungsschwächsten und leistungsstärksten Schülerinnen und Schülern abbilden.

© Waxmann Verlag GmbH

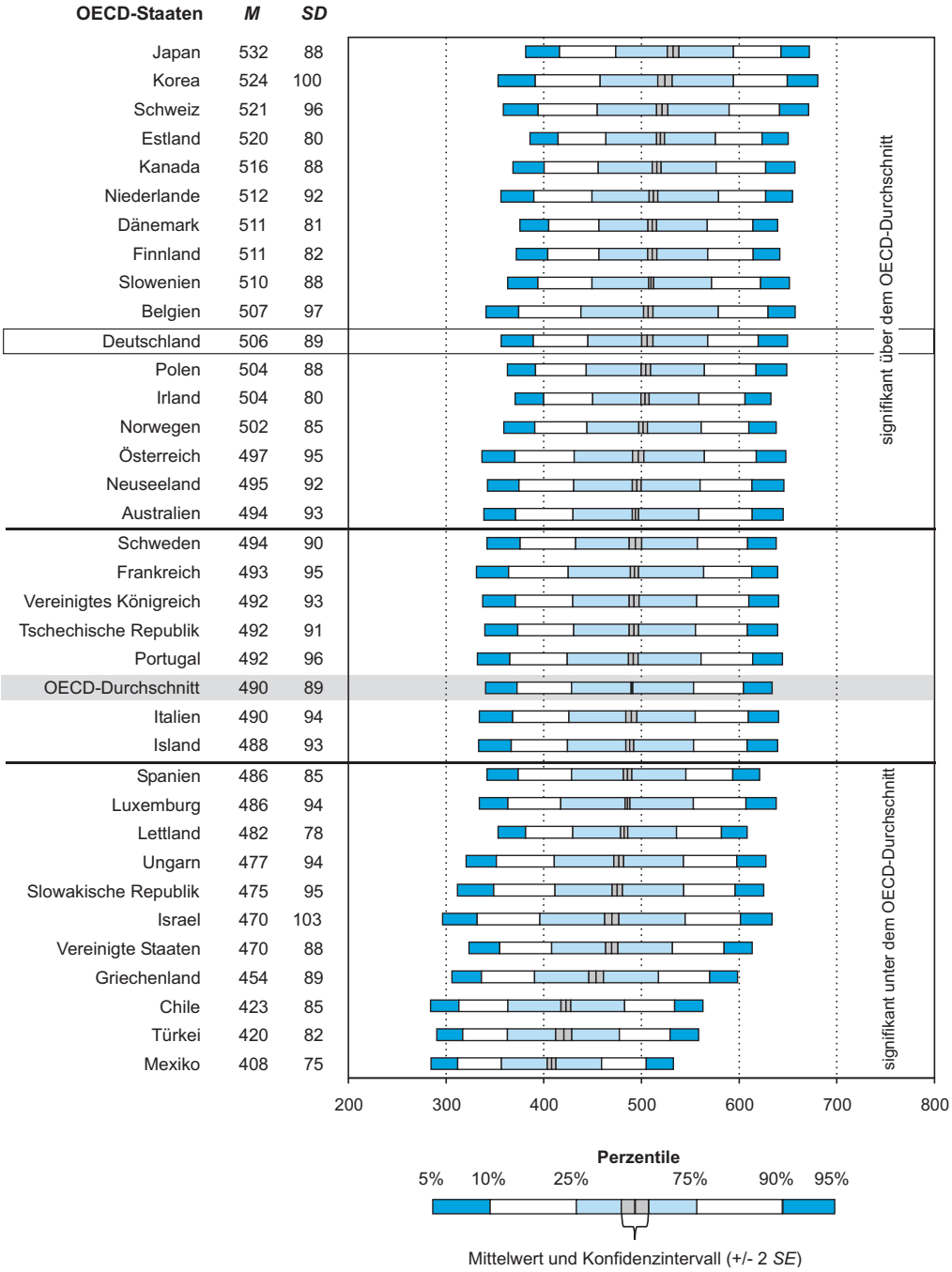


Abbildung 6.2: Perzentilbänder mathematischer Kompetenz in den OECD-Staaten

© Waxmann Verlag GmbH

Ganz allgemein sieht man in Abbildung 6.2 auch, dass sich die OECD-Staaten teilweise deutlich in den Standardabweichungen und damit in der Streuung der Leistungen unterscheiden. So weist Mexiko mit 75 Punkten die geringste Streuung innerhalb der OECD-Staaten auf, Israel dahingegen mit 103 Punkten die größte. Die Standardabweichung in Deutschland beträgt 89 Punkte und entspricht damit genau dem OECD-Durchschnitt. Für Staaten, deren durchschnittliche mathematische Kompetenz höher als der OECD-Mittelwert ist, liegt nicht unbedingt ein großer Unterschied zwischen dem oberen und dem unteren Leistungsspektrum vor. So weisen Staaten wie Estland (SD = 80), Dänemark (SD = 81) und Finnland (SD = 82) eine signifikant geringere Streuung als der OECD-Durchschnitt auf. Diesen Staaten gelingt somit die Förderung zu einem höheren durchschnittlichen Leistungsniveau auf der Grundlage einer höheren Homogenität der Leistungen. Eine eher heterogene Leistungsverteilung mit einer im Vergleich zum OECD-Durchschnitt signifikant größeren Standardabweichung weisen Korea (SD = 100), die Schweiz (SD = 96) und Belgien (SD = 97) auf. Hier unterscheiden sich demnach leistungsstarke und leistungsschwache Schülerinnen und Schüler stärker in Bezug auf die gezeigte Kompetenz.

Verteilung auf die Stufen der mathematischen Kompetenz

Anhand von Kompetenzstufen ist es möglich, die Punktzahl auf der Skala mathematischer Kompetenz mit den entsprechenden Aufgabenanforderungen in Verbindung zu bringen. Diese kriteriumsorientierte Interpretation ist insbesondere für den untersten und obersten Bereich der Kompetenzverteilung interessant. Die entsprechenden Anteile können Hinweise liefern, inwieweit es den jeweiligen Bildungssystemen gelingt, die Potenziale von leistungsstarken Schülerinnen und Schülern bezüglich ihres mathematischen Verständnisses zu fördern und eine Leistungsspitze auszubauen. Ebenso ist entscheidend, ob die Staaten einen geringen Anteil an Jugendlichen aufweisen, der nur sehr grundlegende mathematische Tätigkeiten beherrscht: Bestenfalls können Schülerinnen und Schüler, deren mathematische Kompetenz auf Kompetenzstufe I oder darunter liegt, einfache Berechnungen ausführen, und sie können mathematische Fragen in der Regel nur dann beantworten, wenn alle relevanten Informationen gegeben sind sowie die klare Fragestellung aus einem ihnen bekannten Kontext stammt. Es ist davon auszugehen, dass diese Jugendlichen für ihre weitere Ausbildung unzulänglich vorbereitet sind und nicht das Leistungsniveau erreichen, das für eine selbstbestimmte Teilhabe in einer modernen Gesellschaft notwendig ist.

In Abbildung 6.3 werden für alle OECD-Staaten die Anteile der Schülerinnen und Schüler auf den untersten Kompetenzstufen (I oder darunter) beziehungsweise auf den obersten Kompetenzstufen (V oder VI) dargestellt. Eine besonders geringe mathematische Kompetenz (Kompetenzstufe I oder darunter) weisen über alle OECD-Staaten hinweg 23 Prozent der Schülerinnen und Schüler auf. Dieser Anteil ist in Deutschland signifikant niedriger und beträgt 17 Prozent. In anderen OECD-Staaten der Spitzengruppe wie Japan (11 Prozent) und Estland (11 Prozent) gibt es einen noch einmal deut-

© Waxmann Verlag GmbH

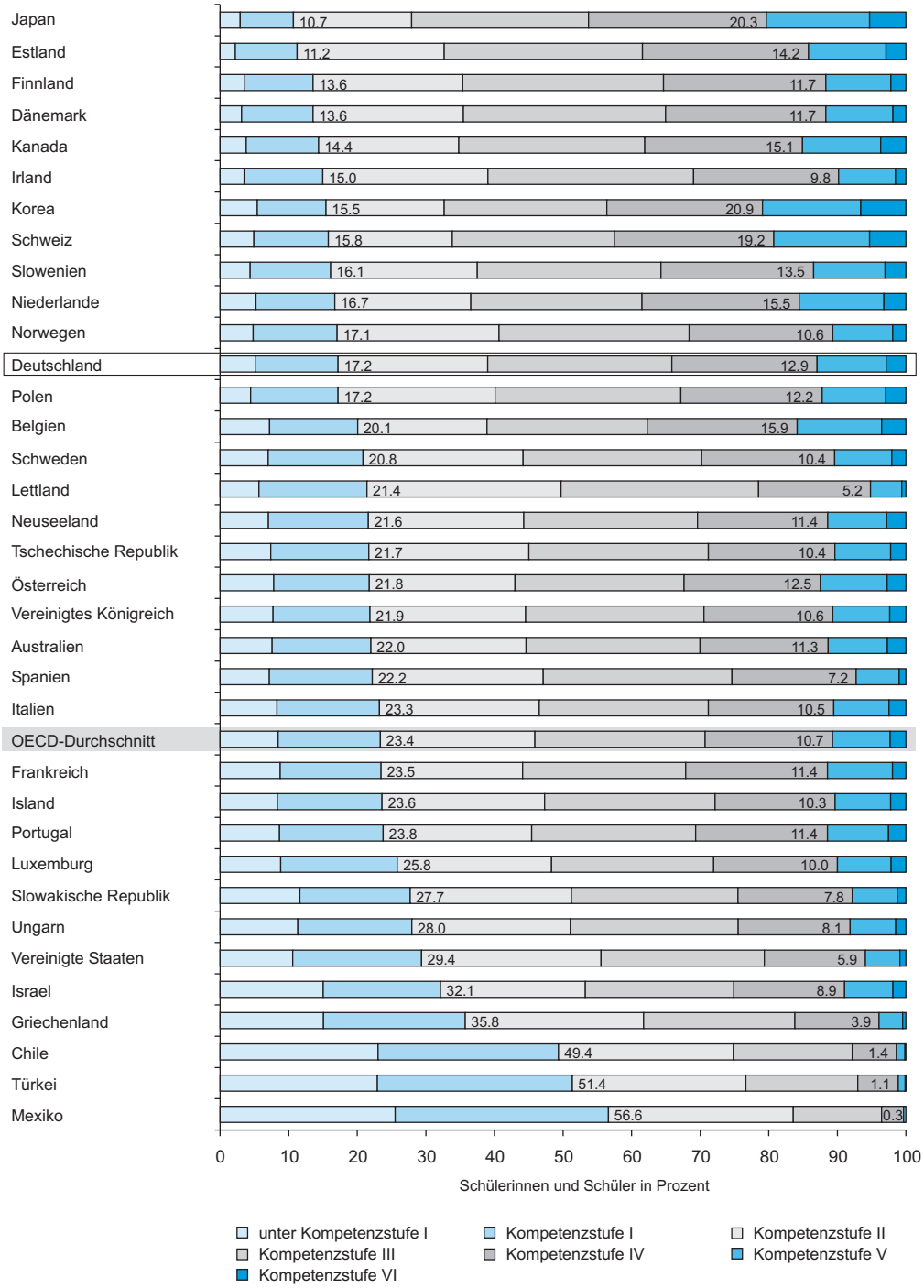


Abbildung 6.3: Prozentuale Anteile der Schülerinnen und Schüler auf Kompetenzstufe I oder darunter beziehungsweise auf Kompetenzstufe V oder VI

© Waxmann Verlag GmbH

lich geringeren Anteil leistungsschwacher Jugendlicher. Sehr niedrige Anteile im untersten Kompetenzbereich zeigen darüber hinaus leistungsstarke OECD-Partner wie Macao (7 Prozent), Singapur (8 Prozent) und Hongkong (9 Prozent). Es scheint für Bildungssysteme also prinzipiell möglich zu sein, herausragende Leistungen in der Spitzengruppe mit einem geringen Anteil leistungsschwacher Schülerinnen und Schüler zu verbinden.

Herausragende Leistungen sind im OECD-Durchschnitt bei 11 Prozent der Schülerinnen und Schüler festzustellen (Kompetenzstufe V und VI). Deutschland weist mit 13 Prozent eine im Vergleich zum OECD-Durchschnitt große Spitzengruppe im Bereich Mathematik auf. Einen wesentlich größeren Anteil Jugendlicher auf den obersten beiden Kompetenzstufen findet man in Korea (21 Prozent), Japan (20 Prozent) und der Schweiz (19 Prozent). Auch in den Niederlanden, Kanada und Estland erreichen signifikant mehr Jugendliche herausragende Leistungen als in Deutschland. Schließt man ergänzend in die Betrachtung noch leistungsstarke OECD-Partner mit ein, zeigt sich in Chinesisch Taipeh (28 Prozent) und Singapur (35 Prozent) ein sehr hoher Anteil an leistungsstarken Schülerinnen und Schülern, der in den OECD-Staaten nicht erreicht wird. Die Leistungsspitze und ebenso die Schülerinnen und Schüler im unteren Leistungsspektrum erfolgreich zu fördern, gelingt innerhalb der OECD-Staaten vor allem Japan und Estland besonders gut. Hervorzuheben ist, dass in den vier leistungsstärksten OECD-Staaten (Japan, Korea, Schweiz und Estland) signifikant mehr Schülerinnen und Schüler auf den beiden obersten Kompetenzstufen anzufinden sind als auf Kompetenzstufe I oder darunter. Zu den Staaten, in denen ein großer Anteil der Jugendlichen besonders leistungsschwach und nur ein geringer Anteil leistungsstark ist, gehören Mexiko, die Türkei, Chile und Griechenland.

Unterschiede zwischen Mädchen und Jungen in der mathematischen Kompetenz

Im OECD-Durchschnitt erreichen die Jungen mit 494 Punkten bei PISA 2015 einen signifikant höheren Wert als die Mädchen mit 486 Punkten (vgl. Abbildung 6.4). Eine solche Geschlechterdifferenz wurde in allen bisherigen PISA-Erhebungsrunden gefunden und ist nicht auf Deutschland beschränkt: In PISA 2015 liegt in 18 der 35 OECD-Staaten die mathematische Kompetenz der Jungen über der Kompetenz der Mädchen. Dieser Unterschied ist statistisch abgesichert. Deutschland gehört, wie auch Spanien, Irland, Chile, Italien und Österreich, zu den sechs Staaten, in denen die Leistungsdifferenz mehr als 15 Punkte beträgt. Die Kompetenzunterschiede zwischen Mädchen und Jungen in diesen sechs Staaten unterscheiden sich signifikant von der im OECD-Durchschnitt gefundenen Differenz von acht Punkten. Trotz des Unterschieds der mittleren Kompetenzwerte von 17 Punkten zwischen Mädchen und Jungen in Deutschland zeigen die Streuungen der beiden Gruppen, dass sich die Leistungsverteilungen zu einem großen Teil überlappen.

In 16 OECD-Staaten lässt sich kein signifikanter Leistungsunterschied zwischen Mädchen und Jungen feststellen. Zu diesen Ländern gehört beispielsweise Korea, aber auch in Norwegen, Schweden, Lettland, Griechenland oder Israel zeigen Mädchen und

© Waxmann Verlag GmbH

OECD-Staaten	Mädchen M	Mädchen (SE)	Jungen M	Jungen (SE)
Finnland	515	(2.6)	507	(2.6)
Korea	528	(3.9)	521	(5.2)
Norwegen	503	(2.3)	501	(2.9)
Schweden	495	(3.3)	493	(3.8)
Lettland	483	(2.5)	481	(2.6)
Island	489	(2.4)	487	(2.9)
Griechenland	454	(3.6)	454	(4.7)
Niederlande	511	(2.5)	513	(2.6)
Slowenien	508	(2.2)	512	(1.9)
Estland	517	(2.3)	522	(2.7)
Slowakische Republik	472	(3.6)	478	(3.0)
Australien	491	(2.5)	497	(2.1)
Türkei	418	(4.9)	423	(4.6)
Frankreich	490	(2.6)	496	(2.9)
Tschechische Republik	489	(2.8)	496	(3.3)
Mexiko	404	(2.4)	412	(2.7)
OECD-Durchschnitt	486	(0.5)	494	(0.6)
Ungarn	473	(3.0)	481	(3.6)
Israel	466	(4.0)	474	(5.4)
Vereinigte Staaten	465	(3.4)	474	(3.6)
Neuseeland	491	(2.7)	499	(3.4)
Kanada	511	(2.6)	520	(2.9)
Dänemark	506	(2.8)	516	(2.5)
Portugal	487	(2.7)	497	(3.0)
Luxemburg	480	(2.0)	491	(2.0)
Polen	499	(2.8)	510	(2.8)
Vereinigtes Königreich	487	(3.1)	498	(2.9)
Schweiz	515	(3.5)	527	(3.2)
Japan	525	(3.1)	539	(3.8)
Belgien	500	(2.8)	514	(3.1)
Spanien	478	(2.8)	494	(2.4)
Irland	495	(2.4)	512	(3.0)
Deutschland	498	(3.0)	514	(3.5)
Chile	413	(3.0)	432	(3.1)
Italien	480	(3.4)	500	(3.5)
Österreich	483	(3.6)	510	(3.8)

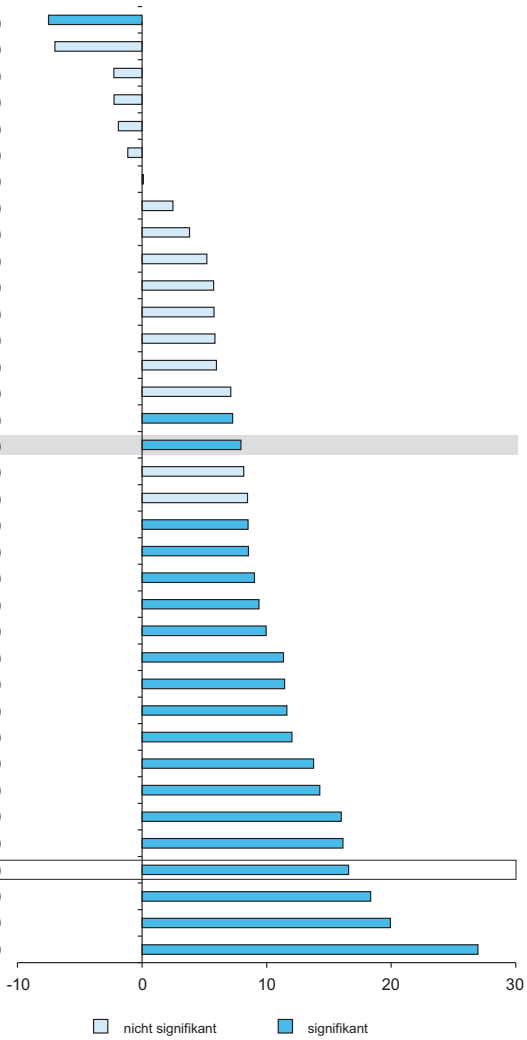


Abbildung 6.4: Mittelwerte mathematischer Kompetenz von Mädchen und Jungen in den OECD-Staaten

Jungen gleiche mathematische Kompetenzen. Auch in Estland, einem OECD-Staat der Spitzengruppe, scheint die Förderung von Mädchen und Jungen gleich gut zu gelingen. Es sei angemerkt, dass auch unter Einbeziehung der OECD-Partner bis auf Japan alle teilnehmenden asiatischen Bildungssysteme in dieser Gruppe sind, bei der sich Mädchen und Jungen in der Leistung nicht signifikant unterscheiden. Finnland ist der einzige OECD-Staat, in dem die Mädchen über eine signifikant höhere mathematische Kompetenz verfügen als die Jungen. In der letzten PISA-Erhebungsrunde mit Mathematik

© Waxmann Verlag GmbH

als Hauptdomäne (PISA 2012) hatten sich Mädchen und Jungen in Finnland bezüglich ihrer mathematischen Kompetenz nicht signifikant voneinander unterschieden.

6.3 Vertiefende Analysen der mathematischen Kompetenz in Deutschland

In den folgenden Abschnitten werden differenzierte Ergebnisse zur mathematischen Kompetenz der fünfzehnjährigen Schülerinnen und Schüler in Deutschland dargestellt. Es werden zu Beginn Befunde zu Unterschieden zwischen den nicht gymnasialen Schularten und dem Gymnasium sowie zwischen Mädchen und Jungen hinsichtlich der mathematischen Kompetenz präsentiert. Anschließend wird aufgezeigt, wie sich die mathematische Kompetenz seit PISA 2003, als Mathematik erstmalig die Hauptdomäne war, entwickelt hat.

6.3.1 Unterschiede zwischen Schularten

Die Sekundarstufe I umfasst in Deutschland unterschiedliche Schularten, die in den verschiedenen Bundesländern nicht einheitlich ausgelegt sind. So wurden die Haupt- und Realschulen in vielen Ländern der Bundesrepublik durch kombinierte Schularten wie etwa Schulen mit mehreren Bildungsgängen oder Integrierte Gesamtschulen ersetzt (Autorengruppe Bildungsberichterstattung, 2016). Vor diesem Hintergrund werden in diesem Berichtsband die Ergebnisse für die vier Schularten Hauptschule, Realschule, Integrierte Gesamtschule und Schule mit mehreren Bildungsgängen zusammengefasst und damit für alle *nicht gymnasialen Schularten* gemeinsam sowie für die *Gymnasien* dargestellt. Außerdem können *Veränderungen* der Testleistung in Mathematik seit den letzten PISA-Erhebungen für die einzelnen nicht gymnasialen Schularten kaum sinnvoll interpretiert werden. Die Befunde der fünfzehnjährigen Schülerinnen und Schüler an Berufsschulen und Förderschulen werden nicht separat präsentiert, da die entsprechenden Stichprobengrößen zu gering sind (vgl. Kapitel 1).

Mathematische Kompetenz: Mittelwerte und Streuungen im Vergleich

Die Mittelwerte und Standardabweichungen der mathematischen Kompetenz von Schülerinnen und Schülern an verschiedenen Schularten sind in Tabelle 6.3 dargestellt. Der erreichte Mittelwert an nicht gymnasialen Schularten unterscheidet sich erwartungsgemäß von dem an Gymnasien. Jugendliche am Gymnasium erreichen durchschnittlich 573 Punkte, der Mittelwert aller nicht gymnasialen Schularten liegt mit 476 Punkten deutlich unter dem der Gymnasien. Der mittlere Leistungsunterschied zwischen Gymnasien und nicht gymnasialen Schularten entspricht im Bereich Mathematik einem Unter-

© Waxmann Verlag GmbH

Tabelle 6.3: Mittelwerte und Standardabweichungen nach Schulform und für die Gesamtstichprobe

Schulart	*n*	*M*	*(SE)*	*SD*	*(SE)*
Nicht gymnasiale Schularten	3944	476	(3.0)	76	(1.4)
Gymnasium	2266	573	(3.1)	69	(1.9)
Gesamtstichprobe	6504	506	(2.9)	89	(1.4)

schied von etwa eineinhalb Kompetenzstufen (97 Punkte). Generell ist bei der Interpretation der Ergebnisse für die jeweiligen Schularten die Bildungsbeteiligung, also der prozentuale Anteil fünfzehnjähriger Jugendlicher an den nicht gymnasialen Schularten und an den Gymnasien, zu berücksichtigen. Ordnet man beispielsweise den Mittelwert der Gymnasien international ein, muss beachtet werden, dass an dieser Stelle Leistungen einer besonders leistungsstarken Teilgruppe von 33 Prozent der Schülerinnen und Schüler in Deutschland mit denen aller Fünfzehnjähriger in anderen Staaten verglichen werden. Die Bildungsbeteiligung ist in Abbildung 6.5 mit angegeben.

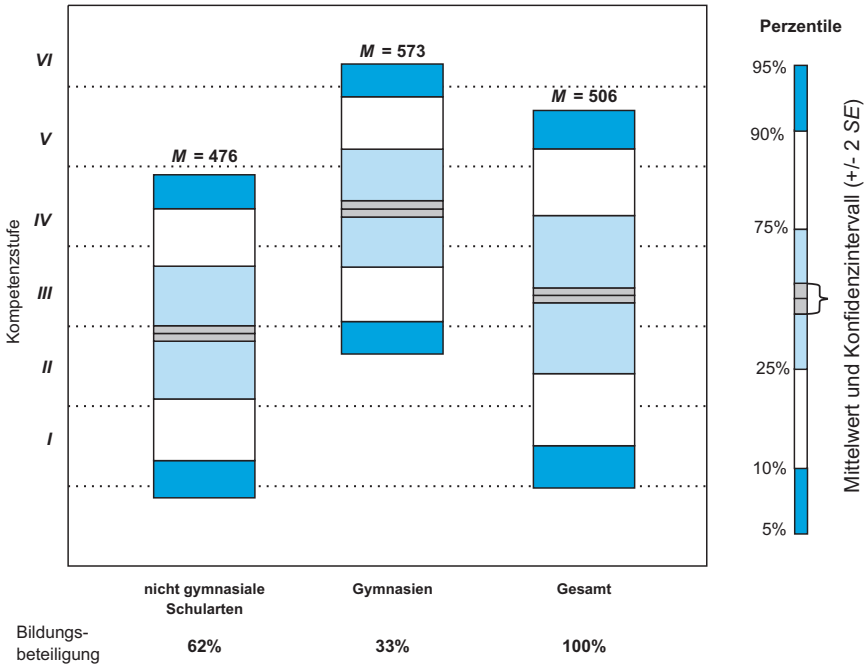


Abbildung 6.5: Perzentilbänder mathematischer Kompetenz in Deutschland nach Schulart und für die Gesamtstichprobe

© Waxmann Verlag GmbH

Neben den Unterschieden in den Mittelwerten zwischen den Schularten zeigen sich auch unterschiedlich große Streuungen der mathematischen Kompetenz. Die Streuung der mathematischen Kompetenz ist in den nicht gymnasialen Schularten (*SD* = 76) größer als am Gymnasium (*SD* = 69), aber insgesamt geringer als die mittlere Streuung in den OECD-Staaten. Dies ist nicht verwunderlich, da verschiedene Bildungsgänge zusammengefasst wurden. Auch strukturell sowie konzeptuell wird hier über die 16 Länder hinweg ein breiteres Spektrum – insbesondere im Vergleich zum Gymnasium – repräsentiert.

Die jeweils großen Streuungen der mathematischen Kompetenz *innerhalb* der Schularten gehen mit einer Überlappung der Kompetenzverteilungen zwischen den beiden Gruppen von Schularten einher. Dies wird in Abbildung 6.5 anhand von Perzentilbändern verdeutlicht. Es werden Gymnasien und nicht gymnasiale Schularten einander gegenübergestellt. An nicht gymnasialen Schularten übertreffen 10 Prozent der Schülerinnen und Schüler den Mittelwert, der für das Gymnasium festgestellt wurde. Außerdem verfügt das leistungsstärkste Viertel der Schülerinnen und Schüler an nicht gymnasialen Schularten über eine höhere mathematische Kompetenz als das leistungsschwächste Viertel der Jugendlichen am Gymnasium. Der mittlere Leistungsbereich an den nicht gymnasialen Schularten liegt in den Kompetenzstufen II und III. Außerdem erreichen 95 Prozent der Jugendlichen an diesen Schularten maximal die Kompetenzstufe IV. Umgekehrt verfügen 95 Prozent der Gymnasiastinnen und Gymnasiasten mindestens über eine mathematische Kompetenz im oberen Bereich der Kompetenzstufe II. Die Leistungen der Schülerinnen und Schüler im mittleren Bereich des Leistungsspektrums am Gymnasium liegen vorwiegend im Bereich der Kompetenzstufe IV.

Verteilung auf die Stufen der mathematischen Kompetenz

Für eine weitere Beurteilung der erreichten Kompetenzniveaus in den untersuchten Schularten wird der prozentuale Anteil der Schülerinnen und Schüler auf den Stufen mathematischer Kompetenz betrachtet (vgl. Abbildung 6.6). Hierbei sind vor allem die untersten und obersten Stufen von Interesse. In Deutschland scheitern etwa 17 Prozent der Fünfzehnjährigen an den Anforderungen der Kompetenzstufe II. Schülerinnen und Schüler, deren Mathematikkompetenz auf der Stufe I oder darunter liegt, verfügen lediglich über elementares mathematisches Wissen. Bestenfalls können sie aus einer einzigen Quelle relevante Informationen entnehmen und einfache Algorithmen, Formeln, Verfahren oder Regeln zur Lösung von Aufgaben mit ganzen Zahlen anwenden. Jugendliche auf dieser Kompetenzstufe können beispielsweise mit großer Wahrscheinlichkeit keine Aufgaben lösen, die die Verwendung von Darstellungen oder deren Vergleich erfordern. Am Gymnasium beträgt der Anteil Jugendlicher auf Kompetenzstufe I oder darunter nur 2 Prozent. Der entsprechende Anteil fällt in den nicht gymnasialen Schularten mit 23 Prozent sehr viel höher aus. Knapp jeder vierte Jugendliche an nicht gymnasialen Schularten hat demnach nur eine sehr geringe mathematische Kompetenz. Es ist davon auszugehen, dass dies mit Schwierigkeiten in Ausbildung und beruflicher Karriere sowie

© Waxmann Verlag GmbH

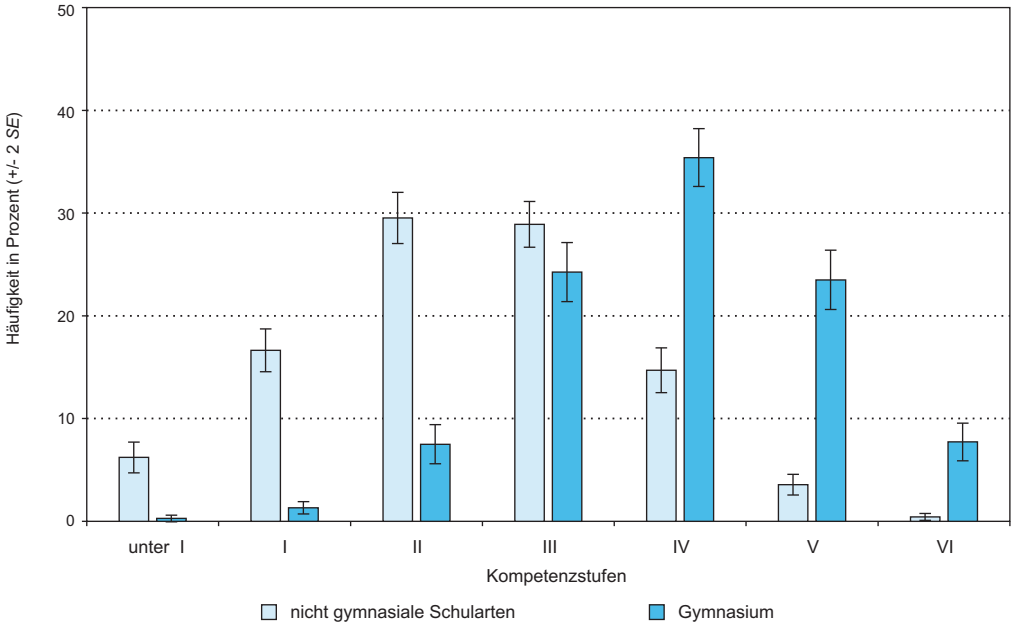


Abbildung 6.6: Prozentuale Anteile von Schülerinnen und Schülern auf den Stufen mathematischer Kompetenz in Deutschland nach Schulart

bei der Teilhabe am gesellschaftlichen Leben einhergeht. Demzufolge besteht besonders an nicht gymnasialen Schularten Handlungsbedarf, weitere Anstrengungen zu unternehmen, um die Gruppe leistungsschwächerer Schülerinnen und Schüler zu reduzieren und die betreffenden Jugendlichen auf ein Kompetenzniveau zu heben, das eine günstige Prognose für einen erfolgreichen Übergang in Ausbildung und Erwerbsleben erlaubt.

Schülerinnen und Schüler, welche die höchsten Kompetenzstufen (V und VI) erreichen, sind in der Lage, komplexe Problemstellungen zu lösen, Modelle für die Problemsituation zu konzipieren und flexibel damit zu arbeiten. Sie können geeignete Strategien heranziehen und das Ergebnis deuten sowie interpretieren. Sie verfügen über gut ausgebildete Denk- und Argumentationsfähigkeiten und reflektieren über ihr eigenes Handeln. Es kann davon ausgegangen werden, dass diese Jugendlichen sehr gut auf Anforderungen in ihrer Ausbildung und beruflichen Laufbahn vorbereitet sind, die eine fundierte mathematische Kompetenz erfordern. Im Gymnasium lassen sich 31 Prozent diesem Kompetenzniveau zuordnen, in den nicht gymnasialen Schularten sind es 4 Prozent.

© Waxmann Verlag GmbH

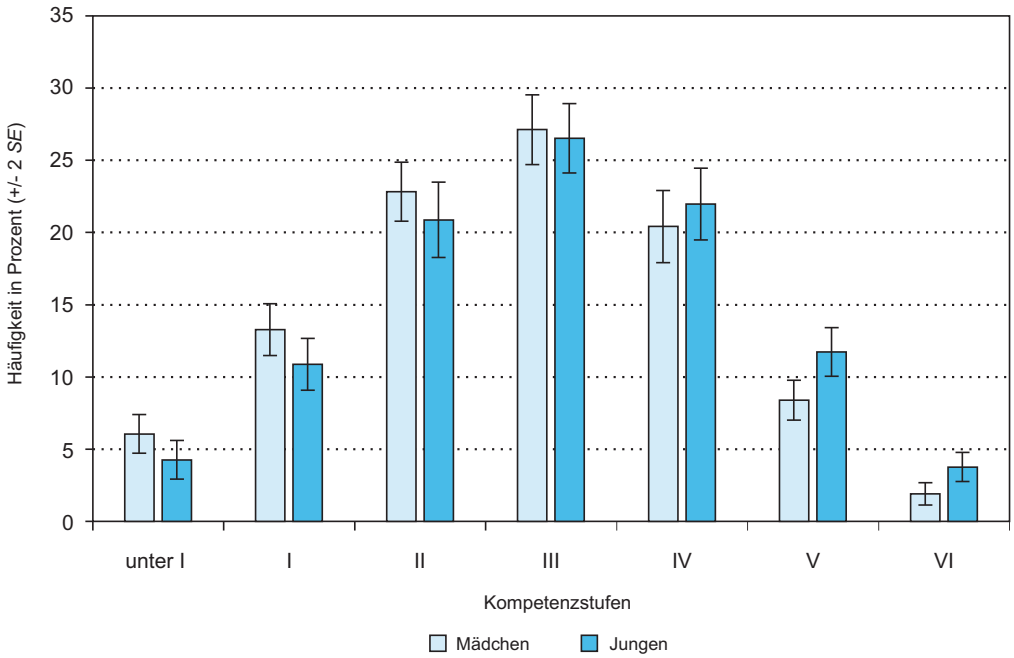


Abbildung 6.7: Prozentuale Anteile der Mädchen und Jungen auf den Stufen der mathematischen Kompetenz in Deutschland

6.3.2 Unterschiede zwischen Mädchen und Jungen

Wie in PISA 2012 erzielen Mädchen im aktuellen PISA-Mathematiktest signifikant niedrigere Werte als Jungen. Mädchen erreichen 498 Punkte (*SD* = 88) und Jungen 514 Punkte (*SD* = 90). Die Streuungen der beiden Gruppen zeigen aber, dass es eine nicht unerhebliche Zahl von Mädchen gibt, deren Leistung über derjenigen von Jungen liegt.

Betrachtet man in einer differenzierten Analyse die prozentualen Anteile von Mädchen und Jungen auf den Stufen der mathematischen Kompetenz, so sind ebenfalls Geschlechterdifferenzen zu erkennen. Auf Stufe I und darunter befinden sich deutschlandweit insgesamt 19 Prozent Mädchen und 15 Prozent Jungen. Auf den beiden höchsten Kompetenzstufen V und VI ist der Anteil der Jungen (16 Prozent) signifikant größer als derjenige der Mädchen (10 Prozent) (vgl. Abbildung 6.7).

6.3.3 Veränderung der mathematischen Kompetenz seit PISA 2003

Die Daten aus PISA ermöglichen eine Beschreibung, wie sich die mathematische Kompetenz in den letzten Jahren verändert hat. Solche Trendaussagen sind besonders vor dem Hintergrund der vielfältigen Maßnahmen zur Sicherung und Weiterentwicklung der Unterrichtsqualität von Interesse. Referenz für die Trendbetrachtung in PISA ist

© Waxmann Verlag GmbH

immer diejenige Erhebungsrunde, in der ein Kompetenzbereich zum letzten Mal Hauptdomäne war – für die mathematische Kompetenz ist dies PISA 2012. Für einen umfassenderen Eindruck der Entwicklung wird teilweise auch ein Vergleich zum Jahr 2003 hergestellt.

Über alle OECD-Staaten hinweg zeigt sich bei PISA 2015 keine signifikante Veränderung der mathematischen Kompetenz im Vergleich zu PISA 2012 und zu PISA 2003. Auf der Ebene der einzelnen OECD-Staaten haben sich aber durchaus bedeutsame Veränderungen ergeben, die im Folgenden dargestellt werden.

Signifikante *Zuwächse* gegenüber PISA 2012 sind in vier OECD-Staaten zu beobachten: Slowenien (+9 Punkte), Dänemark (+11 Punkte), Norwegen (+12 Punkte) und Schweden (+16 Punkte). Dänemark zählte dabei 2012 zu den OECD-Staaten mit überdurchschnittlichen Leistungen in Mathematik, Norwegen lag im Bereich des OECD-Durchschnitts und Schweden deutlich darunter. In sechs der 35 OECD-Staaten zeigen sich in PISA 2015 im Vergleich zu 2012 signifikante *Verringerungen* der mathematischen Kompetenz. Das größte Absinken mit bis zu knapp 30 Punkten haben Korea und die Türkei zu verzeichnen. Korea als der zweitstärkste OECD-Staat konnte in PISA 2015 den Zugewinn von 2003 auf 2012 nicht weiter ausbauen und erreicht in PISA 2015 (524 Punkte) sogar einen Wert, der 18 Punkte unter dem von 2003 liegt. Auch die niederländischen Schülerinnen und Schüler liegen bei PISA 2015 weiterhin über dem OECD-Durchschnitt, erzielen aber sowohl im Vergleich zu 2012 als auch zu 2003 signifikant niedrigere Werte. Seit 2003 ist die mathematische Kompetenz in den Niederlanden um 26 Punkte gesunken. Signifikante Verringerungen der mathematischen Kompetenz finden sich zudem in Australien, Polen und den Vereinigten Staaten. Im Gegensatz dazu sind Japan und die Schweiz Beispiele dafür, dass es durchaus möglich ist, die mathematische Kompetenz *unverändert hoch* zu halten. Der erzielte Wert im Mathematiktest in diesen beiden Ländern ist seit 2003 nahezu konstant.

Auch in Deutschland finden sich keine signifikanten Veränderungen gegenüber 2012. Die Entwicklung der mathematischen Kompetenz der Fünfzehnjährigen in Deutschland seit PISA 2003 ist in Abbildung 6.8 dargestellt. Schülerinnen und Schüler in Deutschland erreichten im Mathematiktest der PISA-Erhebung 2003 einen Wert, der im Bereich des OECD-Durchschnitts lag. In PISA 2006 betrug der Mittelwert der mathematischen Kompetenz in Deutschland 504 Punkte und liegt seitdem bei jeder PISA-Erhebung – auch jetzt 2015 – signifikant über dem OECD-Durchschnitt. Im Zeitraum von 2003 bis 2012 verbesserte sich die mathematische Kompetenz der Schülerinnen und Schüler in Deutschland insgesamt signifikant. Die Veränderungen zwischen den Kohorten aufeinanderfolgender PISA-Erhebungsrunden, also zum Beispiel von PISA 2006 zu PISA 2009, waren jedoch nicht signifikant. Es handelt sich von Erhebungsrunde zu Erhebungsrunde also um Veränderungen, die noch im Bereich des Zufalls liegen, und erst der Vergleich von Hauptdomäne zu Hauptdomäne (2003 bis 2012) weist eine ausreichend große Differenz auf, um von einer bedeutsamen Verbesserung zu sprechen. In der aktuellen PISA-Erhebungsrunde 2015 ist die mathematische Kompetenz in Deutschland nun im Vergleich zu PISA 2012 um acht Punkte gesunken. Da diese Verringerung jedoch

© Waxmann Verlag GmbH

ebenfalls nicht statistisch signifikant ist, kann nicht auf eine bedeutsame Veränderung geschlossen werden. Die beobachtete Veränderung bewegt sich im Bereich dessen, was auch rein zufällig auftreten könnte.

Insgesamt bilden die Veränderungen seit 2003 in Deutschland eine Stabilisierung der mathematischen Kompetenz auf einem gegenüber dem OECD-Mittelwert überdurchschnittlichen Niveau ab. Dass dies keineswegs selbstverständlich ist, zeigt der Vergleich mit Staaten wie Finnland, deren PISA-Kohorte über denselben Zeitraum substanziell an Kompetenz eingebüßt hat. Ebenso positiv herauszustellen ist die veränderte Leistungsstreuung in Deutschland – diese ist 2015 (*SD* = 89) im Vergleich zu 2003 (*SD* = 103) um 14 Punkte und gegenüber 2012 (*SD* = 96) um 7 Punkte bedeutsam gesunken. Folglich ist die Leistungsverteilung der Fünfzehnjährigen inzwischen homogener und der Unterschied der mathematischen Kompetenz zwischen leistungsschwachen und leistungsstarken Schülerinnen und Schülern geringer geworden.

Dass die nominelle Punktzahl auf der Skala mathematischer Kompetenz gegenüber 2012 und 2003 leicht zurückgegangen ist, findet sich sowohl bei den Mädchen als auch bei den Jungen in Deutschland. Der Betrag dieser Differenz ist für beide Geschlechter statistisch nicht bedeutsam und muss daher auch nicht als Absinken interpretiert werden. Es deutet sich jedoch ein Trend an, dass sich die Leistungsdifferenz zwischen Mädchen und Jungen eher vergrößert (von 14 auf 17 Punkte). Auch diese Veränderung kann nicht als statistisch bedeutsam abgesichert werden. Trotzdem gilt es diesen Bereich besonders sorgfältig im Blick zu behalten, insbesondere vor dem Hintergrund, dass sich die Gesamtstreuung der mathematischen Kompetenz in Deutschland seit den letzten PISA-Erhebungen verringert hat – entsprechend mehr fällt die Schere zwischen Mädchen und Jungen damit ins Gewicht und deutet auf entsprechenden Handlungsbedarf hin.

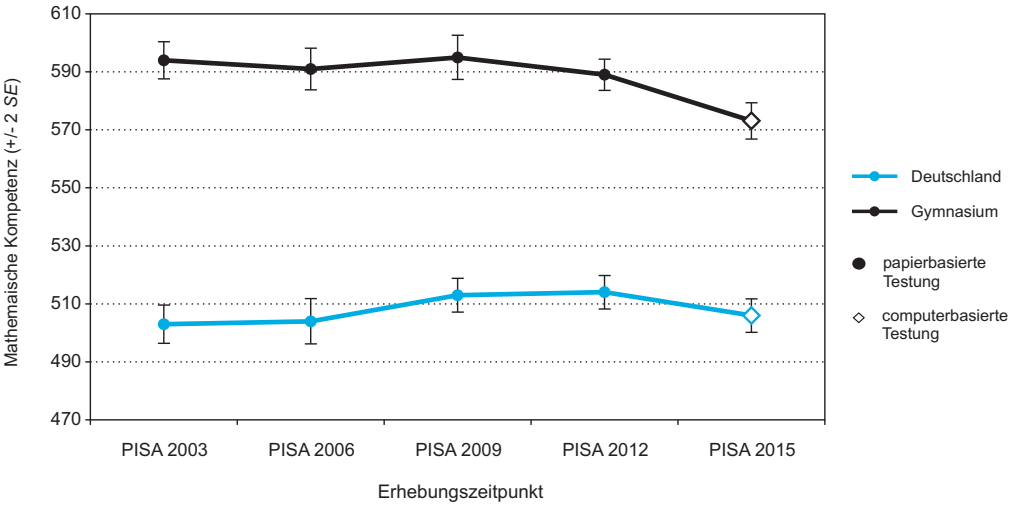


Abbildung 6.8: Mathematische Kompetenz in Deutschland und an Gymnasien von PISA 2003 bis PISA 2015

© Waxmann Verlag GmbH

Die Veränderungen der mittleren mathematischen Kompetenz sind mit einer veränderten Verteilung der Schülerinnen und Schüler auf die Kompetenzstufen verbunden. Vor diesem Hintergrund wird nun in einer vertiefenden Analyse darauf eingegangen, in welchen Bereichen des Leistungsspektrums Änderungen der mathematischen Kompetenz stattgefunden haben. Dabei sind wieder besonders die unteren und oberen Bereiche der Kompetenzverteilung von Interesse. Der Anteil der Jugendlichen auf Kompetenzstufe I oder darunter ist in der aktuellen Erhebung im Vergleich zu 2012 minimal (um einen halben Prozentpunkt, im Vergleich zu 2003 um 4 Prozentpunkte) gesunken. Diese Verringerungen sind aber statistisch nicht gegen den Zufall abzusichern, der Anteil leistungsschwacher Schülerinnen und Schüler ist demnach in den letzten Jahren etwa konstant geblieben. Demgegenüber ist der Anteil der Fünfzehnjährigen in Deutschland im oberen Leistungsspektrum (Kompetenzstufe V und VI) in der aktuellen PISA-Erhebung im Vergleich zu PISA 2012 um 5 Prozentpunkte signifikant gesunken und beträgt nunmehr 13 Prozent. Insgesamt zeigen die Ergebnisse, dass noch immer Handlungsbedarf in Bezug auf diese beiden Leistungsgruppen besteht: In den letzten Jahren konnte weder der Anteil der Leistungsschwächeren verringert noch derjenige der Leistungsstärkeren ausgebaut werden. Beides – die Verringerung der Gruppe leistungsschwacher und die Vergrößerung der Gruppe leistungsstarker Schülerinnen und Schüler – war bereits in früheren PISA-Erhebungsrunden als zentrales Ziel der Förderung formuliert worden, wobei sich bei PISA 2012 stärkerer Bedarf bei der Förderung von Talenten manifestiert hatte.

In diesem Zusammenhang stellt sich die Frage, inwieweit in den gegliederten Schulsystemen Deutschlands insbesondere das Gymnasium die Möglichkeit nutzt, die Leistungsstarken und potenziell Leistungsfähigen bestmöglich zu fördern. Der für den Mathematiktest gemessene Mittelwert der Gymnasien beträgt 573 Punkte und ist damit seit PISA 2012 (589 Punkte) um 16 Punkte und seit PISA 2003 (594 Punkte) um 21 Punkte erstmals signifikant gesunken. Diese Abnahme zeigt sich auch im Anteil der Fünfzehnjährigen am Gymnasium, welche die höchsten Kompetenzstufen (V und VI) erreichen. Dieser ist 2015 um 9 Prozentpunkte bedeutsam geringer als noch 2012 und sogar um 11 Prozentpunkte niedriger als 2003. Dies darf als deutlicher Hinweis darauf verstanden werden, dass gerade hoch leistungsfähige Jugendliche im Bereich Mathematik an Gymnasien adäquate Förderung brauchen.

Bemerkenswert ist insgesamt, dass nach mehreren PISA-Runden mit einem leichten, zwischen 2003 und 2012 signifikanten Anstieg der mathematischen Kompetenzen im Jahr 2015 ein leichtes, wenn auch nicht signifikantes Absinken der mittleren Leistungen beobachtbar ist. Ob diese Veränderung den Start eines Negativtrends bedeutet, ist schwer vorherzusagen, hat sich doch am Erhebungsmodus zwischen 2012 und 2015 einiges geändert, was auch ergebnisrelevant gewesen sein könnte. Der Mathematiktest in der aktuellen PISA-Erhebung war nicht mehr wie in den früheren Erhebungen papier-, sondern computerbasiert. Die Aufgaben, welche die Schülerinnen und Schüler am Computer bearbeiteten, wurden allerdings eins-zu-eins vom Papierformat auf das Bildschirmformat übertragen und bereits in einer vergangenen PISA-Erhebungsrunde erprobt. Für

© Waxmann Verlag GmbH

Deutschland liefert der Feldtest zu PISA 2015 (im Frühjahr 2014) aber Hinweise darauf, dass die PISA-Aufgaben im Mittel am Computer schwieriger waren als auf Papier (vgl. Kapitel 12). Insofern könnte die computerbasierte Testung in PISA 2015 in Deutschland dazu beigetragen haben, dass sich der bis PISA 2012 abzeichnende Positivtrend in PISA 2015 zunächst nicht fortsetzt. Der leichte Rückgang der Punkte gegenüber PISA 2012 ist nicht signifikant; der Leistungsstand darf daher als unverändert gelten. In den kommenden PISA-Runden wird das Testmedium Computer weiterhin eingesetzt, sodass der Darbietungsmodus dann keine Neuerung mehr ist und etwaige Kompetenzveränderungen spezifischer beschrieben werden können.

6.4 Zusammenfassung und Diskussion

Schülerinnen und Schüler in Deutschland erreichen bei PISA 2015 mit 506 Punkten erneut eine gegenüber dem OECD-Mittelwert überdurchschnittlich hohe mathematische Kompetenz. Seit PISA 2012 hat sich dieser Wert nicht signifikant verändert, das heißt, es ist Deutschland als einem von wenigen Bildungssystemen gelungen, die hohe mittlere mathematische Kompetenz zu festigen. Im internationalen Vergleich positionieren sich wie schon 2012 fünf OECD-Staaten signifikant höher als Deutschland, was auf ein Entwicklungspotenzial nach oben hinweist. Im Vergleich zu 2003 hat sich die Leistungsstreuung in Deutschland deutlich reduziert, sodass die Schere zwischen leistungsschwachen und leistungsstarken Schülerinnen und Schülern kleiner geworden ist.

Zu bedenken ist jedoch, dass weiterhin eine signifikante Geschlechterdifferenz in der mathematischen Kompetenz vorhanden ist, die in den letzten Jahren nicht verringert werden konnte. Im Gegenteil deutet sich ein Trend an, dass sich dieser Unterschied eher vergrößert. Deutschland gehört 2015 zu den sechs OECD-Staaten mit der größten Geschlechterdifferenz. Maßnahmen zur besseren Förderung der mathematischen Kompetenz von Mädchen haben offenbar bislang noch nicht erfolgreich gegriffen. Vor allem in Anbetracht der Tatsache, dass sich in vielen OECD-Staaten keine Geschlechterdifferenzen in der mathematischen Kompetenz feststellen lassen, besteht in Deutschland Handlungsbedarf, laufende Maßnahmen zu überdenken und zu ergänzen.

Für den Erfolg des Bildungssystems eines Staates ist es entscheidend, den Anteil derjenigen Schülerinnen und Schüler auszubauen, die eine hohe oder sehr hohe mathematische Kompetenz aufweisen und gleichzeitig den Anteil Leistungsschwacher bestmöglich zu reduzieren. Der Anteil der Schülerinnen und Schüler, deren Kompetenz die Stufe I nicht übersteigt, konnte seit den ersten PISA-Erhebungen bis 2012 verringert werden. Dieser Trend setzt sich in PISA 2015 jedoch nicht fort: Der Anteil der leistungsschwachen Schülerinnen und Schüler ist mit 17 Prozent nach wie vor substanziell und konnte seit 2012 und auch im Vergleich zu 2003 nicht signifikant reduziert werden. Ergebnisse in anderen Staaten (beispielsweise Japan und Estland) machen deutlich, dass es durchaus möglich ist, eine überdurchschnittlich hohe mittlere mathematische Kompetenz zu erzielen, ohne die Leistungsschwachen zu vernachlässigen. Eine der Haupt-

© Waxmann Verlag GmbH

herausforderungen für die nächsten Jahre wird es sein, die Größe dieser Gruppe weiter zu reduzieren, damit für möglichst viele Jugendliche die Barrieren für den Einstieg in das Berufsleben und die Teilhabe am gesellschaftlichen Leben klein bleiben. Demgegenüber stehen in Deutschland 13 Prozent der Fünfzehnjährigen, die über eine mathematische Kompetenz auf mindestens Kompetenzstufe V verfügen. Im OECD-Durchschnitt sind dies nur 11 Prozent. In den Niederlanden waren es hingegen 16 Prozent, in der Schweiz 19 Prozent, in Japan 20 Prozent und Korea 21 Prozent. Diese Werte und insbesondere die signifikante Abnahme im Vergleich zu 2012 um 5 Prozentpunkte machen deutlich, dass die Möglichkeit, die Leistungsspitze bestens zu fördern, in Deutschland noch nicht völlig ausgeschöpft wird. Insbesondere am Gymnasium ist die Situation in Bezug auf die Mathematik verbesserungsfähig: Die durchschnittliche mathematische Leistung insbesondere an dieser Schulart ist in den letzten Jahren kontinuierlich gesunken, die Leistungsspitze ist kleiner geworden. In PISA 2015 schaffen es nur noch 31 Prozent der Gymnasiastinnen und Gymnasiasten, die Anforderungen der oberen beiden Kompetenzstufen erfolgreich zu bewältigen, während es 2012 noch 40 Prozent und 2003 sogar 42 Prozent waren. Es sei angemerkt, dass die Bildungsbeteiligung hier keine Rolle spielen dürfte, denn 2015 waren in der Stichprobe 33 Prozent Gymnasiastinnen und Gymnasiasten, 2012 lag ihr Anteil bei 36 Prozent und 2003 bei 31 Prozent der Fünfzehnjährigen. Auch wenn Schülerinnen und Schüler am Gymnasium über eine deutlich höhere mathematische Kompetenz verfügen als Fünfzehnjährige an nicht gymnasialen Schularten, so darf die Talentförderung nicht aus den Augen verloren werden.

Insgesamt betrachtet ist anzunehmen, dass sich dieser beschriebene Rückgang der Spitzengruppe in Mathematik nicht mit der Umstellung von einem neunjährigen Gymnasium auf ein achtjähriges erklären lässt – dies hätte sich schon bei PISA 2012 und mindestens von der Tendenz her schon 2009 zeigen müssen. Diese Annahme kann durch Ergebnisse empirischer Studien bekräftigt werden, die keine Leistungsunterschiede zwischen den beiden Varianten zeigen (z. B. Trautwein et al., 2015). Weniger die Debatte und öffentliche Diskussion über G8/G9, also darüber wie lange die Schülerinnen und Schüler am Gymnasium sind, sollte im Vordergrund stehen, wenn es um die Leistungsentwicklung, erfolgreiche Lernprozesse und die Motivation der Schülerinnen und Schüler geht. Wesentlicher ist es, die Qualität des Unterrichts weiterzuentwickeln und die dafür eingesetzten Maßnahmen systematisch zu evaluieren. Mit den Ergebnissen von PISA 2015 wird deutlich, dass dabei eine Priorität die Förderung der Leistungsspitze sein muss.

Bildungsgerechtigkeit ist auch in dem Sinn zu verstehen, dass Schülerinnen und Schüler mit Talent ihr Potenzial entfalten können. Mit der aktuellen PISA-Erhebung zeichnet sich wieder ab, dass es noch nicht gelingt, die Leistungsspitze bestmöglich zu fördern und diese zu vergrößern. Es müssen weitere Bemühungen unternommen werden, die Leistungsstarken und potenziell Leistungsfähigen durch Herausforderungen zu unterstützen sowie unterschiedliche Leistungsniveaus beispielsweise durch differenzierte Aufgabenstellungen und zusätzliche Lernangebote anzusprechen. Eine Förderung von Schülerinnen und Schülern mit hohem Potenzial, weit überdurchschnittliche Leistungen

© Waxmann Verlag GmbH

zu erzielen, ist dabei innerhalb und außerhalb des Unterrichts notwendig. Die bundesweite Strategie zur Förderung leistungsstarker Schülerinnen und Schüler ist ein Schritt in diese Richtung (KMK, 2015). In diesem Schreiben wird auf eine Vielzahl erprobter Möglichkeiten zur Begabtenförderung verwiesen. Dazu gehört beispielsweise das schnellere Durchlaufen der Schulzeit etwa durch Überspringen einer Klasse oder einzelner Fächer (*Akzeleration*). Des Weiteren gibt es unterschiedliche Möglichkeiten das Lernangebot für besonders Interessierte und Begabte zu erweitern, indem zum Beispiel zusätzliche Lernmaterialien zum regulären Curriculum zur Verfügung gestellt werden (*Enrichment*).

Auch den Universitäten kommt dabei in der Lehrerausbildung und in der Forschung eine entscheidende Rolle zu. Für eine erfolgreiche Begabtenförderung braucht es qualifizierte Lehrkräfte, die über die professionelle Kompetenz verfügen, besonders begabte Schülerinnen und Schüler zu erkennen und schließlich diese Talente wirksam zu fördern.

Literatur

Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: Bertelsmann.

Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller, O. & Neubrand, J. (1997). *TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich: Deskriptive Befunde*. Opladen: Leske + Budrich. http://dx.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, R. Pekrun, H.-G. Rolff, J. Rost & U. Schiefele (Hrsg.), *PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland: Ergebnisse des zweiten internationalen Vergleichs* (S. 47–92). Münster: Waxmann.

Frey, A., Carstensen, C. H., Walter, O., Rönnebeck, S. & Gomolka, J. (2008). Methodische Grundlagen des Ländervergleichs. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006 in Deutschland. Die Kompetenz der Jugendlichen im dritten Ländervergleich* (S. 375–397). Münster: Waxmann.

Hartig, J. & Frey, A. (2012). Validität des Tests zur Überprüfung des Erreichens der Bildungsstandards in Mathematik. *Diagnostica*, 58 (1), 3–14. http://dx.doi.org/10.1026/0012-1924/a000064

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, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 139–190). Opladen: Leske + Budrich. http://dx.doi.org/10.1007/978-3-322-83412-6_6

KMK (2003) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2003). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss.* Bonn: KMK.

© Waxmann Verlag GmbH

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.* Zugriff am 05.09.2016. Verfügbar unter http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2010/2010_03_04-Foerderstrategie-Leistungsschwaechere.pdf

KMK (2012) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2012). *Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife.* Bonn, Berlin: KMK.

KMK (2015) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2003). *Förderstrategie für leistungsstarke Schülerinnen und Schüler. Beschluss der Kultusministerkonferenz vom 11.06.2015.* Zugriff am 05.09.2016. Verfügbar unter http://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 Mathematical Education* (S. 115–124). Athens: 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. Zugriff am 27.09.2016. Verfügbar unter http://milne.ruc.dk/imfufatekster/pdf/485web_b.pdf

OECD. (2003). *The PISA 2003 Assessment framework – mathematics, reading, science and problem solving knowledge and skills.* Paris: OECD.

OECD. (2013a). *PISA 2012 released mathematics items*. Zugriff am 27.09.2016. Verfügbar unter http://www.oecd.org/pisa/pisaproducts/pisa2012-2006-rel-items-maths-ENG.pdf

OECD. (2013b). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy.* Paris: OECD.

OECD. (2014). *PISA 2012 Ergebnisse: Was Schülerinnen und Schüler wissen und können: Schülerleistungen in Lesekompetenz, Mathematik und Naturwissenschaften* (Band 1, überarb. Aufl.). Bielefeld: Bertelsmann.

OECD. (2016). *PISA 2015 assessment and analytical framework: Mathematics, reading, science and financial literacy* (Band 1, überarb. Aufl.). Paris: OECD.

Prenzel, M., Friedrich, A. & Stadler, M. (2009). *Von Sinus lernen: wie Unterrichtsentwicklung gelingt*. Seelze: Kallmeyer.

Sälzer, C. (2016). *Studienbuch Schulleistungsstudien. Das Rasch-Modell in der Praxis.* Berlin, Heidelberg: Springer Spektrum. http://dx.doi.org/10.1007/978-3-662-45765-8

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.

Trautwein, U., Hübner, N., Wagner, W. & Kramer, J. (2015). *Konsequenzen der G8-Reform – Zusammenfassung zentraler Befunde.* Tübingen: Eberhard Karls Universität, Hector-Institut für Empirische Bildungsforschung.

© Waxmann Verlag GmbH

© Waxmann Verlag GmbH

7 Lesekompetenz in PISA 2015: Ergebnisse, Veränderungen und Perspektiven

Mirjam Weis, Fabian Zehner, Christine Sälzer, Anselm Strohmaier, Cordula Artelt & Maximilian Pfost

Die Ergebnisse der PISA-Studie 2015 zeigen, dass die Lesekompetenz der Jugendlichen in Deutschland signifikant höher ist als die durchschnittliche Lesekompetenz der Jugendlichen aller OECD-Staaten. Insgesamt befindet sich Deutschland im Vergleich mit den anderen OECD-Staaten im oberen Drittel der Rangreihenfolge. Die Gruppe der besonders leistungsstarken Schülerinnen und Schüler ist in Deutschland signifikant größer als im Durchschnitt der OECD-Staaten. Die Gruppe der sehr leistungsschwachen Schülerinnen und Schüler ist hingegen signifikant kleiner als im Durchschnitt der OECD-Staaten. Vertiefende Analysen zeigen eine höhere Lesekompetenz bei Schülerinnen und Schülern am Gymnasium als bei den Fünfzehnjährigen an nicht gymnasialen Schularten. Zudem verfügen Mädchen in Deutschland über eine höhere Lesekompetenz als Jungen, wobei sich dieser Geschlechterunterschied im Vergleich zu früheren Erhebungsrunden deutlich verringert hat. Zuletzt wurde die Lesekompetenz in der PISA-Studie 2009 als Hauptdomäne erhoben. Zu diesem Zeitpunkt lag die Lesekompetenz der Schülerinnen und Schüler in Deutschland noch im Durchschnitt der OECD-Staaten. Seitdem ist die Lesekompetenz in Deutschland deutlich gestiegen. Es sind vor allem die Jungen, die 2015 im Vergleich zur Erhebung im Jahr 2009 eine höhere Lesekompetenz zeigen. Außerdem ist die Gruppe der besonders leistungsstarken Schülerinnen und Schüler in Deutschland seit 2009 signifikant gewachsen.

In dem vorliegenden Kapitel liegt der Fokus auf der Frage, wie sich die Lesekompetenz der Fünfzehnjährigen im internationalen Vergleich und in Deutschland darstellt. Zudem ermöglicht die PISA-Studie eine Evaluation der Entwicklung der Lesekompetenz über die Zeit und damit auch die vorsichtige Abschätzung der Wirksamkeit von Maßnahmen zur Förderung der Lesekompetenz.

Die Effektivität von Bildungssystemen wird danach beurteilt, inwiefern es den teilnehmenden Staaten gelingt, Jugendliche auf das Erwachsenenleben und somit auf eine aktive Teilhabe am gesellschaftlichen Leben sowie auf lebenslanges Lernen vorzubereiten (vgl. Kapitel 1). Daher umfasst *Lesekompetenz* nicht nur die Fähigkeit, Informationen aus Texten zu entnehmen, sondern auch das Reflektieren über Texte sowie die Fähigkeit und Motivation, sich auf Texte einzulassen, sie zu nutzen und sich mit deren Inhal-

© Waxmann Verlag GmbH

ten auseinanderzusetzen (OECD, 2016a). Um diese breit gefächerte Lesekompetenz zu messen, wird sie bei PISA als komplexe Kompetenz mit mehreren Facetten beschrieben. So beinhalten die Testaufgaben sowohl verschiedene *Situationen* (z. B. private, bildungs- und berufsbezogene Situationen) als auch unterschiedliche *Textarten* (z. B. Erzählungen, Anleitungen). Das Lösen der Leseaufgaben erfordert außerdem verschiedene *kognitive Aspekte des Lesens* (z. B. Suchen und Extrahieren von Informationen, Reflektieren und Bewerten).

Relevanz der Lesekompetenz

Lesekompetenz ist für eine aktive Teilhabe am gesellschaftlichen Leben unverzichtbar. Der Erwerb der Lesekompetenz beschränkt sich in der Schule nicht auf das Unterrichtsfach Deutsch (Artelt et al., 2007). Vielmehr werden Aspekte des Reflektierens und Nutzens von Texten etwa bei der Quellenanalyse in den Fächern Geschichte, Ethik oder Religion, aber auch in naturwissenschaftlichen Domänen und der Mathematik geschult (KM Bayern, 2011). Während im frühen Leseunterricht der Schwerpunkt auf dem Erlernen des Lesens liegt, wird das Lesen später genutzt, um zu lernen (Allington, 2002; Chall, 1983). In allen Schulfächern wird Wissen zu einem großen Teil über Texte vermittelt. Wenn Lesekompetenz nicht lediglich als die Fähigkeit verstanden wird, Informationen aus Texten zu entnehmen, sondern Denkprozesse auf Basis von Texten sowie die Bereitschaft, sich mit Texten auseinanderzusetzen, einschließt, wird ihre Bedeutung für alle Schulfächer noch deutlicher. So gilt eine solch umfassend definierte Lesekompetenz beispielsweise als wichtiger Bedingungsfaktor für die Entwicklung mathematischer Kompetenzen (Schneider, Küspert & Krajewski, 2013) und als ein entscheidendes Werkzeug für den erfolgreichen Umgang mit mathematischen Anforderungssituationen (Prediger, Wilhelm, Büchter, Gürsoy & Benholz, 2015). Lesekompetenz wird letztendlich in allen Fächern benötigt, um Aufgaben und Problemstellungen verstehen und beurteilen zu können. Somit bildet Lesekompetenz fachübergreifend eine wesentliche Voraussetzung für erfolgreiches Lernen (Artelt et al., 2007). Dies gilt auch für das weitere Lernen nach der Schule, zum Beispiel in der Ausbildung, im Studium oder in der Fort- und Weiterbildung im Beruf (Heinze & Schnurr, 2010; Nickolaus, Geissel & Gschwendtner, 2008).

Im Hinblick auf eine aktive Teilhabe an der Gesellschaft ist Lesen unverzichtbar, weil darüber sowohl Informationen als auch Ideen, Wertvorstellungen und kulturelle Inhalte vermittelt werden (Artelt et al., 2007; Beck, von Dewitz & Titz, 2015). Es eröffnet den Zugang zu einer großen Vielfalt an kulturellen, sozialen und medialen Gelegenheiten, sodass Einsichten in die eigene und gesellschaftliche Entwicklung genauso wie der Ausblick auf andere Kulturen angeregt werden (KMK, 2004). Hinzu kommt, dass bei der Kommunikation im Privat- und Berufsleben oft gelesen werden muss (OECD, 2016c). Dies ist beim klassischen Briefwechsel ebenso der Fall wie bei der Nutzung moderner Kommunikationsmedien, also etwa beim Lesen von E-Mails, Textnachrichten, Chats oder Blogs im Internet. Ob als kulturelle Bereicherung, zum Wissenserwerb oder als

© Waxmann Verlag GmbH

Kommunikationsmittel – fest steht, dass das Lesen in unserer Gesellschaft in nahezu allen Bereichen eine unverzichtbare Rolle spielt. Neben der Bedeutsamkeit der Lesekompetenz für das Individuum spielt Lesekompetenz zudem eine wichtige Rolle für die gesellschaftliche und volkswirtschaftliche Entwicklung von Staaten (Coulombe, Tremblay & Marchand, 2004; UNESCO, 2005).

Beispiele für Maßnahmen zur Leseförderung seit PISA 2000

In Deutschland wurden im Anschluss an die PISA-Studie im Jahr 2000, als die Lesekompetenz der Fünfzehnjährigen in Deutschland unter dem Durchschnitt der OECD-Staaten lag (Artelt, Stanat, Schneider & Schiefele, 2001), umfangreiche Maßnahmen zur Förderung der Lesekompetenz ergriffen. So wurden im Jahr 2003 in mehreren Fächern bundesweit geltende Bildungsstandards von der Kultusministerkonferenz eingeführt. Die Bildungsstandards im Fach Deutsch sprechen der Lesekompetenz und dem Umgang mit Texten und Medien eine wichtige Rolle zu (KMK, 2004, 2005a, 2005b, 2014). Darüber hinaus definierte die KMK sieben zentrale Handlungsfelder, um gezielte Maßnahmen zur Qualitätsverbesserung im Bildungsbereich einzuleiten (KMK, 2002). Auch hier wurde der Förderung der Lesekompetenz mit den Handlungsfeldern „Maßnahmen zur Förderung der Sprachkompetenz“ sowie „Maßnahmen zur Verbesserung der Grundschulbildung und durchgängige Verbesserung der Lesekompetenz“ ein wichtiger Stellenwert eingeräumt (KMK, 2002). Seither wurden zahlreiche Projekte zur Förderung der Lesekompetenz in Deutschland initiiert und durchgeführt, zu einem großen Teil auf Ebene der Bundesländer. Als unmittelbare Reaktionen auf die PISA-Studie 2000 sind beispielsweise die Projekte *Leseinitiative Thüringen*, *Leselust in Rheinland-Pfalz* und *Bist du auch lesekalisch?* entstanden (Artelt et al., 2007). Bei letzterem handelt es sich um ein Projekt zur frühkindlichen Leseförderung in Nordrhein-Westfalen, bei welchem vor allem Kinder aus bildungsfernen Familien erreicht werden sollten. Kindergärten, Bibliotheken und andere Einrichtungen wie Arztpraxen schlossen sich zusammen, stellten Medienboxen mit Bilder- und Vorlesebüchern zur Verfügung und boten Fortbildungen zur Leseförderung an (Artelt et al., 2007). Länderübergreifend beteiligten sich zwischen 2008 und 2010 alle Bundesländer mit insgesamt ca. 140 Projektschulen an dem Projekt *ProLesen*, bei welchem die Leseförderung als eine Aufgabe aller Schulfächer gesehen wurde und durch fächerübergreifenden Projektunterricht sowie in Zusammenarbeit mit öffentlichen Bibliotheken umgesetzt wurde (KM Bayern, 2011). In öffentlichen Bibliotheken werden beispielsweise Gruppen von Kindern im Kindergarten- und Grundschulalter durch Bilder und Geschichten mit spielerischen Elementen zum eigenständigen Umgang mit Texten angeregt (Keller-Loibl & Brandt, 2015). Auch gegenwärtig werden Förderprogramme initiiert. So sei erwähnt, dass es seit 2015 eine umfassende deutschlandweite Maßnahme zur Lese- und Sprachförderung mit dem *BISS-Programm* (Bildung durch Sprache und Schrift) gibt, das eine gemeinsame Initiative von Bund und Ländern ist (Bundesministerium für Bildung und Forschung, 2015; Schneider et al., 2012). Diese Initiative ist als Forschungs- und Entwicklungsprogramm angelegt und hat das Ziel,

© Waxmann Verlag GmbH

deutschlandweit eine durchgängige Förderung der sprachlichen Kompetenzen von Kindern und Jugendlichen von der Kindertageseinrichtung bis zum Ende der Sekundarstufe zu ermöglichen. Hierbei werden Maßnahmen sowohl zur Leseförderung in Kindertagesstätten und Schulen als auch zur Fortbildung der Erzieherinnen und Erzieher sowie der Lehrkräfte entwickelt und wissenschaftlich überprüft (Bundesministerium für Bildung und Forschung, 2015). Eine Übersicht über aktuelle Informationen zur Leseförderung gibt die Webseite „Lesen in Deutschland – Projekte und Initiativen zur Leseförderung“[1].

Insgesamt scheinen die seit PISA 2000 eingeführten Maßnahmen in Deutschland erste Früchte zu tragen, denn die Lesekompetenz der deutschen Jugendlichen lag bei den PISA-Erhebungen 2003, 2006 und 2009 bereits im durchschnittlichen Bereich der OECD-Staaten und bei PISA 2012 sogar über dem Durchschnitt der OECD-Staaten (Hohn, Schiepe-Tiska, Sälzer & Artelt, 2013). Ein Blick in die Entwicklung der OECD-Staaten insgesamt zeigt, dass die durchgängig positive Entwicklung der Lesekompetenz bei den Schülerinnen und Schülern in Deutschland in allen Erhebungen keineswegs selbstverständlich war.

Im Folgenden wird betrachtet, wie hoch die Lesekompetenz der Fünfzehnjährigen in Deutschland im Jahr 2015 im Vergleich zu früheren PISA-Erhebungsrunden sowie im Vergleich zu anderen Staaten ist. Im ersten Teil wird die PISA-2015-Rahmenkonzeption der Lesekompetenz sowie deren Umsetzung im Lesekompetenztest und in den Lesekompetenzstufen erläutert. Im Anschluss werden die aktuellen Ergebnisse zur Lesekompetenz berichtet. Hierbei wird zuerst auf die Lesekompetenz im internationalen Vergleich und anschließend auf vertiefende Analysen innerhalb Deutschlands eingegangen. Am Ende des Kapitels wird diskutiert, welche Schlussfolgerungen aus den Ergebnissen der PISA-Studie 2015 für die weitere Förderung der Lesekompetenz in Deutschland gezogen werden können.

7.1 Lesekompetenz in PISA 2015

Unter Lesekompetenz wird in der PISA-Studie 2015 die Fähigkeit verstanden, geschriebene Texte zu verstehen, zu nutzen 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, 2016a). Diese Definition von Lesekompetenz bringt zum Ausdruck, dass es sich um eine sehr breite und komplexe Kompetenz handelt, die Teil einer Grundbildung ist. Lesekompetenz beinhaltet zum einen die Fähigkeit, relevante Informationen aus Texten herauszusuchen und zu extrahieren. Zum anderen gehört auch das Verstehen und Nutzen von Texten sowie das Reflektieren über Texte zur Lesekompetenz. Zudem wird die Fähigkeit und Moti-

1 Das Portal „Lesen in Deutschland – Projekte und Initiativen zur Leseförderung“ wird vom Deutschen Institut für Internationale Pädagogische Forschung (DIPF) herausgegeben und kooperiert mit dem Deutschen Bildungsserver sowie dem Innovationsportal. Das Portal kann unter http://www.lesen-in-deutschland.de/html/index.php aufgerufen werden.

© Waxmann Verlag GmbH

vation, sich auf Texte einzulassen und sich mit den Inhalten auseinanderzusetzen, in der Rahmenkonzeption der PISA-Studie als Teil der Lesekompetenz gesehen (OECD, 2016a). Schließlich drückt die Definition entsprechend der Leitfrage von PISA aus, dass Lesekompetenz auf eine Teilhabe an der Gesellschaft vorbereitet, indem sie Menschen sowohl befähigt, eigene Ziele zu erreichen als auch einen Beitrag für die Gesellschaft zu leisten (OECD, 2016a; Sälzer & Prenzel, 2013). Eigene Ziele können etwa sein, dass ein Bildungsabschluss erreicht oder ein Arbeitsplatz gefunden wird, aber auch, dass das Privatleben bereichert und lebenslanges Lernen vorbereitet wird. Ein Beitrag für die Gesellschaft kann durch soziales, kulturelles oder politisches Engagement geleistet werden (OECD, 2016a). Die Entwicklung dieser Kompetenz ist mit dem Erwerb der grundlegenden Lesefertigkeiten in den ersten Schuljahren noch nicht abgeschlossen, sondern der Erwerb komplexerer Lesekompetenz findet vor allem im Verlauf der weiteren Schulzeit und auch noch im Erwachsenenalter statt (Artelt et al., 2007; Beck et al., 2015).

Rahmenkonzeption der Domäne Lesen

In der PISA-Studie 2015 wurde die Lesekompetenz als Nebendomäne erfasst. Die theoretische Rahmenkonzeption zur Erfassung der Lesekompetenz wurde dabei aus der PISA-Studie 2009, in der Lesekompetenz zuletzt als Hauptdomäne erfasst wurde, übernommen. In der Rahmenkonzeption wird hinsichtlich (1) *Situationen*, (2) *Textarten* und (3) *kognitiven Aspekten des Lesens* unterschieden.

Als *Situationen* werden die Kontexte, in denen Lesen stattfindet, sowie die Zwecke des Lesens bezeichnet. Es geht dabei um Kontexte und Anlässe des Lesens und nicht nur um den Ort, an dem die Lesetätigkeit ausgeführt wird. Die Situationen werden danach kategorisiert, für welche Zielgruppe und welchen Zweck der Text geschrieben wurde. Hierbei werden vier Bereiche unterschieden:

- *Private Situationen* beschreiben Texte, die mit persönlichen Interessen und Inhalten zu tun haben. Dies umfasst unter anderem Texte, die zum Ziel haben, persönliche Beziehungen mit anderen Menschen aufrechtzuerhalten oder zu entwickeln. Zu diesen Texten gehören unter anderem persönliche Briefe, E-Mails und Romane.
- *Öffentliche Situationen* bezeichnen Texte, die Aktivitäten und Anliegen der Gesellschaft zugeordnet werden. Offizielle Dokumente, Informationen über öffentliche Veranstaltungen und Zeitungsartikel in gedruckter Form oder im Internet sind Beispiele für solche Texte.
- *Bildungsbezogene Situationen* sind für den Zweck des Lehrens und Lernens ausgelegt. Dazu zählen beispielsweise Lehrbuchtexte und Lernsoftware.
- *Berufsbezogene Situationen* umfassen meist Leseaufgaben, die das unmittelbare Ausführen einer Aufgabe im beruflichen Kontext mit sich bringen. Beispiele für berufsbezogene Situationen sind das Lesen von Stellenausschreibungen in der Zeitung oder im Internet und das Befolgen schriftlicher Anweisungen am Arbeitsplatz.

© Waxmann Verlag GmbH

Die vier Situationen können sich überschneiden. So kann beispielsweise ein Text einer berufsbezogenen Situation auch allgemeine Informationen beinhalten (berufsbezogen und öffentlich) und ein Text kann sowohl private Interessen betreffen als auch lehrreich sein (privat und bildungsbezogen).

Da es im alltäglichen Gebrauch viele verschiedene Arten von Texten gibt, wird im Rahmen der PISA-Studie eine möglichst große Bandbreite von *Textarten* einbezogen. Die *Textarten* werden bezüglich *Textformaten* und *Texttypen* spezifiziert.

- Bezüglich der *Textformate* werden *kontinuierliche* und *nichtkontinuierliche Texte* differenziert. Während *kontinuierliche Texte* aus Sätzen bestehen, die in Absätze gegliedert sind (z. B. Zeitungsartikel, Aufsätze, Romane und Briefe), sind *nichtkontinuierliche Texte* in Form von Listen, Tabellenformaten oder anderen visuell-strukturierenden Formen gegliedert (Tabellen, Grafiken, Diagramme, Stundenpläne, Fahrpläne). Bei vielen Texten handelt es sich um *gemischte Texte*, die Elemente des *kontinuierlichen* und des *nichtkontinuierlichen* Formats verbinden, wie zum Beispiel längere kontinuierliche Textabschnitte kombiniert mit Tabellen und Grafiken (Berichte, Zeitschriftenartikel, Webseiten, Online-Foren). Außerdem können mehrere unabhängige Texte zum gleichen Inhalt zu *multiplen Texten* zusammengesetzt werden. Ein Beispiel hierfür ist die Kombination mehrerer Webseiten von verschiedenen Reiseveranstaltern zu einem Reiseziel.
- Es werden sechs verschiedene *Texttypen* unterschieden: *Beschreibungen* (Schilderung eines bestimmten Orts in einem Tagebuch, Katalog, Landkarte), *Erzählungen* (Roman, Kurzgeschichte, Biografie), *Darlegungen* (wissenschaftlicher Aufsatz, Informationsgrafik, Lexikoneintrag), *Argumentationen* (Leserbrief, Beitrag in einem Online-Forum, Film- oder Buchkritik), *Anleitungen* (Rezept, Richtlinien zu einer Software) und *Transaktionen* (Austausch von Informationen, beispielsweise der Austausch von E-Mails zwischen Kollegen oder Freunden, um Vereinbarungen zu treffen).

Mit den *kognitiven Aspekten des Lesens* werden die kognitiven Prozesse beschrieben, welche die Leserinnen und Leser bei der Auseinandersetzung mit dem Text anwenden müssen, um die Fragestellung zu bearbeiten und zur Aufgabenlösung zu gelangen. Dabei werden die drei Aspekte (1) *Informationen suchen und extrahieren*, (2) *Kombinieren und Interpretieren* und (3) *Reflektieren und Bewerten* unterschieden.

- Beim *Informationen suchen und extrahieren* werden bestimmte Informationen aus dem Text lokalisiert und abgerufen. Leserinnen und Leser suchen dabei die benötigten Informationen heraus, die explizit im Text genannt werden (z. B. eine Liste von Begriffen oder der Name einer Person). Aufgaben, bei denen Informationen gesucht und extrahiert werden müssen, verlangen beispielsweise von den Schülerinnen und Schülern, die Anforderungen des Arbeitgebers in einer Stellenausschreibung zu finden oder eine bestimmte Aussage in einem Text zu identifizieren. Neben vergleichsweise einfachen Suchaufgaben, bei denen die Informationen deutlich sichtbar darge-

© Waxmann Verlag GmbH

stellt sind, können diese Aufgaben auch aufgrund verstreuter Informationen im Text oder durch bestimmte Navigationsanforderungen von höherem Schwierigkeitsgrad sein. Zur Beantwortung dieser Aufgaben müssen jedoch keine über den Text hinausgehenden Schlussfolgerungen gezogen werden.

- Durch *Kombinieren und Interpretieren* wird der Sinn des gelesenen Textes erschlossen. Informationen aus dem Text werden miteinander verknüpft (*Kombinieren*) und das Verständnis des Gelesenen wird durch das Aufstellen von nicht explizit im Text genannten Schlussfolgerungen gezeigt (*Interpretieren*). Sowohl *Kombinieren* als auch *Interpretieren* sind für das tiefere Verständnis eines Textes notwendig. Beim *Kombinieren* werden zuerst Zusammenhänge zwischen verschiedenen Informationen innerhalb des Textes hergestellt, dann werden diese Teilinformationen zusammengeführt, wodurch eine Interpretation des Textes als Ganzes möglich wird. Beim *Interpretieren* ziehen Leserinnen und Leser Schlussfolgerungen aus dem Text, indem sie beispielsweise die implizierte Bedeutung eines Wortes oder eines Satzes interpretieren, die dem Text eine bestimmte Wertung gibt. Ein anderes Beispiel für *Interpretieren* ist das Ziehen von Schlussfolgerungen über die Absichten der Autorin oder des Autors.
- *Reflektieren und Bewerten* beinhaltet das Einbeziehen eigenen Wissens sowie eigener Ideen und Einstellungen, die über den Text hinausgehen. Beim *Reflektieren* werden eigene Erfahrungen und eigenes Wissen hinzugezogen, um zu vergleichen oder Hypothesen gegenüberzustellen. *Bewerten* bedeutet, dass ein Urteil abgegeben wird, welches auf Kriterien beruht, die über den Text hinausgehen. Sowohl beim *Reflektieren* als auch beim *Bewerten* müssen Informationen aus dem Text mit Wissen aus anderen Quellen verbunden werden. Die Leserin beziehungsweise der Leser prüft dabei, ob die Aussagen des Textes mit dem eigenen Wissen über die Welt übereinstimmen und wie diese Anknüpfungspunkte zur Beantwortung der Frage beitragen können.

Der Lesekompetenztest in PISA 2015

Der Lesekompetenztest in der PISA-Studie 2015 setzt alle in der PISA-Rahmenkonzeption spezifizierten Anforderungen um. Durch den Einbezug von Aufgaben, welche die verschiedenen *Situationen*, *Textarten* und *kognitiven Aspekte des Lesens* abdecken, wird gewährleistet, dass die Lesekompetenz als Konstrukt mit verschiedenen Facetten möglichst breit erfasst wird (OECD, 2016a). Zur Erfassung der Lesekompetenz wurden in der PISA-Studie 2015 insgesamt 29 Aufgaben eingesetzt, die sich aus 103 Teilaufgaben (Items) zusammensetzen.[2] Diese Aufgaben kamen bereits in früheren Erhebungsrunden zum Einsatz, was ermöglicht, die Leistungen der Schülerinnen und Schüler im Jahr 2015 mit denen der Schülerinnen und Schüler aus früheren Erhebungen von PISA zu vergleichen. Da die Lesekompetenz im Rahmen der PISA-Studie 2015 als Nebendomäne und nicht als Hauptdomäne erfasst wurde, ist es aufgrund der zu geringen Menge von

2 Eine Übersicht zu den eingesetzten Aufgaben und Teilaufgaben findet sich im Anhang B.

© Waxmann Verlag GmbH

Aufgaben nicht zulässig, Subskalen der Lesekompetenz zu bilden. Es wird daher ausschließlich der Gesamtmittelwert der Lesekompetenz betrachtet. Dieser wurde so gebildet, dass er mit den Gesamtmittelwerten der Lesekompetenz aus den vorherigen PISA-Erhebungen (2000 bis 2012) verglichen werden kann.

Übergang vom papierbasierten zum computerbasierten Lesen

Eine Neuerung der PISA-Studie 2015 im Vergleich zu früheren Erhebungen besteht darin, dass die Lesekompetenz (wie auch die naturwissenschaftliche und die mathematische Kompetenz) in Deutschland ausschließlich computerbasiert erfasst wurde (vgl. Kapitel 12). Um die Vergleichbarkeit der computerbasiert erfassten Lesekompetenz 2015 mit den vorherigen PISA-Erhebungen sicherzustellen, wurde darauf geachtet, dass die Aufgaben- und Antwortformate der Computerversion der Papierversion aus früheren PISA-Erhebungen möglichst gut entsprechen. Daher wurden die Antwortformate unverändert beibehalten. Somit wurden, wie bei den bisherigen PISA-Erhebungen, sowohl geschlossene (z. B. Multiple-Choice) als auch offene Antwortformate eingesetzt. Während bei Multiple- bzw. Single-Choice-Antworten eine oder mehrere Antworten aus vorgegebenen Antwortmöglichkeiten ausgewählt werden, verfassen die Schülerinnen und Schüler bei offenen Antwortformaten selbst kurze Antworten. Da der zu lesende Text in einigen Fällen in Papierform auf mehreren Seiten dargestellt worden war, wurden lange Texte in der computerbasierten Form ebenfalls auf mehreren Seiten dargestellt und nicht auf einer Seite mit Scroll-Funktion. Trotzdem kann nicht ausgeschlossen werden, dass die Umstellung von papier- auf computerbasiertes Testen die Testergebnisse beeinflusst. In mehreren Studien wurde bereits untersucht, ob Schülerinnen und Schüler bei computerbasierten Testungen besser oder schlechter lesen als bei papierbasierten Testungen. Auf Ebene empirischer Primärstudien ist dabei eine divergierende Befundlage zu konstatieren (OECD, 2016a; Wang, Jiao, Young, Brooks & Olson, 2008). Die Ergebnisse einer Metaanalyse zeigten jedoch, dass über 36 Studien hinweg die Art der Testung (ob computer- oder papierbasiert) keinen statistisch signifikanten Effekt auf die Lesekompetenz der Schülerinnen und Schüler hatte (Wang et al., 2008). Demnach hing das Abschneiden in den Kompetenztests nicht generell damit zusammen, ob die Schülerinnen und Schüler mit Aufgaben auf Papier oder am Computer getestet wurden. Allerdings könnte der Effekt der Umstellung von papierbasierten auf computerbasierte Lesekompetenztests in jeder Studie je nach Art der Testaufgaben und abhängig von der Umsetzung der Aufgaben am Computer unterschiedlich sein, möglicherweise auch nur für einzelne Gruppen von Schülerinnen und Schülern. So gibt es Hinweise darauf, dass mit computerbasierten Lesekompetenztests kleinere Geschlechterunterschiede in der Lesekompetenz gefunden werden als mit papierbasierten Tests (Horne, 2007; OECD, 2011).

© Waxmann Verlag GmbH

Die Kompetenzstufen

Die Antworten der Jugendlichen zu den Aufgaben des Lesekompetenztests wurden in der PISA-Studie 2015 – im Prinzip[3] wie in den vorherigen PISA-Erhebungen auch – auf der Grundlage der *Item-Response-Theorie* skaliert (vgl. Kapitel 12). Dies hat zum einen den Vorteil, dass verschiedene Schülerinnen und Schüler unterschiedliche Aufgaben bearbeiten können und somit eine viel größere Anzahl an Aufgaben eingesetzt werden kann, als eine einzelne Person bearbeiten könnte. Nur so ist es möglich, die breit angelegte Definition der Lesekompetenz überhaupt durch die Aufgaben abzudecken. Zum anderen bietet die Skalierung den Vorteil, dass die Kompetenzen der Jugendlichen und die Schwierigkeit der Items (Teilaufgaben) auf einer gemeinsamen Skala abgebildet werden können. Dies bedeutet auch, dass die Kompetenzen der Schülerinnen und Schüler direkt durch die Inhalte und Anforderungen der diesen Kompetenzen entsprechenden (d.h. gleiche Skalenwerte aufweisenden) Testitems interpretiert werden. Im Rahmen der ersten PISA-Erhebungsrunde im Jahr 2000 wurden die Lesekompetenzwerte so normiert, dass eine Skala mit einem OECD-Mittelwert von 500 und einer Standardabweichung (als Maß der Streuung) von 100 entsteht. Durch die leicht unterschiedliche Zusammensetzung der OECD-Staaten bei jeder PISA-Erhebung und weitere Einflussfaktoren liegt der Mittelwert in späteren Erhebungen nicht mehr genau bei 500, sondern kann etwas verschoben sein.

Zur Veranschaulichung, welche Teilaufgaben Schülerinnen und Schüler mit unterschiedlich stark ausgeprägter Lesekompetenz typischerweise lösen können, wird die Lesekompetenzskala in mehrere Kompetenzstufen unterteilt. Nachdem im ersten PISA-Zyklus (PISA 2000, 2003 und 2006) zunächst fünf Stufen der Lesekompetenz unterschieden worden waren, ist die Skala im aktuellen zweiten Zyklus beginnend mit der PISA-Studie 2009 differenzierter, nämlich in sieben Kompetenzstufen unterteilt. Jede Kompetenzstufe umfasst in etwa 73 Punkte auf der Lesekompetenzskala. Mit steigender Kompetenzstufe der Jugendlichen steigt die Schwierigkeit der Items. Je schwieriger ein Item ist, desto geringer ist die Wahrscheinlichkeit, dass Jugendliche mit einer geringen Lesekompetenz das Item lösen (vgl. Kapitel 12). Die Lesekompetenz der Jugendlichen sowie die Itemschwierigkeiten können durch die ihnen zugrunde liegende gemeinsame Skala den sieben Kompetenzstufen zugeordnet werden. Tabelle 7.1 gibt einen Überblick über die typischen Anforderungen der sieben Kompetenzstufen der Lesekompetenz.

Um die Kompetenzstufen und Aufgaben zu veranschaulichen, sind jeweils zwei Items aus den Aufgaben „Geizhals“ und „Heißluftballon“, welche in der PISA-Studie 2009 eingesetzt wurden, abgebildet. Die Texte dieser beiden Aufgaben sind in den Abbildungen 7.1 und 7.2 dargestellt. Abbildung 7.3 stellt die entsprechende Zuordnung der vier Items zu den Kompetenzstufen dar. Der Text „Geizhals“ ist, gemäß der Rahmenkonzeption des Lesekompetenztests, ein *kontinuierlicher Text*, der einer *privaten Situation* zuzuordnen ist. Diese Zuordnung schließt jedoch nicht aus, dass der Text auch in *bildungsbezogenen*

3 Mit dem sogenannten *2-Parameter-Logistischen-Modell* (2 PL-Modell) wurde eine leicht geänderte Variante verwendet, die auch die Trennschärfe der Items berücksichtigt.

© Waxmann Verlag GmbH

Tabelle 7.1: Überblick über die typischen Anforderungen der sieben Kompetenzstufen (Ib – VI) der Lesekompetenz (vgl. Naumann, Artelt, Schneider & Stanat, 2010, S. 27–28)

Kompetenzstufe	Wozu die Schülerinnen und Schüler auf der jeweiligen Kompetenzstufe im Allgemeinen in der Lage sind
VI **≥ 698 Punkte**	Jugendliche auf dieser Stufe können Schlussfolgerungen, Vergleiche und Gegenüberstellungen detailgenau und präzise anstellen. Dabei entwickeln sie ein volles und detailliertes Verständnis eines oder mehrerer Texte und verbinden dabei unter Umständen gedanklich Informationen aus mehreren Texten miteinander. Hierbei kann auch die Auseinandersetzung mit ungewohnten Ideen gefordert sein, genauso wie der kompetente Umgang mit konkurrierenden Informationen und abstrakten Interpretationskategorien sowie hohe Präzision im Umgang mit zum Teil unauffälligen Textdetails.
V **626–697 Punkte**	Jugendliche auf dieser Stufe können sowohl mehrere tief eingebettete Informationen finden, ordnen und herausfinden, welche davon jeweils relevant sind, als auch ausgehend von Fachwissen eine kritische Beurteilung oder Hypothese anstellen. Die Aufgaben dieser Stufe setzen in der Regel ein volles und detailliertes Verständnis von Texten voraus, deren Inhalt oder Form ungewohnt ist. Zudem muss mit Konzepten umgegangen werden können, die im Gegensatz zum Erwarteten stehen.
IV **553–625 Punkte**	Aufgaben dieser Kompetenzstufe erfordern vom Leser/von der Leserin, linguistischen oder thematischen Verknüpfungen in einem Text über mehrere Abschnitte zu folgen, oftmals ohne Verfügbarkeit eindeutiger Kennzeichen im Text, um eingebettete Informationen zu finden, zu interpretieren und zu bewerten oder um psychologische oder philosophische Bedeutungen zu erschließen. Insgesamt muss ein genaues Verständnis langer oder komplexer Texte, deren Inhalt oder Form ungewohnt sein kann, unter Beweis gestellt werden.
III **480–552 Punkte**	Aufgaben dieser Kompetenzstufe erfordern vom Leser/von der Leserin, vorhandenes Wissen über die Organisation und den Aufbau von Texten zu nutzen, implizite oder explizite logische Relationen (z. B. Ursache-Wirkungs-Beziehungen) über mehrere Sätze oder Textabschnitte zu erkennen, mit dem Ziel, Informationen im Text zu lokalisieren, zu interpretieren und zu bewerten. Einige Aufgaben verlangen vom Leser/von der Leserin, einen Zusammenhang zu begreifen oder die Bedeutung eines Wortes oder Satzes zu analysieren. Häufig sind die benötigten Informationen dabei nicht leicht sichtbar oder Passagen des Textes laufen eigenen Erwartungen zuwider.
II **408–479 Punkte**	Jugendliche auf dieser Stufe können innerhalb eines Textabschnitts logischen und linguistischen Verknüpfungen folgen, mit dem Ziel, Informationen im Text zu lokalisieren oder zu interpretieren; im Text oder über Textabschnitte verteilte Informationen aufeinander beziehen, um die Absicht des Autors zu erschließen. 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.
Ia **335–407 Punkte**	Aufgaben dieser Kompetenzstufe erfordern vom Leser/von der Leserin, in einem Text zu einem vertrauten Thema eine oder mehrere unabhängige, explizit ausgedrückte Informationen zu lokalisieren, das Hauptthema oder die Absicht des Autors zu erkennen oder einen einfachen Zusammenhang zwischen den im Text enthaltenen Informationen und allgemeinem Alltagswissen herzustellen. Die erforderlichen Informationen sind in der Regel leicht sichtbar, und es sind nur wenige beziehungsweise keine konkurrierenden Informationen vorhanden. Der Leser wird explizit auf die entscheidenden Elemente in der Aufgabe und im Text hingewiesen.
Ib **262 –334 Punkte**	Jugendliche auf dieser Stufe können in einem kurzen, syntaktisch einfachen Text aus einem gewohnten Kontext, dessen Form vertraut ist (z. B. in einer einfachen Liste oder Erzählung), eine einzige, explizit ausgedrückte Information lokalisieren, die leicht sichtbar ist. Der Text enthält in der Regel Hilfestellungen für den Leser, wie Wiederholungen, Bilder oder bekannte Symbole. Es gibt kaum konkurrierende Informationen. Bei anderen Aufgaben müssen einfache Zusammenhänge zwischen benachbarten Informationsteilen hergestellt werden.
unter Ib **≤ 261 Punkte**	

© Waxmann Verlag GmbH

Situationen vorkommen könnte. Die Beispielitems 5 und 7 zu dem Text „Geizhals" erfordern von den Fünfzehnjährigen frei formulierte Kurztextantworten. Offene Antworten wie diese werden von intensiv geschulten und qualifizierten Kodiererinnen und Kodierern anhand standardisierter Kodiervorschriften ausgewertet (vgl. Kapitel 1). Um die richtigen Lösungen der Beispielitems zu dem Text „Geizhals" nachvollziehbar zu machen, sind in Tabelle 7.2 die Kodieranweisungen zu diesen Items dargestellt.

Der Text „Heißluftballon" ist ein *nichtkontinuierlicher Text*, der einer *bildungsbezogenen Situation* zugeordnet wurde. Die Beispielitems zum Text „Heißluftballon" weisen ein Single-Choice-Antwortformat auf, bei welchem die Jugendlichen die richtige aus vier Antwortmöglichkeiten auswählen sollen. Bei Beispielitem 6 müssen die Schülerinnen und Schüler über den Inhalt des Textes reflektieren und diesen bewerten. Die richtige Lösung ist Antwortoption B. Beispielitem 8 erfordert, Inhalte des Textes zu kombinieren und zu interpretieren. Dabei ist die richtige Lösung ebenfalls Antwortoption B. Die Auswertung dieser Antworten konnte aufgrund der computerbasierten Datenerfassung automatisiert und digital erfolgen (vgl. Kapitel 1). Die ausgewählten Beispielitems veranschaulichen die unteren und mittleren Kompetenzstufen (Ib, Ia, II und III). Es handelt sich also um Items mit niedrigem bis mittlerem Schwierigkeitsgrad. Beispielitems, die höheren Kompetenzstufen zugeordnet sind, wurden in den PISA-Berichtsbänden 2009 (Naumann et al., 2010) und 2012 (Hohn et al., 2013) dargestellt.

Der Geizhals und sein Gold

Eine Fabel von Aesop

Ein Geizhals verkaufte alles, was er hatte, und kaufte einen Klumpen Gold, den er in einem Loch in der Erde neben einer alten Mauer vergrub. Jeden Tag ging er, um danach zu sehen. Einer seiner Arbeiter bemerkte die regelmäßigen Besuche des Geizhalses an dem Ort und beschloss, dessen Kommen und Gehen zu beobachten. Der Arbeiter entdeckte bald das Geheimnis des versteckten Schatzes, grub im Boden, fand den Klumpen Gold und stahl ihn. Der Geizhals fand das Loch bei seinem nächsten Besuch leer vor und begann, sich die Haare zu raufen und laut zu klagen. Ein Nachbar, der ihn in seiner Verzweiflung sah und den Grund dafür erfuhr, sagte zu ihm: „Bitte grämt Euch nicht so; nehmt Euch einen Stein, legt ihn ins Loch und stellt Euch vor, das Gold läge noch dort. Er wird Euch den gleichen Dienst erweisen, denn auch, als das Gold noch da war, besaßt Ihr es nicht, da Ihr nicht den geringsten Gebrauch davon gemacht habt."

Abbildung 7.1: Der Text „Geizhals" zu den in Abbildung 7.3 dargestellten Items 5 und 7

© Waxmann Verlag GmbH

Heißluftballon

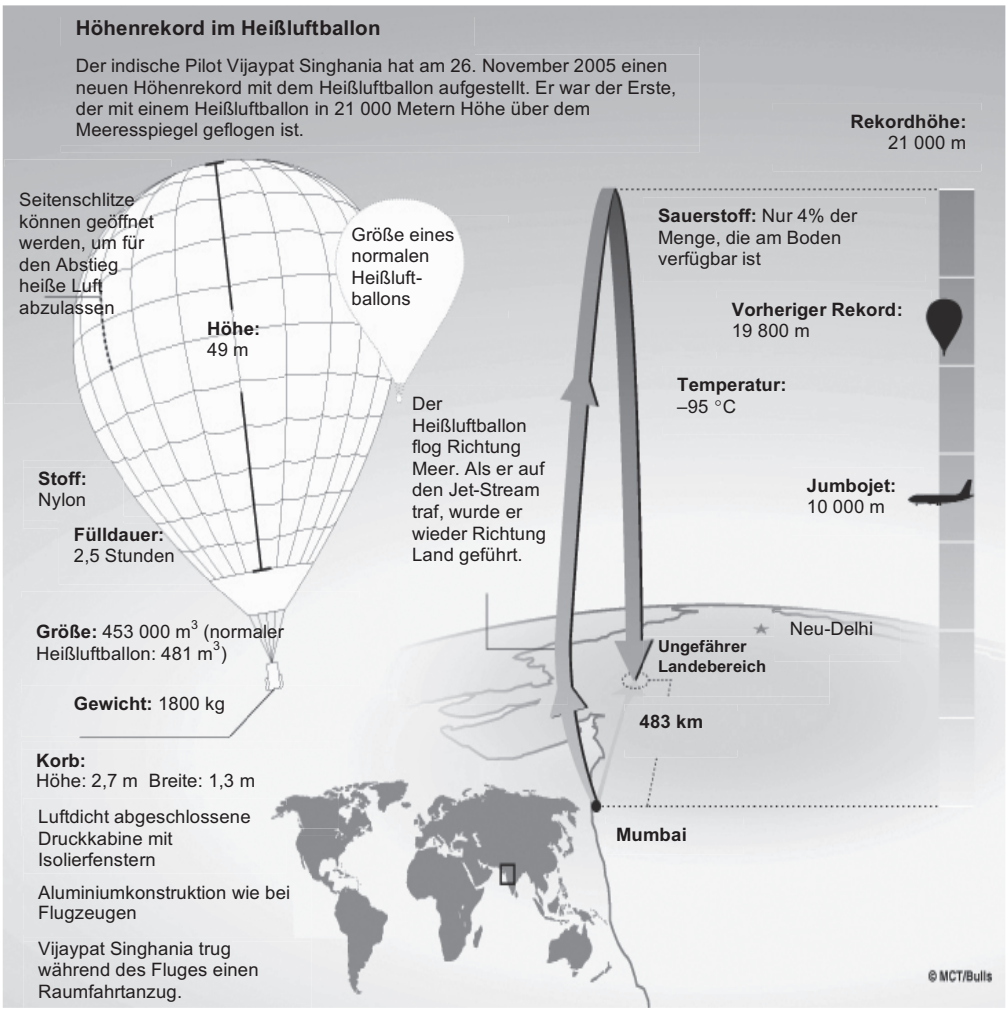
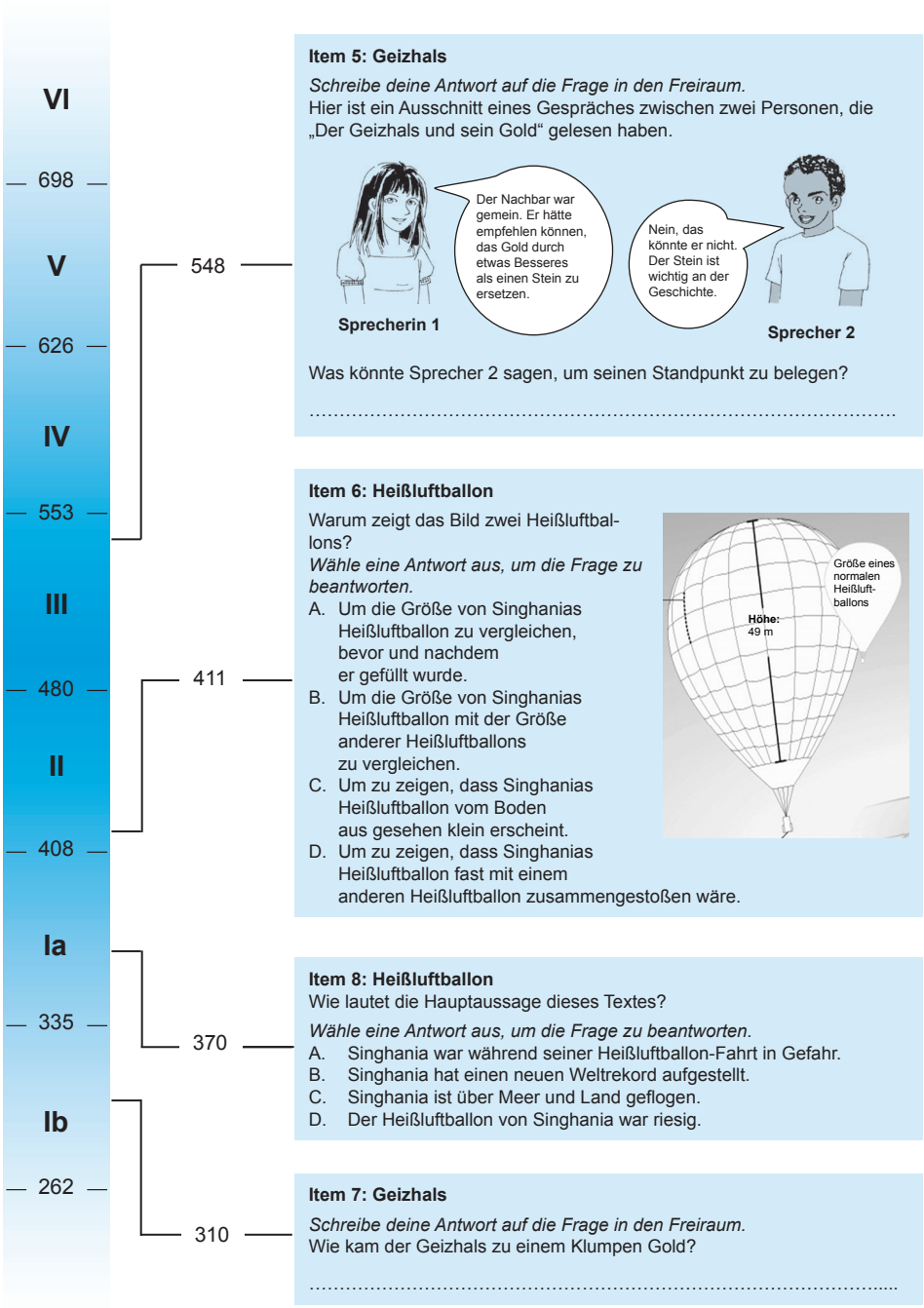


Abbildung 7.2: Der Text „Heißluftballon“ zu den in Abbildung 7.3 dargestellten Items 6 und 8

© Waxmann Verlag GmbH



Anmerkung: Die Werte an den Verbindungslinien zwischen den Beispielitems und der Kompetenzsäule geben das Kompetenzniveau an, mit dem eine Schülerin oder ein Schüler das Item mit einer Wahrscheinlichkeit von 62% lösen würde.

Abbildung 7.3: Itembeispiele zur Veranschaulichung der Zuordnung von Items zu den Kompetenzstufen der Lesekompetenz

© Waxmann Verlag GmbH

Tabelle 7.2: Kodieranweisungen zu den in Abbildung 7.3 dargestellten Items 5 und 7 zum Text „Geizhals“

	Kodieranweisung zu Item 5: Geizhals	**Kodieranweisung zu Item 7: Geizhals**
Absicht der Frage	Kombinieren und Interpretieren: Eine Interpretation entwickeln Ein Detail in Beziehung zur Kernaussage einer Fabel setzen.	Informationen suchen und extrahieren: Informationen ermitteln Eine am Anfang eines kurzen Textes explizit angegebene Information heraussuchen.
Vollständig gelöst: Code 1	Erkennt, dass die Kernaussage der Geschichte davon abhängt, dass das Gold durch etwas Unnützes oder Wertloses ersetzt wird. • Es musste durch etwas Wertloses ersetzt werden, damit es Sinn macht. • Der Stein ist wichtig in der Geschichte, denn der entscheidende Punkt ist, dass er genauso gut einen Stein hätte vergraben können für all das, was das Gold ihm gebracht hat. • Wenn man es mit etwas Besserem als einem Stein ersetzt, geht der Kern der Geschichte verloren, weil die vergrabene Sache etwas wirklich Nutzloses sein muss. • Ein Stein ist nutzlos, aber für den Geizhals war es das Gold auch! • Etwas Besseres wäre etwas, das er gebrauchen könnte – er hat vom Gold keinen Gebrauch gemacht, und das wollte der Mann zeigen. • Weil Steine überall gefunden werden können. Das Gold und der Stein sind das Gleiche für den Geizhals. *[„können überall gefunden werden“ impliziert, dass der Stein keinen besonderen Wert hat]*	Sagt aus, dass er alles verkaufte, was er besaß. Kann den Text paraphrasieren oder zitieren. • Er verkaufte alles, was er besaß. • Er verkaufte sein ganzes Zeug. • Er kaufte ihn. *[impliziert, dass er alles verkaufte, was er besaß]*
Nicht gelöst: Code 0	Antwort ungenügend oder vage • Der Stein war wichtig in der Geschichte. *[wiederholt die Aussage des Sprechers]* • Es musste ein Stein sein. *[Erklärung fehlt.]* • Es wäre nicht das Gleiche gewesen. *[vage]* • Der Stein hat einen symbolischen Wert in der Geschichte *[Erklärt nicht, was der symbolische Wert ist.]* • Zeigt ungenaues Verständnis des Materials oder gibt eine unplausible oder irrelevante Antwort. • Es musste ein Stein sein, weil ein Stein schwer ist.	Antwort ungenügend oder vage • Es war seines. • Er hat es verdient. • Zeigt ungenaues Verständnis des Materials oder gibt eine unplausible oder irrelevante Antwort. • Er hat es gestohlen.
Code 9:	Keine Antwort	Keine Antwort

© Waxmann Verlag GmbH

7.2 Lesekompetenz im internationalen Vergleich

Im Folgenden wird berichtet, wie die Lesekompetenz der Fünfzehnjährigen in Deutschland im Vergleich zur Lesekompetenz der Fünfzehnjährigen in anderen Staaten und insbesondere den OECD-Staaten im Jahr 2015 ausgeprägt ist. Hierzu wird zunächst der Mittelwert der Lesekompetenz der Fünfzehnjährigen in Deutschland mit den Mittelwerten aus anderen Staaten, die an der PISA-Studie 2015 teilnahmen, verglichen. Außerdem werden die Streuungen der Lesekompetenz innerhalb der verschiedenen Staaten betrachtet und miteinander verglichen. Im Anschluss wird die Verteilung der Jugendlichen auf die Kompetenzstufen der Lesekompetenz in den einzelnen Staaten beschrieben. Zuletzt werden Unterschiede zwischen Mädchen und Jungen im internationalen Vergleich präsentiert.

Lesekompetenz: Mittelwerte im internationalen Vergleich

Die Mittelwerte, Standardabweichungen und Perzentilbänder der Lesekompetenz aller OECD-Staaten sind in Abbildung 7.4 dargestellt. Dabei sind die Staaten nach ihrem Mittelwert der Lesekompetenz absteigend angeordnet und werden in drei Gruppen in Abhängigkeit ihrer Relation zum OECD-Mittelwert unterteilt: Statistisch signifikant über oder unter dem OECD-Durchschnitt oder statistisch nicht von diesem zu unterscheiden. Die durchschnittliche Lesekompetenz über alle OECD-Staaten liegt in der PISA-Studie 2015 bei 493 Punkten. In Deutschland zeigen die Fünfzehnjährigen mit einem Mittelwert von 509 Punkten ein signifikant besseres Ergebnis: Der Mittelwert liegt deutlich über dem Mittelwert der OECD-Staaten. In der PISA-Studie 2012 übertraf die Lesekompetenz der Fünfzehnjährigen in Deutschland mit einem Mittelwert von 508 Punkten den OECD-Durchschnitt zum ersten Mal signifikant, dieses Ergebnis konnte somit – fast punktgleich – repliziert werden. In den PISA-Erhebungen in den Jahren 2003, 2006 und 2009 hatte die durchschnittliche Lesekompetenz der Jugendlichen in Deutschland im mittleren Leistungsbereich und bei der ersten Erhebung im Jahr 2000 sogar signifikant unter dem OECD-Durchschnitt gelegen. In der PISA-Studie 2015 stabilisiert sich somit die positive Entwicklung der Lesekompetenz in Deutschland im internationalen Vergleich. Deutschland liegt zudem bei Betrachtung der Lesekompetenz aller OECD-Staaten an neunter Stelle. Bei der Erhebung im Jahr 2012 lag Deutschland dagegen noch an dreizehnter Stelle.

Es sind insgesamt 19 OECD-Staaten, bei denen die durchschnittliche Lesekompetenz der Jugendlichen statistisch signifikant über dem Durchschnitt aller OECD-Staaten liegt. An der Spitze dieser Gruppe stehen Kanada (527 Punkte), Finnland (526 Punkte) und Irland (521 Punkte). Zwischen den Staaten dieser Gruppe gibt es jedoch große Unterschiede in der Lesekompetenz. So beträgt der Abstand zwischen den Mittelwerten des Vereinigten Königreichs und Kanada 29 Punkte.

Nicht statistisch signifikant vom OECD-Mittelwert verschieden ist die durchschnittliche Lesekompetenz von Fünfzehnjährigen in den Vereinigten Staaten (497 Punkte), in

© Waxmann Verlag GmbH

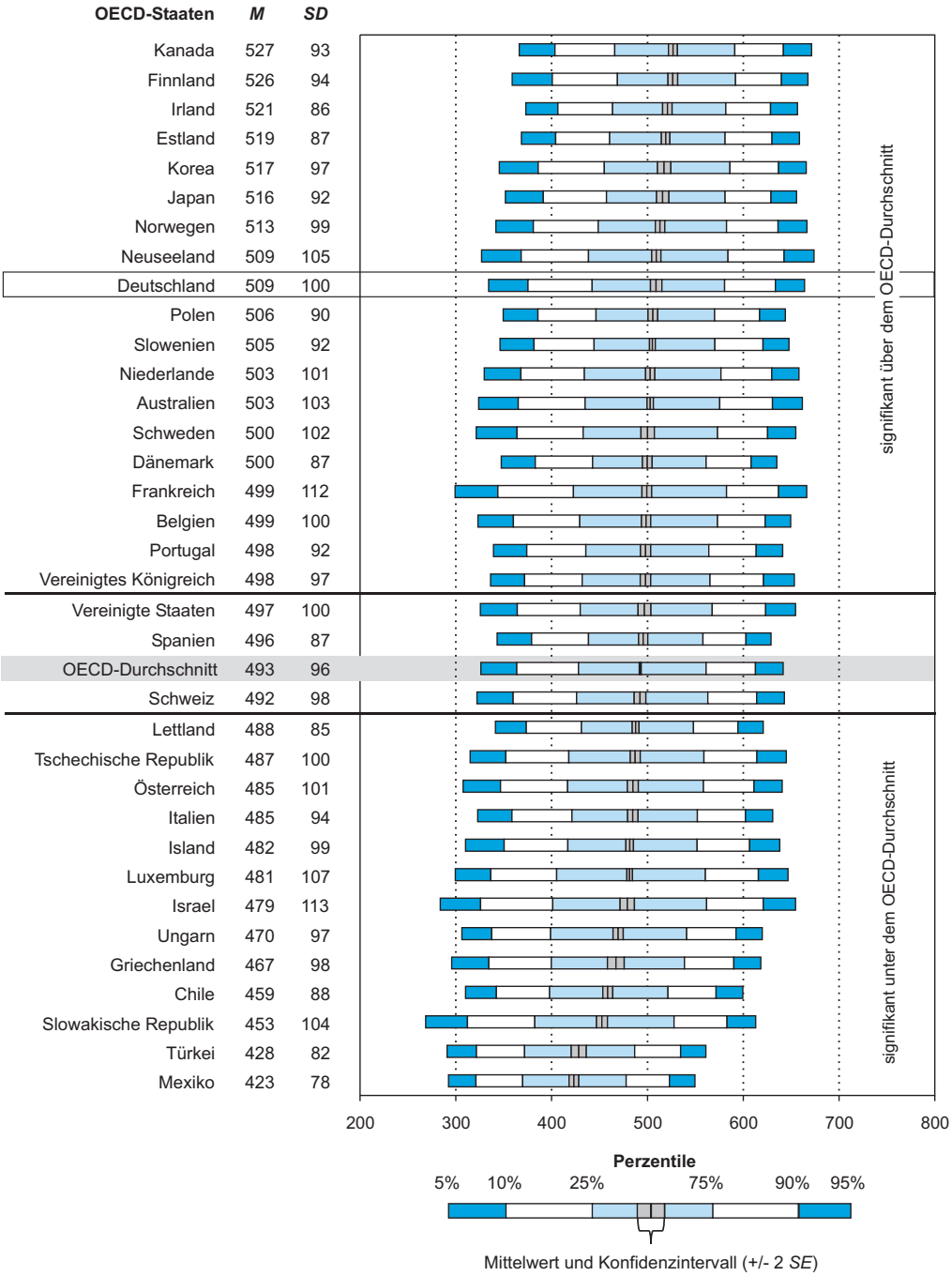


Abbildung 7.4: Perzentilbänder der Lesekompetenz in den OECD-Staaten

© Waxmann Verlag GmbH

Spanien (496 Punkte) und in der Schweiz (492 Punkte). Die Gruppe von Staaten, deren durchschnittliche Lesekompetenz der Schülerinnen und Schüler statistisch signifikant unter dem OECD-Durchschnitt liegt, umfasst 14 Staaten. Dabei reichen die Mittelwerte von 488 Punkten in Lettland bis zu 423 Punkten in Mexiko. In Mexiko liegt somit der Mittelwert mehr als eine Kompetenzstufe unter dem OECD-Mittelwert.

Betrachtet man die Lesekompetenz in den OECD-Partnerstaaten (vgl. Tabelle A6 im Anhang), so ist festzustellen, dass Jugendliche in Singapur mit 535 Punkten eine besonders hohe durchschnittliche Lesekompetenz erreichen, welche noch höher ist als die Lesekompetenz der Jugendlichen in Kanada. Zu den leistungsschwächsten OECD-Partnerstaaten gehören Algerien (350 Punkte), Kosovo (347 Punkte) und Libanon (347 Punkte), bei welchen die durchschnittliche Lesekompetenz der Jugendlichen fast zwei Kompetenzstufen unter dem OECD-Mittelwert liegt.

Die Streuung der Lesekompetenz

Um die Lesekompetenz der Jugendlichen in einzelnen Staaten zu beschreiben, ist zusätzlich zu den erreichten Mittelwerten von Interesse, wie stark die Unterschiede in der Lesekompetenz innerhalb einzelner Staaten ausgeprägt sind. Diese Streuung der Lesekompetenz wird in Form der Standardabweichung (*SD*) in Abbildung 7.4 angegeben. Je höher die Streuung (Standardabweichung) ist, desto größer sind im Mittel die Unterschiede zwischen den leistungsschwächsten und leistungsstärksten Leserinnen und Lesern innerhalb eines Staates. Zusätzlich veranschaulichen die in Abbildung 7.4 dargestellten Perzentilbänder die Bandbreite der Lesekompetenz in den einzelnen Staaten. Je breiter das Perzentilband ist, desto höher ist der Kompetenzunterschied zwischen den 5 Prozent der schwächsten und den 5 Prozent der stärksten Schülerinnen und Schülern eines Staates. In der PISA-Studie 2015 beträgt die Streuung der Lesekompetenz um den Mittelwert über alle OECD-Staaten 96 Punkte. In Deutschland beträgt die Standardabweichung der Lesekompetenz 100 Punkte und ist damit signifikant größer als die durchschnittliche Standardabweichung der OECD-Staaten.

Besonders hohe Streuungen in der Lesekompetenz finden sich in Israel (SD = 113), Frankreich (SD = 112) und Luxemburg (SD = 107). Bereits in den Erhebungsrunden 2009 und 2012 zeigten sich in diesen drei Staaten hohe Streuungen in der Lesekompetenz. Die geringsten Streuungen in der Lesekompetenz haben Mexiko (SD = 78), die Türkei (SD = 82), Lettland (SD = 85) und Irland (SD = 86). Das Beispiel Irland belegt, dass eine hohe durchschnittliche Lesekompetenz bei gleichzeitig vergleichsweise niedrigeren mittleren Unterschieden zwischen den Leistungsschwächsten und Leistungsstärksten innerhalb eines Staates möglich ist.

© Waxmann Verlag GmbH

Verteilung auf die Stufen der Lesekompetenz

Anhand der Kompetenzstufen kann die im PISA-Test erzielte Lesekompetenz der Schülerinnen und Schüler direkt durch die für die jeweilige Stufe typischen Anforderungen der Aufgaben interpretiert werden. Dabei ist es von besonderem Interesse, wie groß die Anteile besonders leseschwacher (Kompetenzstufe Ia oder darunter) und besonders lesestarker (Kompetenzstufen V und VI) Schülerinnen und Schüler in verschiedenen Staaten sind, da diese Anteile Aufschluss darüber geben, wie gut es den Bildungssystemen einzelner Staaten gelingt, sowohl den Anteil besonders leseschwacher Jugendlicher gering zu halten als auch eine Spitzengruppe besonders lesestarker Jugendlicher aufzubauen. Abbildung 7.5 zeigt für alle OECD-Staaten die Anteile von Schülerinnen und Schülern auf den untersten Kompetenzstufen (Ia oder darunter) beziehungsweise auf den obersten Kompetenzstufen (V oder VI).

Im OECD-Durchschnitt befinden sich 20 Prozent der Fünfzehnjährigen auf den untersten Kompetenzstufen (Ia oder darunter). Diese Schülerinnen und Schüler verfügen nur über eingeschränkte Lesefähigkeiten und sind deshalb in der Regel nicht ausreichend auf ihre weitere Schul- und Berufslaufbahn sowie für eine aktive Teilhabe an der Gesellschaft vorbereitet. In Deutschland beträgt der Anteil der Jugendlichen, die sich auf den untersten Kompetenzstufen befinden, 16 Prozent und ist damit signifikant geringer als im OECD-Durchschnitt. Noch geringer ist der Anteil leseschwacher Fünfzehnjähriger an der Spitze der OECD-Staaten. In Kanada und Finnland befinden sich nur 11 Prozent und in Irland 10 Prozent der Fünfzehnjährigen auf den untersten Stufen der Lesekompetenz. Kanada, Finnland und Irland erreichen die höchsten Mittelwerte der Lesekompetenz aller OECD-Staaten und weisen gleichzeitig einen vergleichsweise geringen Anteil leseschwacher Schülerinnen und Schüler auf. Auch in Singapur, dem OECD-Partnerstaat mit der höchsten durchschnittlichen Lesekompetenz, sind es nur 11 Prozent der Jugendlichen, die sich auf den untersten Stufen der Lesekompetenz befinden.

Auf den obersten Kompetenzstufen (V und VI) befinden sich im OECD-Durchschnitt 8 Prozent der Fünfzehnjährigen. Jugendliche, deren Kompetenz auf den obersten Stufen liegt und die sich somit in der Spitzengruppe befinden, verfügen über sehr gute Voraussetzungen für eigenständiges, flexibles Weiterlernen und die Teilhabe an modernen Gesellschaften, in denen Wissen zu einem großen Teil in Form von Texten weitergegeben wird. In Deutschland gehören 12 Prozent der Schülerinnen und Schüler zu den besonders lesestarken Jugendlichen. Damit ist der Anteil der Spitzengruppe im Lesen in Deutschland signifikant größer als im OECD-Durchschnitt. Ebenfalls große Anteile von Schülerinnen und Schülern auf den obersten Kompetenzstufen finden sich in den OECD-Staaten Kanada und Finnland mit einem Anteil von 14 Prozent und in Irland mit einem Anteil von 11 Prozent. In dem OECD-Partnerstaat Singapur befinden sich sogar 18 Prozent der Fünfzehnjährigen auf den obersten Stufen der Lesekompetenz.

© Waxmann Verlag GmbH

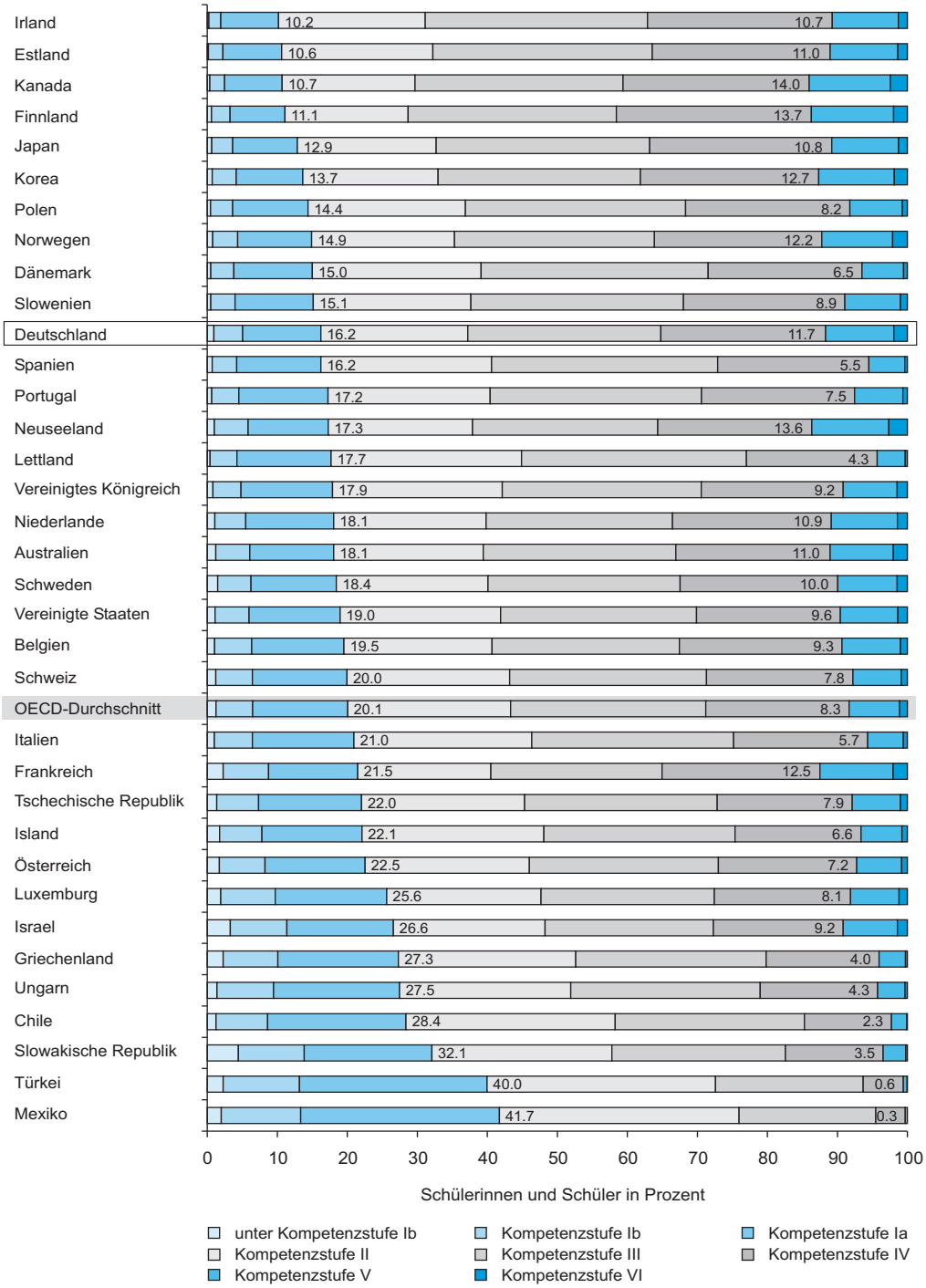


Abbildung 7.5: Prozentuale Anteile von Schülerinnen und Schülern auf Kompetenzstufe Ia, Kompetenzstufe Ib oder darunter sowie auf Kompetenzstufe V und Kompetenzstufe VI

© Waxmann Verlag GmbH

Unterschiede in der Lesekompetenz zwischen Mädchen und Jungen

Wie in den früheren PISA-Erhebungen verfügen Mädchen auch in der PISA-Studie 2015 über eine signifikant höhere Lesekompetenz als Jungen (vgl. Abbildung 7.6). Im OECD-Durchschnitt sind die Mädchen um 27 Punkte besser als die Jungen, was etwas mehr als einem Drittel einer Kompetenzstufe entspricht. Auch in Deutschland verfügen Mädchen mit einem Mittelwert von 520 Punkten über eine deutlich höhere Lesekompetenz als Jungen, die im Durchschnitt 499 Punkte erreichen. Auch wenn dieser Leistungsunterschied von 21 Punkten in Deutschland beträchtlich ist, hat er sich im Vergleich zu dem in der PISA-Studie 2012 signifikant verringert. Im Jahr 2012 betrug die Differenz zwischen Mädchen und Jungen in Deutschland noch 44 Punkte. Nicht nur in Deutschland, sondern auch in vielen anderen Teilnehmerstaaten hat sich der Geschlechterunterschied in der Lesekompetenz in der PISA-Erhebung 2015 im Vergleich zu 2012 erheblich reduziert. Dabei gehört Deutschland neben Israel, Portugal und Italien zu den OECD-Staaten, bei denen die Geschlechterdifferenz seit 2012 besonders stark zurückgegangen ist (um 21 bis 23 Punkte). Finnland und Slowenien sind die beiden OECD-Staaten, bei denen sich in der Erhebung 2015 – wie zuvor schon in der PISA-Erhebung 2012 – die größten Geschlechterunterschiede in der Lesekompetenz zeigen. Auch in diesen Staaten hat der Geschlechterunterschied jedoch von 2012 zu 2015 abgenommen. Während die Geschlechterdifferenz in der PISA-Studie 2015 in Finnland 47 Punkte und in Slowenien 43 Punkte beträgt, belief sich diese 2012 in Finnland noch auf 62 Punkte und in Slowenien auf 56 Punkte. Im OECD-Durchschnitt hat sich die Differenz zwischen Mädchen und Jungen in der PISA-Studie 2015 (28 Punkte Differenz) im Vergleich zu der PISA-Studie 2012 (37 Punkte Differenz) um 9 Punkte verringert. Allerdings hat sich der Unterschied zwischen Mädchen und Jungen nicht konsistent über alle Teilnehmerstaaten hinweg verringert. Insgesamt ist eher unklar, welche Ursachen für die markante Abnahme der Geschlechterunterschiede in einigen Staaten verantwortlich sind und wie die dafür verantwortlichen Faktoren gegebenenfalls zusammenspielten.

Ein in diesem Zusammenhang wichtiger Aspekt ist die Umstellung auf computerbasiertes Testen in der PISA-Studie 2015. Frühere Studien geben Hinweise darauf, dass Jungen bei computerbasierten Lesetests eine bessere Performanz zeigen als bei papierbasierten Lesetests (Horne, 2007; OECD, 2011). Daher könnte möglicherweise in manchen Staaten die Umstellung auf computerbasiertes Testen zu besseren Ergebnisse der Jungen im Lesekompetenztest in der PISA-Studie 2015 geführt haben.

© Waxmann Verlag GmbH

OECD-Staaten	Mädchen M	Mädchen (SE)	Jungen M	Jungen (SE)	Differenz Jungen – Mädchen
Finnland	551	(2.8)	504	(3.0)	
Slowenien	528	(2.1)	484	(2.3)	
Lettland	509	(2.4)	467	(2.3)	
Island	502	(2.6)	460	(2.8)	
Korea	539	(4.0)	498	(4.8)	
Norwegen	533	(2.9)	494	(3.1)	
Schweden	520	(3.5)	481	(4.1)	
Griechenland	486	(4.2)	449	(5.1)	
Slowakische Republik	471	(3.5)	435	(3.3)	
Neuseeland	526	(3.0)	493	(3.3)	
Australien	519	(2.3)	487	(2.3)	
Polen	521	(2.8)	491	(2.9)	
Frankreich	514	(3.3)	485	(3.3)	
Estland	533	(2.3)	505	(2.9)	
Türkei	442	(4.8)	414	(4.5)	
OECD-Durchschnitt	506	(0.5)	479	(0.6)	
Kanada	540	(2.5)	514	(2.6)	
Tschechische Republik	501	(2.9)	475	(3.6)	
Schweiz	505	(3.4)	480	(3.4)	
Ungarn	482	(3.1)	457	(3.7)	
Niederlande	515	(2.9)	491	(3.0)	
Israel	490	(4.6)	467	(5.4)	
Dänemark	511	(3.4)	489	(2.8)	
Vereinigtes Königreich	509	(3.5)	487	(2.9)	
Luxemburg	492	(2.2)	471	(1.9)	
Deutschland	520	(3.1)	499	(3.7)	
Spanien	506	(2.8)	485	(3.0)	
Österreich	495	(3.7)	475	(4.3)	
Vereinigte Staaten	507	(3.9)	487	(3.7)	
Portugal	507	(2.8)	490	(3.1)	
Italien	493	(3.6)	477	(3.5)	
Belgien	507	(2.9)	491	(3.1)	
Mexiko	431	(2.9)	416	(2.9)	
Japan	523	(3.3)	509	(4.2)	
Irland	527	(2.7)	515	(3.2)	
Chile	465	(2.9)	453	(3.4)	

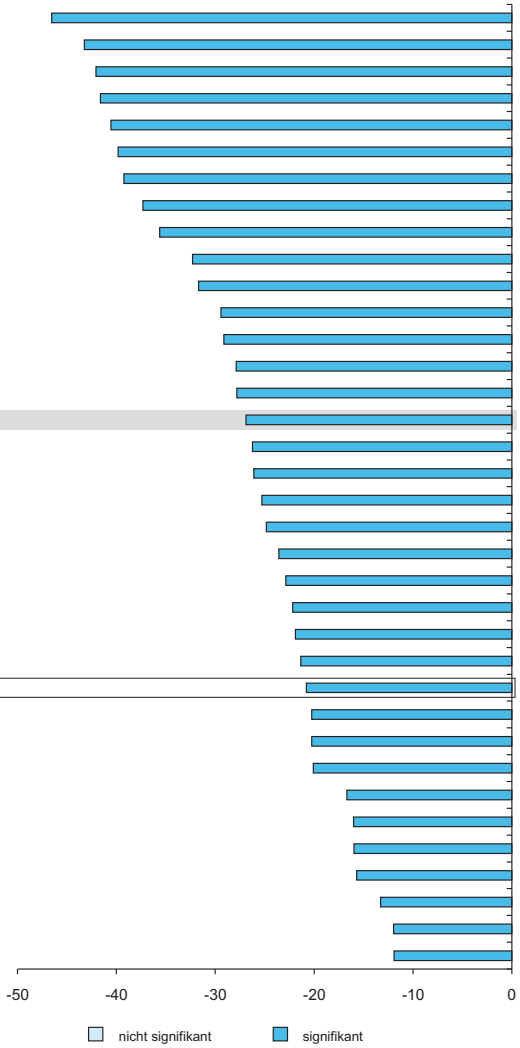


Abbildung 7.6: Mittelwerte der Lesekompetenz von Mädchen und Jungen in den OECD-Staaten

© Waxmann Verlag GmbH

7.3 Vertiefende Analysen zur Lesekompetenz in Deutschland

Neben dem Vergleich der Lesekompetenz fünfzehnjähriger Schülerinnen und Schüler mit anderen an PISA teilnehmenden Staaten ist aus deutscher Perspektive ein differenzierter Blick auf spezifische nationale Befunde von großem Interesse. In den folgenden Abschnitten werden daher die Ergebnisse vertiefender Analysen zur Lesekompetenz innerhalb Deutschlands aufgezeigt. Konkret werden Unterschiede in der Lesekompetenz zwischen Schularten dargestellt sowie Unterschiede zwischen Mädchen und Jungen in Deutschland differenzierter betrachtet. Im Anschluss wird präsentiert, wie sich die Lesekompetenz der Schülerinnen und Schüler seit der PISA-Studie 2000 in Deutschland verändert hat. Die Ergebnisse der Lesekompetenz von PISA 2015 werden vor allem mit denen von PISA 2009 verglichen. Im Jahr 2009 war Lesen Hauptdomäne, sodass eine differenziertere Testung als 2012 und 2015 stattgefunden hat, die eine differenziertere Bewertung der Ergebnisse erlaubt.

7.3.1 Unterschiede zwischen Schularten

Eine Besonderheit im deutschen Schulsystem ist die nach der Grundschule erfolgende Aufgliederung in verschiedene weiterführende Schularten. Zum Zeitpunkt der Stichprobenziehung und Datenerhebung der PISA-Studie 2015 bestanden in den Ländern der Bundesrepublik Deutschland unterschiedliche Schularten für die Sekundarstufe. Dabei ist das Gymnasium die einzige Schulart, die nach wie vor in allen 16 Ländern der Bundesrepublik Deutschland existiert (vgl. Kapitel 1). Ein Vergleich der anderen Schularten über die Länder hinweg scheint den teilweise stark unterschiedlichen Rahmenbedingungen in den 16 Ländern nicht mehr gerecht zu werden. Daher wird in diesem Berichtsband bei der Betrachtung von Unterschieden zwischen Schularten in Deutschland lediglich zwischen Gymnasien und nicht gymnasialen Schularten (Hauptschule, Schule mit mehreren Bildungsgängen, Integrierte Gesamtschule und Realschule) unterschieden. Die Leistungen von Schülerinnen und Schülern an Sonder- und Förderschulen sowie an beruflichen Schulen gehen in den Durchschnitt der Gesamtstichprobe der Fünfzehnjährigen in Deutschland ein, werden jedoch nicht in die Ergebnisse der nicht gymnasialen Schularten einbezogen (vgl. Kapitel 1).

Lesekompetenz: Mittelwerte und Streuungen im Vergleich

Die Ergebnisse der PISA-Studie 2015 zeigen, dass sich die durchschnittliche Lesekompetenz der Jugendlichen zwischen nicht gymnasialen Schularten (Hauptschule, Schule mit mehreren Bildungsgängen, Integrierte Gesamtschule und Realschule) und dem Gymnasium beträchtlich unterscheidet. In Tabelle 7.3 sind die Mittelwerte und Streuungen (Standardabweichungen) der Lesekompetenz für die nicht gymnasialen Schularten, das Gymnasium sowie für die Gesamtstichprobe dargestellt. Die Gesamtstichprobe umfasst

© Waxmann Verlag GmbH

alle in Deutschland getesteten Schülerinnen und Schüler (n = 6 504).[4] Es zeigt sich, dass Schülerinnen und Schüler an Gymnasien durchschnittlich eine höhere Lesekompetenz erzielen. Mit 583 Punkten liegt der Mittelwert der Gymnasiastinnen und Gymnasiasten um 74 Punkte und somit in etwa eine Kompetenzstufe über dem Mittelwert der deutschen Gesamtstichprobe (509 Punkte). Die durchschnittliche Lesekompetenz von Fünfzehnjährigen an nicht gymnasialen Schularten befindet sich mit einem Mittelwert von 478 Punkten fast eine halbe Kompetenzstufe unter dem Mittelwert der deutschen Gesamtstichprobe und mehr als eine Kompetenzstufe unter dem Mittelwert der Gymnasiastinnen und Gymnasiasten.

Tabelle 7.3: Mittelwerte und Standardabweichungen der Lesekompetenz in Deutschland für nicht gymnasiale Schularten, Gymnasium und Gesamtstichprobe

Schulart	*n*	*M*	*(SE)*	*SD*	*(SE)*
Nicht gymnasiale Schularten	3 944	478	(3.4)	88	(1.8)
Gymnasium	2 266	583	(3.0)	73	(1.8)
Gesamtstichprobe	6 504	509	(3.0)	100	(1.6)

Die Lesekompetenz der deutschen Jugendlichen unterscheidet sich aber nicht nur zwischen nicht gymnasialen Schularten und dem Gymnasium, sondern auch innerhalb der Schularten variiert die Lesekompetenz und zeigt jeweils eine große Bandbreite. Diese Streuungen werden in Form der Standardabweichungen (*SD*) in Tabelle 7.3 angegeben. Innerhalb der nicht gymnasialen Schularten (*SD* = 88) findet sich eine größere Streuung der Lesekompetenz als innerhalb des Gymnasiums. Dies ist nicht verwunderlich, da bei den nicht gymnasialen Schularten mehrere Bildungsgänge zusammengefasst werden und somit zu erwarten ist, dass sich die Jugendlichen in ihrer Lesekompetenz stärker unterscheiden als jene, welche die gleiche Schulart besuchen. Allerdings findet sich auch innerhalb des Gymnasiums (*SD* = 73) eine beträchtliche Streuung, die große Unterschiede in der Lesekompetenz widerspiegelt. So verteilen sich beispielsweise die mittleren 50 Prozent der Gymnasiastinnen und Gymnasiasten auf die Kompetenzstufen III bis V.

Darüber hinaus überschneiden sich die Verteilungen der Lesekompetenz der Schülerinnen und Schüler an nicht gymnasialen Schularten und dem Gymnasium. Diese Überschneidungen sind in Abbildung 7.7 erkennbar. Die Abbildung veranschaulicht die Mittelwerte und Streuungen der Lesekompetenz für die nicht gymnasialen Schularten, das Gymnasium und die deutsche Gesamtstichprobe anhand von Perzentilbändern. Man erkennt hier, dass die besten 10 Prozent der Schülerinnen und Schüler an nicht gymnasialen Schularten über eine höhere Lesekompetenz verfügen als mehr als 50 Prozent der Gymnasiastinnen und Gymnasiasten. Außerdem ist aus der Abbildung ersichtlich,

4 Zur Gesamtstichprobe gehören auch die Jugendlichen aus beruflichen Schulen sowie aus Sonder- und Förderschulen.

© Waxmann Verlag GmbH

dass ein Großteil der Schülerinnen und Schüler an nicht gymnasialen Schularten auf den Kompetenzstufen II und III zu finden ist.

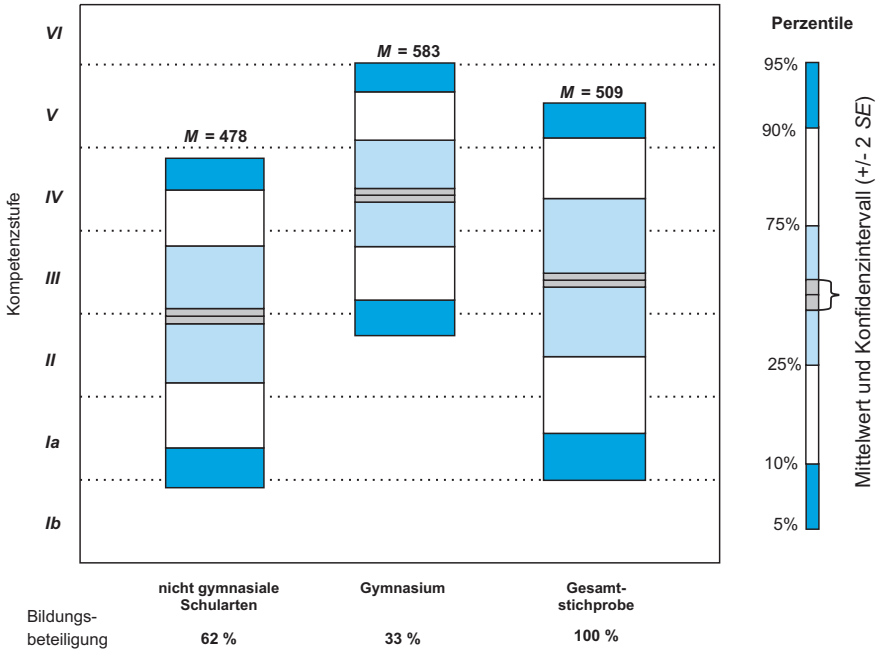


Abbildung 7.7: Perzentilbänder für die Lesekompetenz in Deutschland für nicht gymnasiale Schularten, Gymnasium und Gesamtstichprobe

Verteilung auf die Stufen der Lesekompetenz

Um die Unterschiede in der Lesekompetenz zwischen nicht gymnasialen Schularten und dem Gymnasium genauer aufzuschlüsseln, werden in Abbildung 7.8 die prozentualen Anteile der Schülerinnen und Schüler auf den Kompetenzstufen dargestellt. Aus dieser Abbildung kann somit auch abgelesen werden, wie groß die Anteile besonders leseschwacher (Kompetenzstufe Ia oder darunter) und besonders lesestarker (Kompetenzstufen V und VI) Schülerinnen und Schüler an nicht gymnasialen Schularten und dem Gymnasium sind. Während sich lediglich 1 Prozent der Gymnasiastinnen und Gymnasiasten auf den Kompetenzstufen Ia oder darunter befindet, sind es 21 Prozent der Schülerinnen und Schüler an nicht gymnasialen Schularten. Das heißt, dass der Anteil der leseschwachen Schülerinnen und Schüler an nicht gymnasialen Schularten noch immer beträchtlich ist. Insgesamt hat sich der Anteil der Fünfzehnjährigen auf den untersten Kompetenzstufen seit 2009 nicht signifikant verändert. Schülerinnen und Schülern, die sich auf der Kompetenzstufe Ia oder darunter befinden, fehlen die Kompetenzen, um sich mit ungewohnten Texten auseinanderzusetzen und Schlussfolgerungen über mehrere Textteile zu ziehen. Sie haben mit hoher Wahrscheinlichkeit Schwierigkeiten, Informationen aus einem Text zu interpretieren und Details eines Textes in Beziehung zur

© Waxmann Verlag GmbH

Kernaussage zu setzen (vgl. z. B. Beispielitem 5 in Abbildung 7.3). Außerdem sind leseschwache Jugendliche kaum in der Lage, Vergleiche anzustellen oder Zusammenhänge zwischen dem Text und den nicht im Text enthaltenen Informationen zu erkennen (vgl. z. B. Beispielitem 6 in Abbildung 7.3). Gemäß der in der theoretischen Rahmenkonzeption definierten Lesekompetenz als Teil einer Grundbildung muss davon ausgegangen werden, dass diese Jugendlichen schlechte Voraussetzungen im Hinblick auf ihre Schul- und Berufslaufbahn mitbringen. Ausgehend von den äußerst eingeschränkten Lesefähigkeiten dieser Gruppe ist zu befürchten, dass sie kaum in der Lage sind, sich über das Lesen eigenständig Informationen zu erschließen und alltägliche Routinen der Informationsbeschaffung und Kommunikation zu erwerben, die für einen gelingenden Einstieg in ein selbstbestimmtes Erwachsenenleben Voraussetzung sind.

Eine Verringerung des Anteils dieser besonders leseschwachen Jugendlichen in allen Schularten ist demnach noch immer eines der wichtigsten Ziele der Schulbildung und außerschulischen Förderung in Deutschland. Wie im Einleitungsteil dieses Kapitels erläutert, wurden nach der PISA-Studie 2000 sieben Handlungsfelder von der KMK (KMK, 2002) entwickelt, in welchen die Leseförderung deutlich betont wird und woraus zahlreiche Maßnahmen und Projekte zur Förderung der Lesekompetenz hervorgingen. Diese sollten konsequent weiter vorangetrieben werden. Konkret scheint eine durchgängige Förderung der Lesekompetenz vom vorschulischen Bereich bis zum Ende der Schulpflicht sinnvoll zu sein (Artelt et al., 2007; Beck et al., 2015). Insbesondere sind weitere Anstrengungen angeraten, um Risikogruppen, wie beispielsweise Kinder und Jugendliche mit Zuwanderungshintergrund (vgl. Abbildung A3 im Anhang), gezielt in ihrer Sprach- und Lesekompetenz zu fördern (Gold & Dubowy, 2013; Kreyenfeld & Krapf, 2013; Schröder, Keller, Dintsioudi & List, 2013).

Der Anteil besonders lesestarker Schülerinnen und Schüler ist am Gymnasium deutlich größer als an nicht gymnasialen Schularten. An Gymnasien sind die Leistungen von 28 Prozent der Jugendlichen auf den Kompetenzstufen V und VI anzusiedeln. Damit ist diese Spitzengruppe um 8 Prozent größer als noch bei der PISA-Studie 2009. An nicht gymnasialen Schularten befinden sich 4 Prozent der Schülerinnen und Schüler auf diesen Kompetenzstufen, 2009 waren es nur 2 Prozent. Insgesamt ist die Gruppe der Schülerinnen und Schüler auf den obersten Kompetenzstufen seit 2009 statistisch bedeutsam gewachsen. Die Jugendlichen, die sich in der Spitzengruppe, also auf den Kompetenzstufen V und VI befinden, sind in der Lage, ungewohnte Texte detailliert zu verstehen und können mit Konzepten umgehen, die im Gegensatz zum Erwarteten stehen. Diese Gruppe hat mit fünfzehn Jahren ein Niveau erreicht, das beste Prognosen für eigenständiges Lernen über das Lesen und die künftige Teilhabe an gesellschaftlichen Prozessen zulässt und die Jugendlichen mit einer hohen Flexibilität ausstattet. Lesestarke Jugendliche erkennen, was die Intention eines Textes ist und erfassen darin enthaltene Informationen problemlos. Sie erkennen verborgene Bedeutungen „zwischen den Zeilen“ und sind in der Lage, komplexe Schlussfolgerungen zu ziehen, indem sie Gelesenes an ihr eigenes Weltwissen anknüpfen. In Gesellschaften, die den Großteil ihres Wissens in

© Waxmann Verlag GmbH

Form von Texten sammeln und weitergeben, haben diese Schülerinnen und Schüler sehr gute Voraussetzungen für ein erfolgreiches, selbstbestimmtes Leben.

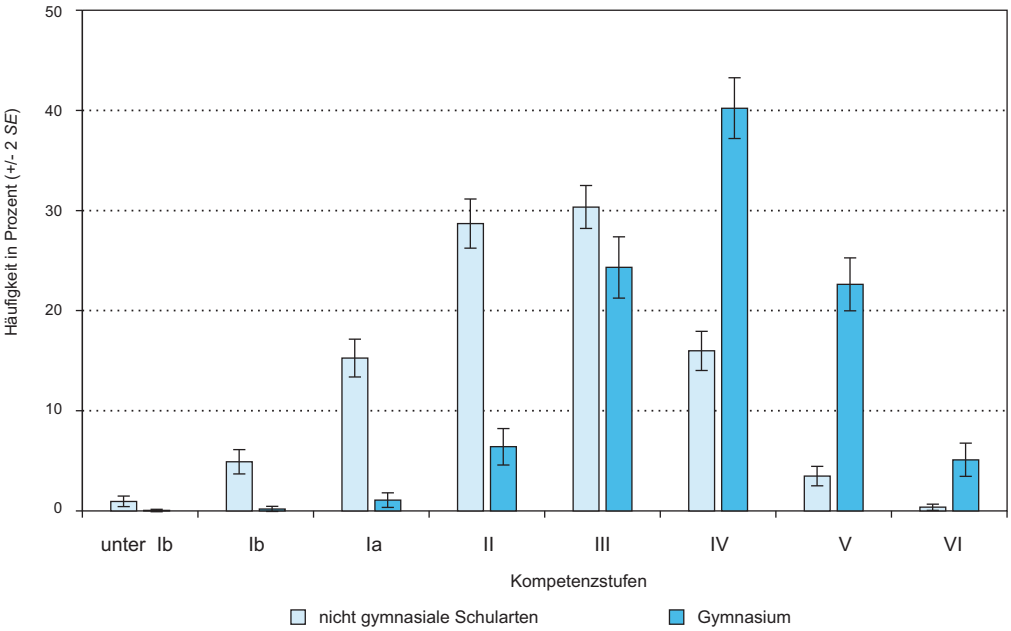


Abbildung 7.8: Prozentuale Anteile der Schülerinnen und Schüler in Deutschland auf den Stufen der Lesekompetenz an nicht gymnasialen Schularten und am Gymnasium

7.3.2 Unterschiede zwischen Mädchen und Jungen in der Lesekompetenz

Wie bereits im Absatz zu Unterschieden zwischen Mädchen und Jungen im internationalen Vergleich berichtet, erreichen Mädchen in Deutschland eine deutlich höhere Lesekompetenz (520 Punkte) als Jungen (499 Punkte). Um diesen Geschlechterunterschied im Lesen für Deutschland differenzierter zu betrachten, werden in Abbildung 7.9 die prozentualen Anteile der Mädchen und Jungen auf den Kompetenzstufen dargestellt. Sie zeigt, dass der Anteil leseschwacher Jungen höher ist als der Anteil leseschwacher Mädchen. Während sich insgesamt 19 Prozent der Jungen auf Kompetenzstufe Ia oder darunter befinden, sind es nur 14 Prozent der Mädchen. Umgekehrt ist ein höherer Anteil der Mädchen auf den beiden obersten Kompetenzstufen (V und VI) vertreten und somit besonders lesestark. Von den Mädchen befinden sich 14 Prozent und von den Jungen 10 Prozent auf den Kompetenzstufen V und VI.

© Waxmann Verlag GmbH

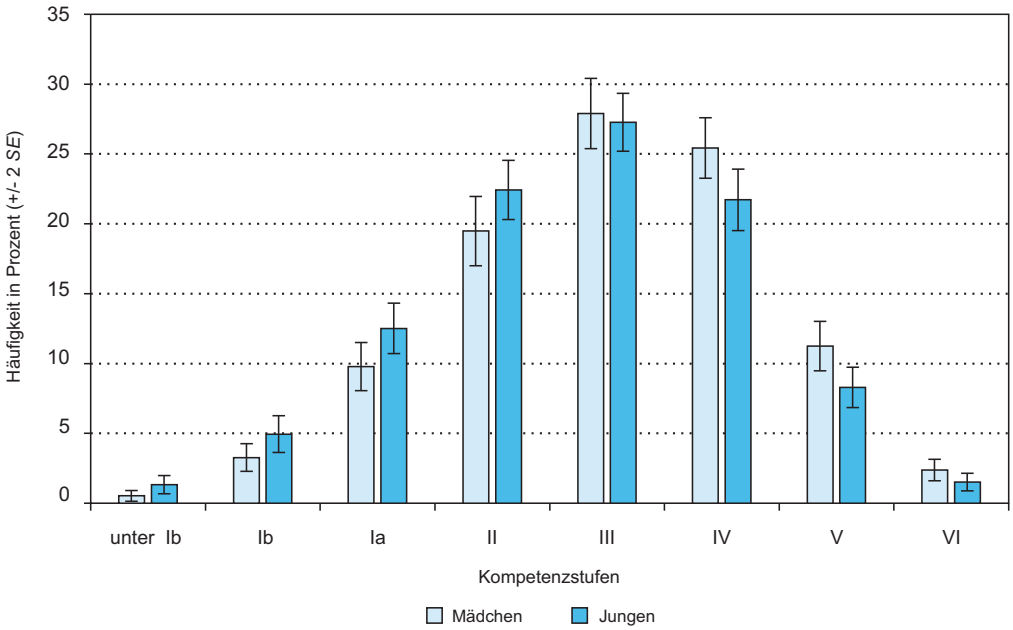


Abbildung 7.9: Prozentuale Anteile der Mädchen und Jungen in Deutschland auf den Stufen der Lesekompetenz

Im Vergleich zu der PISA-Studie 2009 sind 2015 signifikant weniger Jungen auf den untersten Kompetenzstufen (Ia oder darunter) vertreten. Abbildung 7.10 veranschaulicht, dass der Anteil der Jungen auf den untersten beiden Kompetenzstufen seit 2009 (von 24 auf 19 Prozent) gesunken ist. Der Anteil der Mädchen auf den untersten beiden Kompetenzstufen hat sich seit 2009 (von 13 auf 14 Prozent) nicht signifikant verändert. Somit zeichnet sich die Tendenz ab, dass sich der Unterschied zwischen Jungen und Mädchen auf den untersten Kompetenzstufen verringert. Gleichzeitig ist der Anteil der Jungen auf den obersten Stufen (V und VI) der Lesekompetenz seit 2009 (von 4 auf 10 Prozent) signifikant gestiegen. Damit hat sich der Anteil der lesestarken Jungen (Kompetenzstufen V und VI) seit 2009 mehr als verdoppelt. Dass der Wechsel von der Papier-und-Bleistift- auf die Computertestung den Jungen geholfen haben könnte, ist nicht auszuschließen. Der Anteil der Mädchen in der Spitzengruppe (Kompetenzstufen V und VI) hat sich im Vergleich zur PISA-Studie 2009 (von 11 auf 14 Prozent) nicht signifikant verändert.

© Waxmann Verlag GmbH

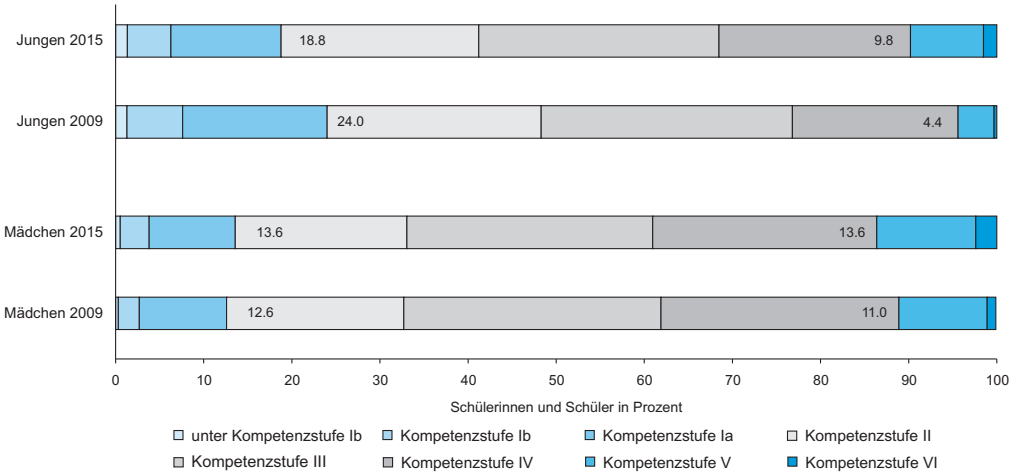


Abbildung 7.10: Prozentuale Anteile von Mädchen und Jungen in Deutschland bei PISA 2015 und 2009 auf Kompetenzstufe Ia, Ib und darunter sowie auf Kompetenzstufe V und VI der Gesamtskala Lesen

7.3.3 Veränderung der Lesekompetenz seit PISA 2000

Die PISA-Studie ermöglicht es festzustellen, wie sich die Lesekompetenz der Fünfzehnjährigen in Deutschland über die Zeit entwickelt. Diese Informationen sind relevant, um Hinweise darauf zu gewinnen, welche Weiterentwicklungen im Bildungssystem sinnvoll sein könnten. Wie im Abschnitt zu den Mittelwerten im internationalen Vergleich bereits berichtet wurde, verbesserte sich die durchschnittliche Lesekompetenz der Fünfzehnjährigen in Deutschland seit der PISA-Studie 2000 über die verschiedenen Erhebungsrunden hinweg. Dabei lag die Lesekompetenz 2009 (497 Punkte) erstmals signifikant über dem Mittelwert der PISA-Erhebung 2000 (484 Punkte). In der PISA-Erhebung 2012 stieg die Lesekompetenz der Schülerinnen und Schüler in Deutschland nochmals an (508 Punkte) und lag sogar signifikant über dem 2009 erzielten Durchschnittswert sowie erstmals über dem OECD-Durchschnitt. In der PISA-Studie 2015 zeigen die Schülerinnen und Schüler in Deutschland mit einem Mittelwert von 509 Punkten eine ähnlich hohe durchschnittliche Lesekompetenz wie in der Erhebungsrunde 2012. Die mittlere Lesekompetenz in Deutschland liegt auch 2015 über dem OECD-Durchschnitt. Zwischen den Durchschnittswerten der Lesekompetenz in den PISA-Erhebungen 2012 und 2015 bestehen keine signifikanten Unterschiede, sodass die Leistungen im Wesentlichen unverändert geblieben sind. Somit wird mit der PISA-Studie 2015 das Ergebnis von 2012 bestätigt, wodurch sich sowohl die positive Entwicklung der Lesekompetenz innerhalb Deutschlands als auch die Position des Mittelwerts der Lesekompetenz in Deutschland über dem OECD-Mittelwert konsolidieren. Dies ist keinesfalls selbstverständlich. Ein Blick in andere Staaten zeigt, dass zum Beispiel die mittlere Lesekompetenz der

© Waxmann Verlag GmbH

Fünfzehnjährigen in der Schweiz im Jahr 2012 mit 509 Punkten noch signifikant über dem OECD-Durchschnitt lag, 2015 aber auf 492 Punkte abfiel. Damit gehört die Schweiz nun zu der Gruppe von Staaten, in denen sich die Mittelwerte der Lesekompetenz nicht statistisch vom OECD-Durchschnitt unterscheiden. Ebenfalls signifikant gesunken ist die durchschnittliche Lesekompetenz der Jugendlichen 2015 im Vergleich zu 2012 in Japan (von 538 auf 516 Punkte), in Korea (von 536 auf 517 Punkte), in Ungarn (von 488 auf 470) und in der Türkei (von 475 auf 428). In Slowenien (von 481 auf 505 Punkte), Schweden (von 483 auf 500 Punkte) und in Chile (von 441 auf 459 Punkte) hingegen zeigen die Schülerinnen und Schüler 2015 eine deutlich höhere Lesekompetenz als noch 2012. Die OECD ordnet Deutschland der kleinen Gruppe von insgesamt fünf Staaten (Israel, Portugal, Deutschland, Hongkong (China), Japan) zu, in welchen sich die durchschnittliche Lesekompetenz der Fünfzehnjährigen über den gesamten Zeitraum von 2000 bis 2015 insgesamt verbessert hat (OECD, 2016b). In Abbildung 7.11 wird die Entwicklung der mittleren Lesekompetenz in Deutschland über die PISA-Erhebungsrunden von 2000 bis 2015 dargestellt. Die durchschnittliche Lesekompetenz an Gymnasien ist bei der PISA-Studie 2015 (583 Punkte) ähnlich hoch wie in den Erhebungsrunden 2012 (579 Punkte) und 2009 (575 Punkte).

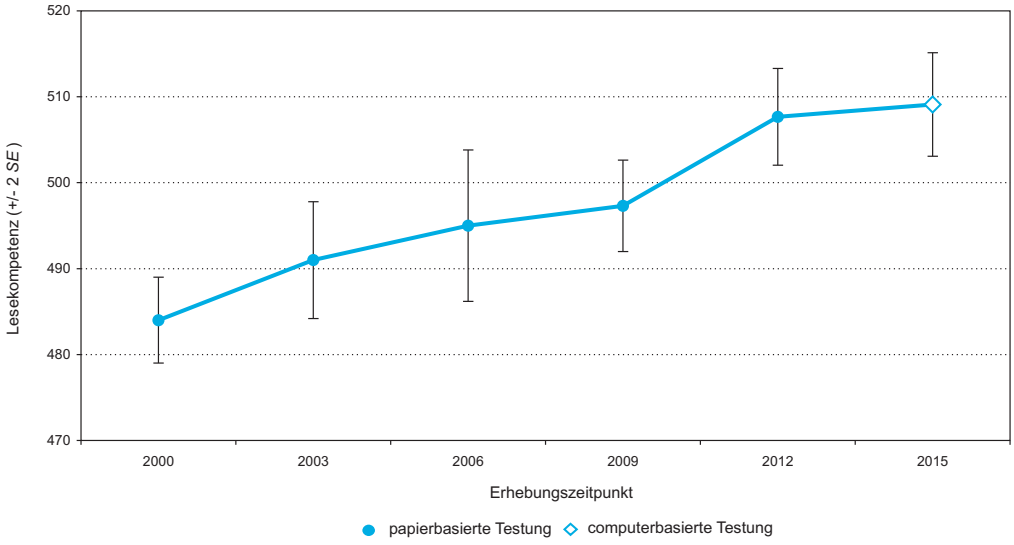


Abbildung 7.11: Veränderung der mittleren Lesekompetenz in Deutschland von 2000 bis 2015

Da sich seit der ersten PISA-Studie im Jahr 2000 und in allen weiteren bisherigen Erhebungen ausgeprägte Geschlechterunterschiede in der Lesekompetenz in Deutschland zeigten, ist es von besonderem Interesse, die Entwicklung der Lesekompetenz von Mädchen und Jungen getrennt zu betrachten. Wie aus Abbildung 7.12 ersichtlich wird, war in den Erhebungen 2000 bis 2012 ein Aufwärtstrend sowohl bei den Jungen als auch

© Waxmann Verlag GmbH

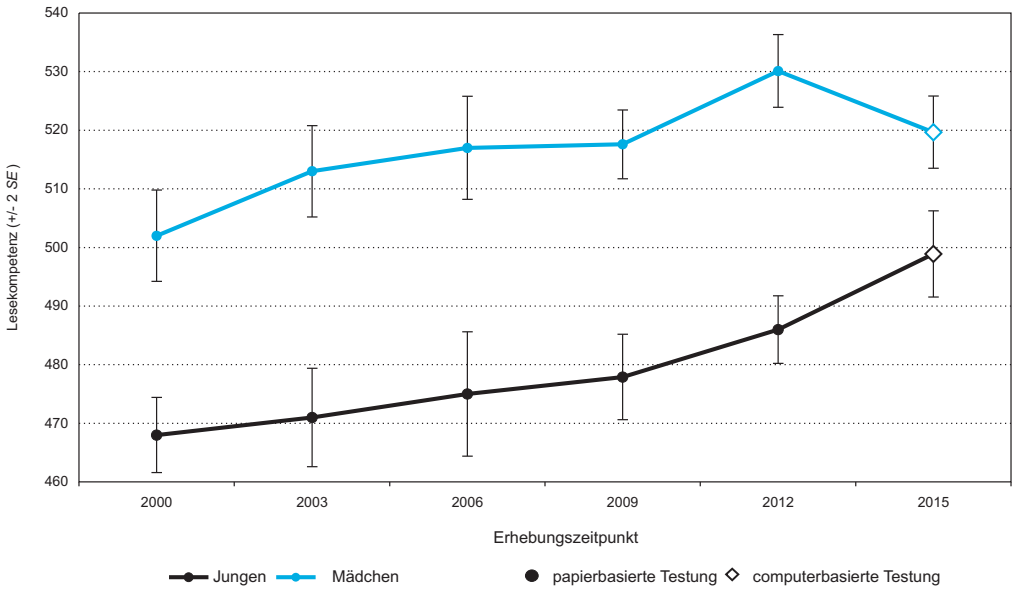


Abbildung 7.12: Veränderung der mittleren Lesekompetenz in Deutschland von 2000 bis 2015 für Mädchen und Jungen

bei den Mädchen erkennbar, bei welchem sich die Jungen den Mädchen nicht annäherten. Auf Basis der Daten der PISA-Erhebung 2015 ändert sich dieses Bild und deutet einen Leistungssprung der Jungen an. Die durchschnittliche Lesekompetenz der Jungen ist 2015 statistisch signifikant höher als in allen früheren Erhebungsrunden. Im Vergleich zu 2009 ist die mittlere Lesekompetenz der Jungen um 21 Punkte und im Vergleich zu 2012 um 13 Punkte gestiegen. Die Mädchen erreichen 2015 eine ähnlich hohe Lesekompetenz wie in den vorherigen PISA-Erhebungen. Im Vergleich zu 2012 ist die durchschnittliche Lesekompetenz der Mädchen zwar leicht gesunken, allerdings ist dieser Unterschied statistisch nicht bedeutsam. Durch die Steigerung der Leseleistung der Jungen nähern sich die mittleren Werte der Lesekompetenz der Mädchen und Jungen 2015 erstmals an. Während die durchschnittliche Lesekompetenz der Mädchen im Jahr 2000 um 34 Punkte, in den Jahren 2003 und 2006 um 42 Punkte, 2009 um 40 Punkte und 2012 um 44 Punkte höher war als jene der Jungen, unterscheidet sich die Lesekompetenz von Jungen und Mädchen in der PISA-Studie 2015 nur noch um 21 Punkte. Wie im Abschnitt zu den Unterschieden zwischen Mädchen und Jungen im internationalen Vergleich bereits erläutert, kann es mehrere Gründe für die Abnahme der Geschlechterunterschiede bei der PISA-Studie 2015 im Vergleich zu früheren Erhebungen geben. Für Deutschland lässt sich nicht ausschließen, dass der Wechsel zur computerbasierten Testung günstigere Effekte bei den Jungen als bei den Mädchen hatte, denn Ergebnisse verschiedener Studien weisen darauf hin, dass Jungen eine höhere Lesekompetenz erreichen, wenn diese am Computer erfasst wird (Horne, 2007; OECD, 2011). Eine mögliche Ursache könnte eine gesteigerte Lesemotivation oder auch Übung der

© Waxmann Verlag GmbH

Jungen am Computer sein. Die PISA-Studie 2009 zeigte, dass Mädchen in Deutschland eine höhere Lesefreude und Lesemotivation hatten als Jungen. Auch die Lesestrategien variierten zwischen Jungen und Mädchen. Diese stärker ausgeprägte Lesemotivation der Mädchen sowie die daraus resultierenden elaborierteren Lesestrategien konnten 2009 die höhere Lesekompetenz der Mädchen zu einem großen Teil erklären (Artelt, Naumann & Schneider, 2010). Außerdem gab es bereits in der Erhebung im Jahr 2009 Anzeichen, dass Jungen in Deutschland häufiger im Internet lesen als Mädchen. So nehmen mehr Jungen als Mädchen regelmäßig an Foren und Online-Diskussionen teil (Artelt et al., 2010). Dies spricht dafür, dass Jungen in Deutschland möglicherweise mehr Freude am Lesen im Internet beziehungsweise am digitalen Lesen haben als am Lesen gedruckter Materialien. Dass in der PISA-Studie 2015 der Test zur Erfassung der Lesekompetenz erstmals am Computer bearbeitet wurde, könnte demnach zu einer Steigerung der Lesefreude und Lesemotivation der Jungen und somit zu einer Verringerung des Geschlechterunterschieds in der Lesekompetenz geführt haben.

7.4 Zusammenfassung und Ausblick

Mit der PISA-Studie 2015 bestätigt sich das positive Bild der Ergebnisse zur Lesekompetenz aus der PISA-Erhebung im Jahr 2012, bei welcher die durchschnittliche Lesekompetenz der Fünfzehnjährigen in Deutschland erstmals über dem Durchschnitt der OECD-Staaten lag. Seit der ersten PISA-Studie im Jahr 2000 steigerte sich die Lesekompetenz der Jugendlichen in Deutschland im Mittel kontinuierlich bis zur Erhebung 2012 und liegt 2015 zum zweiten Mal in Folge über dem OECD-Mittelwert und zudem im oberen Drittel in der Rangreihe aller OECD-Staaten. Die Ergebnisse zur Lesekompetenz der PISA-Studie 2015 lassen sich damit als Stabilisierung auf einem hohen, gegenüber dem OECD-Mittelwert sogar überdurchschnittlichen Niveau charakterisieren. Die Entwicklung der mittleren Lesekompetenz der Fünfzehnjährigen in Deutschland über den Zeitraum von 2000 bis 2015 lässt sich als kontinuierlicher Aufwärtstrend beschreiben (OECD, 2016b).

Diese positive Entwicklung der Lesekompetenz der Jugendlichen in Deutschland ist erfreulich und kann vermutlich als ein Hinweis darauf verstanden werden, dass die Maßnahmen zur Förderung der Lesekompetenz, die seit der PISA-Studie 2000 umgesetzt wurden, nicht erfolglos geblieben sind. Einen wichtigen Stellenwert nehmen hierbei die sieben zentralen Handlungsfelder der KMK (KMK, 2002) ein, in welchen die Förderung der Lesekompetenz betont wurde. Als Konsequenz wurden vor allem auf Länderebene gezielte Handlungen zur Verbesserung der Lesekompetenz ergriffen. Außerdem wurden für ganz Deutschland die Bildungsstandards mit einem starken Fokus auf die Lesekompetenz eingeführt (KMK, 2004; 2005a; 2005b; 2014). Sicherlich ist zu sehen, dass die durchschnittliche Leistungssteigerung nicht alle Gruppen gleichermaßen betrifft. So ist hervorzuheben, dass insbesondere an den Gymnasien die Förderung leistungsstarker Schülerinnen und Schüler gelungen ist.

© Waxmann Verlag GmbH

Auch wenn sich die durchschnittliche Lesekompetenz in Deutschland deutlich verbessert hat, gibt es Staaten, in denen die Lesekompetenz der Jugendlichen noch deutlich höher ist als in Deutschland. Beispielsweise übersteigt die mittlere Lesekompetenz der Fünfzehnjährigen in Singapur, Kanada, Finnland und Irland diejenige der Jugendlichen in Deutschland merklich. Zudem ist in den OECD-Staaten, deren Mittelwert der Lesekompetenz höher ist als in Deutschland, sowohl die Streuung als auch der Anteil leseschwacher Fünfzehnjähriger geringer als in Deutschland. Insbesondere die Beispiele Irland und Kanada zeigen, dass eine hohe durchschnittliche Lesekompetenz bei gleichzeitig geringen mittleren Differenzen zwischen den leistungsschwächsten und leistungsstärksten Schülerinnen und Schülern innerhalb eines Staates möglich ist. Zudem ist der Anteil leseschwacher Schülerinnen und Schüler in Irland und Kanada vergleichsweise gering. Diesen Staaten gelingt es demnach, die Lesekompetenz in der Breite, das heißt sowohl bei leistungsschwachen als auch bei leistungsstarken Schülerinnen und Schülern, zu fördern. In Deutschland ist zwar seit 2009 die Gruppe der besonders lesestarken Schülerinnen und Schüler deutlich gewachsen, allerdings hat sich die Gruppe der besonders leseschwachen Schülerinnen und Schüler insgesamt nicht verändert. Daher zeigt sich im Vergleich mit anderen Staaten, dass es insbesondere bei der Förderung besonders leistungsschwacher Schülerinnen und Schüler in Deutschland noch Verbesserungspotenzial gibt.

Für die Zukunft bleibt also die Aufgabe, Maßnahmen zur Förderung zu ergreifen, die zum einen die mittlere Lesekompetenz der Jugendlichen weiter steigern und dabei die Streuung der Lesekompetenz innerhalb Deutschlands verringern. Die ermutigenden Ergebnisse belegen dabei den möglichen Erfolg geeigneter Maßnahmen. Sie legen nahe, insbesondere die Schulen bei einer offensichtlich erfolgreichen Arbeit weiter zu unterstützen.

In der PISA-Studie 2015 erreichen die Mädchen weiterhin eine höhere Lesekompetenz als die Jungen, allerdings hat sich diese Geschlechterdifferenz im Vergleich zu früheren Erhebungsrunden und auch im Vergleich zu einigen anderen Staaten sowie dem OECD-Durchschnitt stark verringert. Möglicherweise könnte die Umstellung vom papierbasierten auf das computerbasierte Testen einen Beitrag zur Reduktion der Geschlechterdifferenz in Deutschland geliefert haben. Plausible Begründungen für eine im Vergleich zu früheren (papierbasierten) Erhebungen höhere Lesekompetenz von Jungen am Computer könnten eine gesteigerte Lesemotivation und veränderte Lesestrategien der Jungen am Computer sein. Da die Lesekompetenz im Rahmen der PISA-Studie 2015 als Nebendomäne und nicht als Hauptdomäne erfasst wurde, erlauben die Daten keine weiterführenden Analysen zu diesen Annahmen. Anders wird dies in der kommenden PISA-Erhebungsrunde 2018 sein: Dann wird die Lesekompetenz wiederum die Hauptdomäne sein und entsprechend differenziert (und ebenfalls am Computer) erfasst und berichtet werden.

© Waxmann Verlag GmbH

Literatur

Allington, R. L. (2002). You can't learn much from books you can't read. *Educational Leadership, 60* (3), 16–19.

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., Naumann, J. & Schneider, W. (2010). Lesemotivation und Lernstrategien. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 73–112). Münster: Waxmann.

Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz. Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 69–140). Opladen: Leske + Budrich.

Beck, L., von Dewitz, N. & Titz, C. (2015). Durchgängige Leseförderung. *BiSS-Journal, Bildung durch Sprache und Schrift* (3), 4–11. Zugriff am 17.10.2016. Verfügbar unter http://www.biss-sprachbildung.de/pdf/BiSS-Journal_III_November_2015.pdf

Bundesministerium für Bildung und Forschung. (2015). *Bekanntmachung des Bundesministeriums für Familie, Senioren, Frauen und Jugend und des Bundesministeriums für Bildung und Forschung von Richtlinien zur Förderung von Vorhaben zur Evaluation von Konzepten und Maßnahmen der Sprachförderung, Sprachdiagnostik und Leseförderung – Evaluationsvorhaben der zweiten Runde – im Rahmen der Bund-Länder-Initiative „Bildung durch Sprache und Schrift (BiSS)"*. Zugriff am 17.10.2016. Verfügbar unter https://www.bmbf.de/foerderungen/bekanntmachung-1003.html

Chall, J. S. (1983). *Stages of reading development.* New York, NY: McGraw-Hill.

Coulombe, S., Tremblay, J.-F. & Marchand, S. (2004). *Literacy scores, human capital and growth across fourteen OECD countries*. Ottawa: Statistics Canada.

Gold, A. & Dubowy, M. (2013). *Frühe Bildung. Lernförderung im Elementarbereich*. Stuttgart: Kohlhammer.

Heinze, N. & Schnurr, J.-M. (2010). Informationskompetenz als Baustein für lebenslanges Lernen. In M. H. Breitner, C. Voigtländer & K. Sohns (Hrsg.), *Perspektiven des lebenslangen Lernens – dynamische Bildungsnetzwerke, Geschäftsmodelle, Trends* (S. 183–193). Berlin: Gito.

Hohn, K., Schiepe-Tiska, A., Sälzer, C. & Artelt, C. (2013). Lesekompetenz in PISA 2012: Veränderungen und Perspektiven. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 11–45). Münster: Waxmann.

Horne, J. (2007). Gender differences in computerised and conventional educational tests. *Journal of Computer Assisted Learning, 23*, 47–55. http://dx.doi.org/10.1111/j.1365-2729.2007.00198.x

Keller-Loibl, K. & Brandt, S. (2015). *Leseförderung in öffentlichen Bibliotheken*. Berlin: De Gruyter Saur.

KM Bayern (2011) = Bayerisches Staatsministerium für Unterricht und Kultus, Staatsinstitut für Schulqualität und Bildungsforschung (Hrsg.). (2011). *ProLesen. Auf dem Weg zur Le-*

© Waxmann Verlag GmbH

seschule – 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.* Zugriff am 17.10.2016. Verfügbar unter https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2002/2002_10_07-Pisa-2000-Zentrale-Handlungsfelder.pdf

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 (2005a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005a). *Bildungsstandards im Fach Deutsch für den Hauptschulabschluss. Beschluss vom 15.10.2004*. Neuwied: Luchterhand.

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

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.

Naumann, J., Artelt, C., Schneider, W. & Stanat, P. (2010). Lesekompetenz von PISA 2000 bis PISA 2009. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 23–71). Münster: Waxmann.

Nickolaus, R., Geissel, B. & Gschwendtner, T. (2008). Die Rolle der Basiskompetenzen Mathematik und Lesefähigkeit in der beruflichen Ausbildung und die Entwicklung mathematischer Fähigkeiten im ersten Ausbildungsjahr. *Berufs- und Wirtschaftspädagogik – online, 14,* 1–14. Zugriff am 17.10.2016. Verfügbar unter http://www.bwpat.de/ausgabe14/nickolaus_etal_bwpat14.pdf

OECD. (2011). *PISA 2009 Results: Students on line: Digital technologies and performance*. Zugriff am 17.10.2016. Verfügbar unter https://www.oecd.org/pisa/pisaproducts/48270093.pdf

OECD. (2016a). PISA 2015: Reading framework. In OECD (Hrsg.), *PISA 2015 assessment and analytical framework: Science, reading, mathematic and financial literacy* (S. 47–61). Paris: OECD. http://dx.doi.org/10.1787/9789264255425-4-en

OECD. (2016b). *PISA 2015 Results (Volume I): Excellence and equity in education*. Paris: OECD.

OECD. (2016c). *The survey of adult skills. Reader's companion* (2. Aufl.). Paris: OECD. http://dx.doi.org/10.1787/9789264258075-en

Prediger, S., Wilhelm, N., Büchter, A., Gürsoy, E. & Benholz, C. (2015). Sprachkompetenz und Mathematikleistung – Empirische Untersuchung sprachlich bedingter Hürden in

© Waxmann Verlag GmbH

den Zentralen Prüfungen 10. *Journal für Mathematik-Didaktik, 36*, 77–104. http://dx.doi.org/10.1007/s13138-015-0074-0

Sälzer, C. & Prenzel, M. (2013). PISA 2012 – eine Einführung in die aktuelle Studie. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 11–45). Münster: Waxmann.

Schneider, W., Baumert, J., Becker-Mrotzek, M., Hasselhorn, M., Kammermeyer, G., Rauschenbach, T., Roßbach, H.-G., Roth, H.-J., Rothweiler, M., Stanat, P., Schmiedeler, S. & Chilla, S. (2012). *Expertise „Bildung durch Sprache und Schrift (BISS)"*. Berlin: Bundesministerium für Bildung und Forschung. Zugriff am 17.10.2016. Verfügbar unter http://www.bmbf.de/pubRD/BISS_Expertise.pdf

Schneider, W., Küspert, P. & Krajewski, K. (2013). *Die Entwicklung mathematischer Kompetenzen*. Paderborn: Schöningh.

Schröder, L., Keller, H., Dintsioudi, A. & List, M. (2013). *Sprachliche Bildung im Kita-Alltag. Gespräche mit Kindern anregen und lebendig gestalten*. Berlin: Cornelsen.

UNESCO. (2005). *Education for all. Literacy for life*. Paris: Unesco.

Wang, S., Jiao, H., Young, M. J., Brooks, T. & Olson, J. (2008). 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*, 5–24. http://dx.doi.org/10.1177/0013164407305592

© Waxmann Verlag GmbH

© Waxmann Verlag GmbH

8 Soziale Herkunft und Kompetenzerwerb

Katharina Müller & Timo Ehmke

Viele Studien haben gezeigt, dass in der Bildungsbeteiligung und beim Kompetenzerwerb von Schülerinnen und Schülern soziale Disparitäten bestehen. Im Folgenden wird beschrieben, welchen durchschnittlichen sozioökonomischen und soziokulturellen Status Fünfzehnjährige in den teilnehmenden Staaten aufweisen und welche Variabilität der Sozialstruktur es in Deutschland und in den übrigen OECD-Staaten gibt. Der soziale Hintergrund der Fünfzehnjährigen wird in PISA anhand mehrerer Indizes beschrieben. Neben dem auf der *internationalen Standardklassifikation der Berufe* (ISCO) basierenden *höchsten sozioökonomischen Status* (HISEI) und dem *Index of Economic, Social and Cultural Status* (ESCS) kommen für die Stichprobe aus Deutschland auch die sogenannten *EGP*-Klassen zum Einsatz. Im Fokus des Kapitels steht der Zusammenhang zwischen sozioökonomischer Herkunft bzw. ökonomischem, kulturellem und sozialem Status und der naturwissenschaftlichen Kompetenz. Diese Kopplung wird regressionsanalytisch untersucht und die Steigung des sozialen Gradienten sowie die Varianzaufklärung international vergleichend und mit Blick auf den PISA-Zyklus 2006 berichtet. Auf der Basis der EGP-Klassifikation werden Kompetenzunterschiede von Schülerinnen und Schülern in Deutschland in den verschiedenen Sozialschichten in den Blick genommen und die soziale Lage der Jugendlichen in Deutschland differenziert dargestellt. Die Analysen zeigen, dass es in Deutschland einen deutlichen Zusammenhang zwischen der sozialen Herkunft der Eltern und dem naturwissenschaftlichen Kompetenzniveau der fünfzehnjährigen Schülerinnen und Schüler gibt. Auch wenn die Abstände im Kompetenzniveau zwischen sozialen Schichten in den letzten Jahren kleiner geworden sind, bleibt das Bemühen um eine Verringerung sozialer Disparitäten des Kompetenzerwerbs und der Bildungsbeteiligung nach wie vor eine der vorrangigen bildungspolitischen Aufgaben.

Es kann als gesicherter Befund angesehen werden, dass es in allen Bildungssystemen – wenn auch unterschiedlich stark ausgeprägte – soziale Ungleichheiten gibt. In der Bundesrepublik beschäftigte man sich seit der in den 1960er-Jahren einsetzenden Bildungsreform zunehmend mit sozialen Ungleichheiten in der Bildungsbeteiligung und den damit verbundenen sozial- und bildungspolitischen Problemen (Maaz, Baumert & Cortina, 2008). Die wissenschaftliche Auseinandersetzung mit dem Thema durchlief

© Waxmann Verlag GmbH

unterschiedliche Phasen (Maaz, Baumert & Trautwein, 2009; Maaz, Hausen, McElvany & Baumert, 2006). Im Zuge der in PISA erstmals offen zutage getretenen umfassenden Unterschiede zwischen sozialer Herkunft und Bildungsbeteiligung bzw. Kompetenzerwerb (Baumert & Schümer, 2001; Ehmke, Hohensee, Heidemeier & Prenzel, 2004) kam es zu einer breiten öffentlichen Debatte darüber, inwieweit es innerhalb des Bildungssystems gelingt, allen Mitgliedern der Gesellschaft gerechte Chancen zum Lernen und zur Entwicklung von Kompetenzen zu bieten.

Inzwischen ist vielfach belegt, dass die familiäre Herkunft und die damit verbundenen Unterschiede in der ökonomischen, sozialen und kulturellen Ausstattung die Kompetenzentwicklung und die Schulleistungen von Schülerinnen und Schülern sowohl direkt (primäre Herkunftseffekte) als auch vermittelt über bildungs- und sozialschichtabhängige Bildungsentscheidungen (sekundäre Herkunftseffekte) beeinflussen. Auch gelingt es zunehmend, die dahinterliegenden Wirkungsmechanismen (vgl. zusammenfassend etwa Baumert, Maaz & Trautwein, 2009; Maaz, Neumann & Baumert, 2014) empirisch zu fassen. Neben jenen Bereichen, die im deutschen Schulsystem potenziell Ungleichheit verstärken, wie etwa die diversen Bildungsübergänge (Dumont, Maaz, Neumann & Becker, 2014; Faust & Roßbach, 2014; Granato & Ulrich, 2014; Watermann, Daniel & Maaz, 2014), kann auch die Wirksamkeit von Förder- und Interventionsmöglichkeiten etwa im Bereich der vorschulischen Bildung (Hasselhorn & Kuger, 2014; Seyda, 2009), der Sprachförderung (Paetsch, Wolf, Stanat & Darsow, 2014) oder der Ganztagsschulangebote (Strietholt, Manitius, Berkemeyer & Bos, 2015; Züchner & Fischer, 2014) in den Blick genommen werden.

Welchen Beitrag können die aktuellen PISA-Befunde angesichts dieser zunehmend ausdifferenzierten Befundlage zu den sozialen Disparitäten der Bildungsbeteiligung leisten? Der Mehrwert der turnusmäßigen Zusammenhangsanalysen zwischen Kompetenz und sozialer Herkunft kann vor allem in den folgenden Bereichen gesehen werden: Die regelmäßigen PISA-Erhebungen zeigen auf, inwieweit es in den einzelnen OECD-Staaten gelingt, Bildungsgerechtigkeit (*Equity*) hinsichtlich des Erwerbs von Kompetenzen zu erlangen (vgl. OECD, 2007). Im Sinne eines *Benchmarking* können Staaten im Hinblick darauf verglichen werden, in welchem Umfang eine Entkopplung von Kompetenzerwerb und sozialer Herkunft möglich ist. Der internationale Vergleich sozialer Disparitäten ist dann von besonderem Interesse, wenn es Staaten gibt, in denen hohe Kompetenzniveaus mit gleichzeitig geringen sozialen Unterschieden einhergehen. Zudem können im Sinne eines *Monitorings* die Ergebnisse aufeinanderfolgender Erhebungen über die Zeit verglichen werden. Auf diese Weise lässt sich feststellen, ob sich soziale Disparitäten langfristig abschwächen, stabil bleiben oder sich sogar verstärken. Die soziale Herkunft wird dabei operationalisiert über strukturelle Familienmerkmale, von denen man annimmt, dass sie sich auf den Erwerb von Kompetenzen und die Bildungskarriere auswirken. Durch die Ergebnisse von PISA lässt sich so mit statistischen Mitteln beschreiben, wie sehr diese Merkmale der Elternhäuser mit Unterschieden in den Kompetenzen (Naturwissenschaften, Mathematik, Lesen) zusammenhängen. Neben dieser international vergleichenden Perspektive können die Ergebnisse zur sozialen Lage und zum Zusammen-

© Waxmann Verlag GmbH

hang zwischen Kompetenz und sozialer Herkunft aufgrund des repräsentativ angelegten Stichprobendesigns als *Referenz* für nationale Untersuchungen herangezogen werden (Baumert & Maaz, 2010).

In PISA 2000 und PISA 2003 ergaben die Befunde für die Schülerinnen und Schüler aus Deutschland, dass die erreichten Kompetenzen der Jugendlichen eng mit ihrer sozialen Herkunft zusammenhängen. Dies zeigte sich insbesondere an den Unterschieden zwischen Schülerinnen und Schülern aus unterschiedlichen sozialen Lagen in Bezug auf die Lesekompetenz, aber auch in Bezug auf Mathematik und die Naturwissenschaften. Im internationalen Vergleich gab es damals nur wenige OECD-Staaten, in denen die Kopplung zwischen dem sozioökonomischen Status der Eltern und den gemessenen Kompetenzen ihrer Kinder so eng war wie in Deutschland (Baumert & Schümer, 2001; Ehmke et al., 2004; OECD, 2001, 2004).

In der dritten, vierten und fünften Erhebungsrunde von PISA in den Jahren 2006, 2009 und 2012 wurde für Deutschland festgestellt, dass die Kennwerte für den sozialen Gradienten der Lesekompetenz im Vergleich zu den Befunden aus PISA 2000 bedeutsam abgenommen hatten (Ehmke & Baumert, 2007; Ehmke & Jude, 2010; Müller & Ehmke, 2013; OECD, 2007, 2010, 2013). Die nationalen Ländervergleichsstudien zur Überprüfung der Bildungsstandards in Deutschland haben zudem darauf verwiesen, dass zwischen den Bundesländern zum Teil bedeutsame Unterschiede im Grad der Kopplung von sozialer Herkunft und erreichten Kompetenzen bestehen (Knigge & Köller, 2010; Knigge & Leucht, 2010; Kuhl, Haag, Federlein, Weirich & Schipolowksi, 2016; Kuhl, Siegle & Lenski, 2013; Richter, Kuhl & Pant, 2012).

Mit PISA 2015 liegen jetzt Ergebnisse für den sechsten Erhebungszeitpunkt im internationalen Vergleich vor. Veränderungen in den Kennwerten für soziale Disparitäten können inzwischen über einen Zeitraum von fünfzehn Jahren untersucht werden. Der vorliegende Bericht behandelt insbesondere folgende Fragen: (1) Wie hoch ist der Zusammenhang zwischen dem sozioökonomischen Status der Jugendlichen und ihrer naturwissenschaftlichen Kompetenz in den OECD-Staaten? (2) Inwieweit hat sich dieser Zusammenhang zwischen den Erhebungen PISA 2006 und PISA 2015 in den einzelnen OECD-Staaten verändert? (3) Inwiefern unterscheiden sich Jugendliche unterschiedlicher sozialer Herkunft in Deutschland hinsichtlich bildungsrelevanter Merkmale ihrer Elternhäuser? (4) Haben sich zwischen PISA 2000 und PISA 2015 die sozialen Disparitäten in der Lesekompetenz und in der Bildungsbeteiligung für die Kohorte der Fünfzehnjährigen verringert oder vergrößert?

Zu den beiden ersten Fragestellungen wird über Analysen berichtet, bei denen die OECD-Staaten miteinander verglichen werden. Dabei wurde geprüft, wie sich der Zusammenhang zwischen der sozialen Herkunft und der naturwissenschaftlichen Kompetenz der Jugendlichen in den Erhebungen seit PISA 2006 darstellt. Das Jahr 2006 wurde deshalb gewählt, da hier Naturwissenschaften erstmals als Hauptdomäne erhoben wurden. Den Fragestellungen drei und vier wird im Rahmen nationaler Analysen nachgegangen. Diese zielen auf die detailliertere Untersuchung sozialer Disparitäten in Deutschland ab. Um langfristige Entwicklungen beschreiben zu können, wird bei diesen

© Waxmann Verlag GmbH

Analysen die Lesekompetenz einbezogen. Da Lesen im Jahr 2000 erstmals Hauptdomäne war, lassen sich damit Trendanalysen in Bezug auf den Zusammenhang zwischen sozialer Herkunft und Kompetenz – wenn auch mit der gebotenen Vorsicht – über einen Zeitraum von fünfzehn Jahren darstellen.

8.1 Erfassung der sozialen Herkunft in PISA

Um Effekte der sozialen Herkunft auf den Kompetenzerwerb möglichst umfassend untersuchen zu können, erfolgt die Erfassung in der Regel über mehrere Dimensionen. Von zentraler Bedeutung ist die sozioökonomische Stellung der Familien, aus der die Schülerinnen und Schüler stammen. Die sozioökonomische Stellung beschreibt die relative Position der Eltern in einer sozialen Hierarchie, deren Rangfolge sich daraus ergibt, in welchem Umfang sie über finanzielle Mittel oder bestimmte Gestaltungs- und Einflussmöglichkeiten verfügen. Der sozioökonomische Status wird, da detaillierte Informationen über Einkommen, Macht oder soziale Anerkennung schwer zugänglich sind, in der Regel über die Berufstätigkeit erfasst. Im Anschluss an die sozialkapitaltheoretischen Ansätze von Bourdieu (1982, 1983) und Coleman (1988, 1996) werden zunehmend auch Merkmale des kulturellen und des sozialen Kapitals der Familien einbezogen, um die soziale Herkunft von Individuen etwas differenzierter zu beschreiben. Kulturelles und soziales Kapital umfassen die erworbenen Fertigkeiten, Fähigkeiten und Kenntnisse von Individuen, die deren Handlungs- und Erwerbsmöglichkeiten erweitern und die sozioökonomische Stellung positiv beeinflussen können. Zum kulturellen Kapital können die kulturbezogenen Sachgüter, Bildungszertifikate, die in Familien vorhandenen Wahrnehmungs-, Denk- und Deutungsmuster sowie Wertorientierungen und Einstellungen gezählt werden. Das soziale Kapital bestimmt sich durch das soziale Netzwerk und die sozialen Beziehungen, in denen Kinder und Jugendliche aufwachsen. Die sozioökonomische Stellung spiegelt die Ressourcen in einer Familie wider, die mit den Einkommensverhältnissen und dem materiellen Wohlstand einhergehen. Dagegen sind die mit dem sozialen und kulturellen Kapital verbundenen Lebensbedingungen mit dem Lebensstil und dem sozialen Milieu verknüpft. Auch diese Merkmale können als Ressourcen verstanden werden, durch die sich unterschiedlich lernförderliche Umgebungen und damit differenzielle Entwicklungsmilieus abbilden lassen.

In der PISA-Berichterstattung werden verschiedene Indikatoren zur Messung der sozialen Herkunft verwendet: der *International Socio-Economic Index of Occupational Status* (ISEI), der Index of *Economic, Social and Cultural Status* (ESCS) und die *Erikson-Goldthorpe-Portocarero-Klassifikation* (EGP). Ausgangsbasis der verwendeten Indikatoren ist die internationale Standardklassifikation der Berufe (*International Standard Classification of Occupations*, ISCO-08) (ILO, 1969, 1990, 2012). Das von der internationalen Arbeitsorganisation (*International Labour Office*, ILO) für die amtliche Statistik und international vergleichende Forschung konzipierte Instrument wurde in der zuletzt angepassten Version eingesetzt (ILO, 2012). Mit dem ISCO werden die Angaben der

© Waxmann Verlag GmbH

Schülerinnen und Schülern zur Berufstätigkeit der Eltern auf vier Ebenen hierarchisch klassifiziert und die einzelnen Berufe verschiedenen Berufsgruppen zugeordnet.

Anhand des von Ganzeboom, de Graaf, Treimann und de Leeuw (1992) entwickelten *International Socio-Economic Index of Occupational Status* (ISEI) wird der sozioökonomische Status der beruflichen Tätigkeit erfasst (Ganzeboom & Treimann, 2003). Dabei werden die Berufe in eine eindimensionale Ordnung gebracht werden, welche die Stellung des Berufsinhabers in einer sozialen Hierarchie zum Ausdruck bringen (Schimpl-Neimanns, 2004). Hierbei wird angenommen, dass Berufe spezifische Qualifikationen erfordern und ein bestimmtes Einkommen zur Folge haben und insofern Bildungsabschlüsse vermittelt über die berufliche Tätigkeit das Einkommen bestimmen. Der Index ist so konstruiert, dass er den indirekten, über den Beruf vermittelten Einfluss des Bildungsabschlusses auf das Einkommen maximiert und den direkten Effekt minimiert. Die kontinuierliche, hierarchische Skala des ISEI, die entsprechend der Revision der ISCO-Codes aktualisiert wurde (Ganzeboom & Treimann, 2012), reicht von 11 Punkten (Reinigungskraft) bis 90 Punkten (Richter) (Ganzeboom & Treimann, 1996), wobei höhere Werte einen höheren sozioökonomischen Status zum Ausdruck bringen. Sofern Werte von Vater und Mutter vorlagen, ist der jeweils höhere Wert (*Highest International Socio-Economic Index of Occupational Status*; HISEI) in die Berechnungen eingegangen.

Neben dem kontinuierlich konstruierten ISEI kommt in den nationalen Analysen außerdem das von Erikson, Goldthorpe und Portocarero entwickelte *EGP-Klassifikationsschema* zum Einsatz (Erikson & Goldthorpe, 2002; Erikson, Goldthorpe & Portocarero, 1979). Bei diesem kategorialen Ansatz werden Berufe in diskrete Klassen unterteilt. Hier wird angenommen, dass sich die Mitglieder innerhalb der Klassen relativ ähnlich sind (internale Homogenität) und zwischen den Klassen unterscheiden (externale Heterogenität) (Ganzeboom, de Graaf & Treimann, 1992). Grundlage der EGP-Klassifikation ist ein Kategoriensystem, mit dem die Berufe nach der Art der Tätigkeit, der Stellung im Beruf und der Weisungsbefugnis geordnet werden. Im Vergleich zur eindimensionalen Gliederung des sozioökonomischen Index (HISEI) können so qualitative Abstufungen zwischen sozialen Klassen vorgenommen werden. Durch die Verbindung von gegliederter Abstufung und typologischer Klassifikation, können Berufsgruppen theoretisch fundiert und anschaulich beschrieben werden. Anhand der EGP-Klassifikation können Unterschiede in der sozialen Entwicklungsumwelt der Schülerinnen und Schüler sichtbar gemacht werden, die mit dem Einkommen der Eltern und deren Zugang zu Bildung, Macht und gesellschaftlicher Anerkennung verknüpft sind. Deshalb wurde in den nationalen Analysen zu PISA seit Beginn auch die EGP-Klassifikation eingesetzt und berichtet. Dabei wurde und wird statt des elfstufigen Modells von Erikson und Kollegen ein Klassifikationsschema mit sechs Klassen verwendet (1979).

Zur Quantifizierung sozialer Disparitäten auf internationaler Ebene wird in PISA ein globaler Index gebildet, der sowohl soziokulturelle als auch sozioökonomische Merkmale der sozialen Herkunft beinhaltet. Der *Index of Economic, Social and Cultural Status* (ESCS) berücksichtigt den auf der ISCO-Kodierung basierenden höchsten sozioökonomischen Status (HISEI) der Eltern, den Bildungsabschluss der Eltern und Informationen

© Waxmann Verlag GmbH

über den Besitz von Kultur- und Wohlstandsgütern (OECD, 2016; OECD, in Vorbereitung). Im Vergleich zur Vorhersagekraft der Einzelindikatoren ermöglicht es der (über OECD-Staaten hinweg *z*-standardisierte) ESCS, Unterschiede in den Kompetenzen, die auf Herkunftsmerkmale zurückzuführen sind, möglichst breit vorherzusagen. Während die differenzierte Betrachtung der Einzelindikatoren die Möglichkeit bietet, die Bedeutung der verschiedenen Komponenten der sozialen Herkunft abzuschätzen, erlaubt der aggregierte ESCS die gleichzeitige Berücksichtigung unterschiedlicher Herkunftsmerkmale (Ehmke & Siegle, 2005). Dabei ist zum einen zu beachten, dass bei der Interpretation von Analysen globaler Indikatorensysteme wie dem ESCS die bessere Vorhersagekraft mit Unschärfen in der theoretischen Grundlegung und der konzeptionellen Bedeutung einhergeht (Caro & Cortés, 2012). Zum anderen sollte berücksichtigt werden, dass es sich bei der Operationalisierung des kulturellen Kapitals um das in Form von Kulturgütern objektivierte kulturelle Kapital handelt. Kenntnisse über den Besitz von Kultur- oder Wohlstandsgütern sind allenfalls als Annäherung an das relationale Konzept Bourdieus zu verstehen (Kramer, 2011).

Da der ESCS über die verschiedenen PISA-Erhebungen hinweg zum Teil unterschiedlich operationalisiert und um länderspezifische Items ergänzt wurde, basieren vergleichende Analysen zwischen den Erhebungsrunden auf einem eigens skalierten Trend-ESCS. Dazu wurden die ESCS-Werte der vergangenen PISA-Erhebungsrunden im Rahmen einer konkurrenten rückwärtigen Skalierung neu berechnet. Dies erfolgte dadurch, dass die Variablen, die in den ESCS eingehen, anhand der aktuellen Vorschriften einheitlich umgerechnet und die ESCS-Werte für die vorangegangenen PISA-Erhebungszeiträume seit 2000 neu ermittelt wurden (OECD, in Vorbereitung).

8.2 Der internationale Vergleich: Zusammenhang zwischen sozialer Herkunft und naturwissenschaftlicher Kompetenz

Im Folgenden steht die soziale Herkunft der fünfzehnjährigen Schülerinnen und Schüler in den OECD-Staaten im Vordergrund. Dazu wird zunächst auf den sozioökonomischen Status der Schülerinnen und Schüler eingegangen und anhand des HISEI die mittlere Ausprägung und die Streuung über die Staaten hinweg vergleichend beschrieben. Zudem wird untersucht, ob Jugendliche, deren Eltern eine höhere sozioökonomische Stellung haben, auch über eine höhere naturwissenschaftliche Kompetenz verfügen. Anschließend wechseln wir vom HISEI zum ESCS-Index, der im Vergleich zum HISEI neben sozioökonomischen auch soziokulturelle Merkmale der sozialen Herkunft integriert. Zentral ist nicht nur die Frage, wie sich der Zusammenhang zwischen naturwissenschaftlicher Kompetenz und ökonomischer, sozialer und kultureller Herkunft verhält, sondern auch, wie sich Veränderungen in den sozialen Disparitäten der naturwissenschaftlichen Kompetenz zwischen PISA 2006 und PISA 2015 beschreiben lassen.

Um die Kopplung zwischen der sozialen Herkunft und der Kompetenz zu analysieren, kommen in PISA zwei Kennwerte, der *soziale Gradient* und das *Maß der aufge-*

© Waxmann Verlag GmbH

klärten Varianz, zum Einsatz. Die beiden Kennwerte sind das Ergebnis einer linearen Regression, anhand derer die Unterschiedlichkeit in einem abhängigen Merkmal durch Unterschiede in einem oder mehreren anderen unabhängigen Merkmalen vorhergesagt wird. Als Prädiktoren werden in den folgenden internationalen Analysen der HISEI sowie der ESCS verwendet. Die in der linearen Regression ermittelte *Steigung* des sozialen Gradienten gibt an, inwieweit gemäß der Regressionsgeraden die durchschnittliche Kompetenz der Schülerinnen und Schüler zunehmen würde, wenn sich der Prädiktor (HISEI bzw. ESCS) um eine Standardabweichung vergrößert. Der soziale Gradient quantifiziert damit den Zusammenhang zwischen Kompetenz und dem Merkmal der sozialen Herkunft. Ein niedriger sozialer Gradient bzw. eine schwach ansteigende Regressionsgerade deutet auf geringe Kompetenzunterschiede im Zusammenhang mit der sozialen Herkunft hin, ein hoher sozialer Gradient dagegen auf umfängliche Kompetenzunterschiede. Die *Varianzaufklärung* gibt im Vergleich dazu als statistischer Kennwert an, zu welchem Anteil sich Unterschiede in der erreichten Kompetenz in einem Staat durch die soziale Herkunft vorhersagen lassen. Dabei wird die Gesamtvarianz der Kompetenzwerte in zwei Anteile zerlegt, nämlich einen, der auf die soziale Herkunft als Vorhersagekriterium zurückgeht und beschreibt, welcher Anteil der Gesamtvarianz dadurch erklärt wird, sowie einen Restanteil, der sich nicht durch die soziale Herkunft vorhersagen lässt. Je geringer der prozentuale Anteil der Gesamtvarianz der Kompetenz ist, desto schlechter kann das Kompetenzniveau vorhergesagt werden. Die beiden Indikatoren erfassen also unterschiedliche Aspekte der Kopplung zwischen Kompetenz und sozialer Herkunft. Der soziale Gradient gibt Auskunft über die Stärke des Zusammenhangs zwischen Kompetenz und sozialer Herkunft. Die Varianzaufklärung spiegelt wider, wie präzise sich Unterschiede in der Kompetenz durch die soziale Herkunft vorhersagen lassen.

8.2.1 Kopplung zwischen naturwissenschaftlicher Kompetenz und sozioökonomischem Status (HISEI)

Bevor auf den Zusammenhang zwischen sozioökonomischem Status und naturwissenschaftlicher Kompetenz eingegangen wird, beschreiben wir zunächst die durchschnittliche Lage und Variabilität des sozioökonomischen Status in den OECD-Staaten. Die Perzentilbänder in Abbildung 8.1 veranschaulichen grafisch für jeden OECD-Staat den Abstand zwischen den 5 Prozent der aus den sozioökonomisch schwächsten und stärksten Elternhäusern stammenden Jugendlichen. Die Länge des Perzentilbandes illustriert damit die Streuung des sozioökonomischen Status: Je schmaler das Perzentilband, desto homogener ist die Verteilung innerhalb des Staates.

In Deutschland liegt der mittlere sozioökonomische Status bei 51.5 Punkten und ist damit wie auch in Österreich, Frankreich (jeweils M = 51.3), Belgien (M = 52.7) oder Irland (M = 52.8) nicht signifikant vom OECD-Durchschnitt (M = 51.8) verschieden. In den skandinavischen Staaten Finnland (M = 53.0), Schweden (M = 57.7) oder Norwegen (M = 61.8) sowie der Schweiz (M = 53.0) oder Kanada (M = 58.1) liegt der sozioökono-

© Waxmann Verlag GmbH

mische Status signifikant über dem OECD-Mittelwert. In Griechenland (*M* = 50.0), Italien (*M* = 49.7), Portugal (*M* = 48.6), Spanien (*M* = 48.0) sowie der Türkei (36.7) finden sich unterdurchschnittliche Werte.

Der Blick auf die Standardabweichung veranschaulicht die Unterschiedlichkeit der Verteilung des sozioökonomischen Status innerhalb der Staaten. Homogene soziale Lagen finden sich etwa in den sieben Staaten mit dem höchsten sozioökonomischen Status sowie den Niederlanden, Estland, Korea, Japan, der Slowakei oder der Tschechischen Republik. Auch in Deutschland liegt die Streuung im sozioökonomischen Status mit 20.4 Punkten signifikant unter dem OECD-Durchschnitt (*SD* = 21.1). Deutlich darüber liegen etwa die Streuungen in Griechenland (*SD* = 23.1) und Spanien (*SD* = 23.4).

Verglichen mit der vorangegangenen PISA-Erhebung 2012, als zum ersten Mal mit der neuen und angepassten Berufskodierung ISCO-08 gearbeitet wurde, zeigt sich in PISA 2015, dass der mittlere HISEI-Wert in der OECD tendenziell höher ist (PISA 2012: *M* = 50.6; PISA 2015: *M* = 51.8). Bei der Standardabweichung, die in PISA 2012 deutlich größer ausfiel als in den vorangegangenen Zyklen, zeigt sich im OECD-Durchschnitt eine leichte Zunahme (PISA 2012: *SD* = 20.8; PISA 2015: *SD* = 21.1). Bei der Interpretation von Trendaussagen zum sozioökonomischen Status, die auf der alten Berufskodierung ISCO-88 basieren, muss jedoch Folgendes berücksichtigt werden: Der damit verbundene, zum Teil umfängliche Anstieg in Bezug auf Mittelwert und Standardabweichung hat einen Einfluss auf die Zusammenhangsanalysen zwischen sozioökonomischer Herkunft und Kompetenz. Auf ausführliche Trendanalysen, die auf dem HISEI beruhen, wird daher an dieser Stelle verzichtet. Für den ESCS hingegen wurde rückwirkend für vorangegangene Zyklen eine umfangreiche Rekodierung der eingehenden Indizes sowie eine Trendskalierung vorgenommen, die vergleichende Analysen über alle PISA-Zyklen hinweg erlaubt (OECD, in Vorbereitung). Deshalb wird bei den Veränderungen in Bezug auf die sozialen Disparitäten primär auf den ESCS zurückgegriffen.

Die Ergebnisse der linearen Regression, in die als Prädiktor für die naturwissenschaftliche Kompetenz der höchste sozioökonomische Status (HISEI) eingegangen ist, sind in Tabelle 8.1 wiedergegeben. Darin wird bezogen auf die naturwissenschaftliche Kompetenz für alle OECD-Staaten (Spalte 1) die Steigung des sozialen Gradienten (Spalte 4) sowie die Varianzaufklärung (Spalte 6) berichtet. Der Achsenabschnitt (Spalte 2) sagt dabei die Höhe des Kompetenzwertes vorher, den eine Jugendliche bzw. ein Jugendlicher bei einem durchschnittlichen sozioökonomischen Status erlangen würden.

Anhand der Tabelle wird deutlich, dass es in allen OECD-Staaten einen positiven Zusammenhang zwischen dem sozioökonomischen Status und der naturwissenschaftlichen Kompetenz gibt. Im OECD-Durchschnitt liegt die Steigung des sozialen Gradienten bei 31 Punkten. Um diesen Wert würde sich die naturwissenschaftliche Kompetenz der Fünfzehnjährigen vergrößern, wenn sich der sozioökonomische Status um eine Standardabweichung (*SD* = 21.1, vgl. Abbildung 8.1) erhöhte. Mit 499 Punkten wird jener Kompetenzwert vorhergesagt, den eine Schülerin beziehungsweise ein Schüler bei einem mittleren sozioökonomischen Status (*M* = 51.8) erlangt.

© Waxmann Verlag GmbH

OECD-Staaten	*MW*	*(SE)*	*SD*
Norwegen	61.8	(0.4)	19.7
Island	60.4	(0.4)	19.4
Israel	59.8	(0.6)	20.8
Kanada	58.1	(0.4)	20.2
Schweden	57.7	(0.5)	20.4
Australien	57.5	(0.3)	20.8
Neuseeland	57.4	(0.4)	20.0
Vereinigtes Königreich	56.4	(0.5)	21.4
Dänemark	56.2	(0.5)	21.9
Niederlande	54.5	(0.4)	20.0
Vereinigte Staaten	54.1	(0.6)	21.6
Schweiz	53.0	(0.5)	21.7
Finnland	53.0	(0.6)	21.3
Irland	52.8	(0.6)	21.4
Belgien	52.7	(0.5)	22.0
Estland	52.3	(0.4)	20.5
Slowenien	52.3	(0.3)	21.4
Korea	52.1	(0.5)	19.0
OECD-Durchschnitt	51.8	(0.1)	21.1
Deutschland	51.5	(0.5)	20.4
Österreich	51.3	(0.5)	21.1
Frankreich	51.3	(0.5)	21.6
Lettland	51.0	(0.5)	22.2
Japan	50.6	(0.3)	19.9
Griechenland	50.0	(0.6)	23.1
Italien	49.7	(0.4)	21.5
Luxemburg	48.8	(0.2)	22.4
Portugal	48.6	(0.6)	22.9
Spanien	48.0	(0.7)	23.4
Slowakei	47.8	(0.5)	20.7
Ungarn	47.6	(0.5)	21.5
Tschechische Republik	47.5	(0.3)	19.4
Polen	45.7	(0.5)	22.3
Chile	44.6	(0.5)	22.1
Mexiko	40.0	(0.5)	21.3
Türkei	36.7	(0.7)	19.8

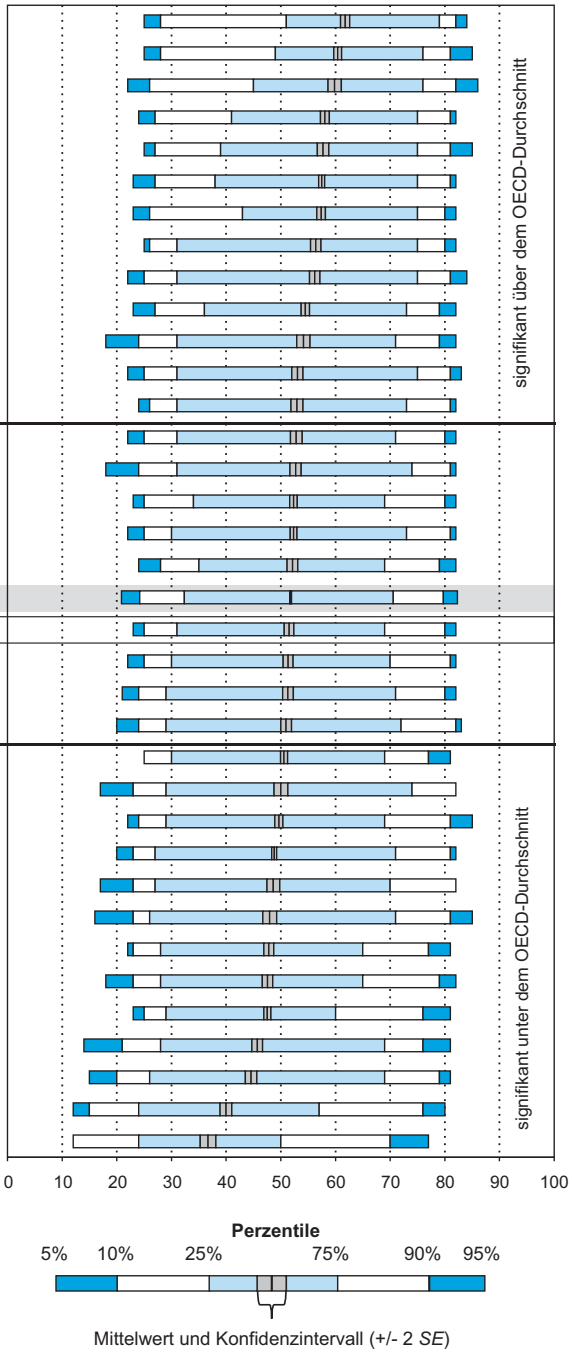


Abbildung 8.1: Verteilung des sozioökonomischen Status (HISEI) in den OECD-Staaten

© Waxmann Verlag GmbH

Tabelle 8.1: Zusammenhang zwischen Naturwissenschaftskompetenz und ökonomischem Status im internationalen Vergleich (Prädiktorvariable: HISEI)

OECD-Staaten	Naturwissenschafts-kompetenz- *Achsen-abschnitt*	*(SE)*	Steigung des sozialen Gradienten *Steigung*	*(SE)*	Stärke des Zusammenhangs R^2	*(SE)*
Island	470	(1.8)	18.2	(2.0)	3.2	(0.7)
Japan	545	(2.7)	21.2	(1.6)	4.5	(0.6)
Korea	518	(2.8)	26.2	(2.0)	5.8	(0.9)
Norwegen	491	(2.0)	26.7	(1.5)	6.5	(0.7)
Kanada	526	(1.8)	26.3	(1.2)	7.3	(0.6)
Australien	509	(1.4)	28.7	(1.3)	7.6	(0.7)
Finnland	533	(1.9)	27.0	(1.7)	7.8	(0.9)
Mexiko	429	(2.1)	20.4	(1.5)	7.9	(1.1)
Lettland	495	(1.4)	22.9	(1.4)	8.2	(1.0)
Estland	537	(2.0)	26.9	(1.6)	8.3	(0.9)
Polen	512	(2.3)	25.3	(1.6)	8.4	(1.0)
Türkei	447	(4.1)	25.4	(2.8)	8.4	(1.9)
Vereinigte Staaten	499	(2.5)	28.4	(1.8)	8.5	(1.0)
Dänemark	503	(2.1)	25.9	(1.5)	8.7	(1.0)
Vereinigtes Königreich	510	(2.1)	29.7	(1.6)	8.9	(0.9)
Irland	505	(2.1)	27.8	(1.4)	9.7	(0.9)
Spanien	501	(1.7)	25.4	(1.2)	9.8	(0.9)
Italien	487	(2.4)	29.5	(1.7)	10.3	(1.1)
Niederlande	509	(2.1)	35.1	(2.0)	10.6	(1.2)
Israel	465	(2.7)	35.1	(2.1)	10.7	(1.2)
Schweden	492	(2.5)	35.2	(2.0)	11.1	(1.2)
Neuseeland	511	(2.2)	37.9	(2.0)	11.7	(1.1)
Slowakei	478	(2.3)	34.5	(2.1)	12.1	(1.3)
Österreich	501	(2.0)	35.2	(1.8)	12.7	(1.2)
Griechenland	462	(3.2)	30.6	(1.9)	12.7	(1.3)
Chile	462	(2.3)	30.3	(1.5)	13.0	(1.2)
Deutschland	521	(2.2)	37.9	(1.8)	13.2	(1.1)
Slowenien	515	(1.3)	34.7	(1.4)	13.2	(1.0)
Schweiz	509	(2.5)	37.0	(1.8)	14.3	(1.2)
Portugal	509	(2.0)	33.4	(1.6)	14.9	(1.4)
Tschechische Republik	506	(2.0)	41.0	(1.9)	15.1	(1.2)
Frankreich	507	(1.6)	39.1	(1.6)	16.0	(1.3)
Belgien	508	(1.6)	40.6	(1.6)	18.0	(1.2)
Ungarn	491	(2.1)	43.3	(1.9)	20.4	(1.5)
Luxemburg	495	(1.2)	46.3	(1.3)	23.0	(1.1)
OECD-Durchschnitt	499	(0.4)	31.1	(0.3)	10.9	(0.2)

Anmerkung: Prädiktorvariable ist 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

Wenn auch mit zunehmendem sozioökonomischem Status in der Tendenz das Kompetenzniveau steigt, so wird anhand der Ergebnisse doch sichtbar, dass sich die Steigungen der sozialen Gradienten zwischen den Staaten unterscheiden. Zu den Staaten, in denen die Steigung nicht signifikant vom OECD-Mittelwert verschieden ist, gehören Australien, die Vereinigten Staaten, das Vereinigte Königreich, Italien, Israel, die Slowakei, Griechenland, Chile und Portugal. Niedriger als im OECD-Mittel sind sie in Island,

© Waxmann Verlag GmbH

Japan, Korea, Norwegen, Kanada, Finnland, Mexiko, Lettland, Estland, Polen, der Türkei, Dänemark, Irland und Spanien. Deutlich ausgeprägter als im OECD-Durchschnitt ist die Steigung in den Niederlanden, Schweden, Neuseeland, Österreich, Slowenien, der Schweiz, der Tschechischen Republik, Frankreich, Belgien, Ungarn und Luxemburg. Auch in Deutschland ist die Steigung des sozialen Gradienten bedeutsam höher als im OECD-Mittelwert.

Die Varianzaufklärung, also jener Anteil der Unterschiede in der naturwissenschaftlichen Kompetenz, der durch den sozioökonomischen Status vorhergesagt werden kann, liegt über alle OECD-Staaten hinweg bei durchschnittlich 10.9 Prozent. In Norwegen, Kanada, Finnland, Polen oder Dänemark liegt die Varianzaufklärung statistisch bedeutsam darunter. In Irland, Spanien, Italien, den Niederlanden, Schweden oder Österreich unterscheidet sich dieser Anteil nicht signifikant und in Deutschland, der Schweiz, Frankreich, Belgien oder Luxemburg liegt der Wert signifikant über dem OECD-Mittelwert.

8.2.2 Kopplung zwischen naturwissenschaftlicher Kompetenz und ökonomischem, kulturellem und sozialem Status (ESCS)

Werden die Berechnungen des sozialen Gradienten auf Grundlage des ESCS-Indexes vorgenommen, der neben sozioökonomischen auch soziokulturelle Merkmale beinhaltet, fallen Steigung des sozialen Gradienten und Varianzaufklärung erwartungsgemäß höher aus. Tabelle 8.2 sind die Ergebnisse für den sozialen Gradienten zu entnehmen, die anhand des erweiterten Index ermittelt wurden.

Die Steigung liegt im OECD-Mittelwert bei 38 Punkten. Würde der ESCS also um eine Standardabweichung (SD = 1.0) höher ausfallen, läge die naturwissenschaftliche Kompetenz der Fünfzehnjährigen in den OECD-Staaten um 38 Punkte höher. Dieser Wert liegt in Deutschland bei 42 Punkten und unterscheidet sich damit wie auch in der Slowakei, Polen, Irland, Israel, dem Vereinigten Königreich, Japan, Finnland und Norwegen nicht signifikant vom OECD-Mittelwert. In Ungarn, Luxemburg, Frankreich, Belgien, der Tschechischen Republik, Österreich, der Schweiz, Neuseeland, Slowenien, den Niederlanden, Schweden, Australien und Korea ist der Zusammenhang zwischen Merkmalen der sozioökonomischen bzw. soziokulturellen Herkunft und naturwissenschaftlicher Kompetenz stärker ausgeprägt als im OECD-Durchschnitt, während er in Chile, Portugal, Spanien, Griechenland, den Vereinigten Staaten, Mexiko, Dänemark, Italien, der Türkei, Kanada, Lettland, Estland und Island signifikant weniger stark ausgeprägt ist.

Anhand der farblichen Markierung in Tabelle 8.2 wird ersichtlich, in welchen Staaten sich die Varianzaufklärung (Spalte 6) bedeutsam vom OECD-Mittelwert von 13 Prozent unterscheidet und wo sie im Bereich des Durchschnitts liegt. In Deutschland wird ein Varianzanteil der naturwissenschaftlichen Kompetenz von 16 Prozent durch den ökonomischen, sozialen und kulturellen Status (ESCS) aufgeklärt. Damit liegt Deutschland

© Waxmann Verlag GmbH

Tabelle 8.2: Zusammenhang zwischen Naturwissenschaftskompetenz und ökonomischem, kulturellem und sozialem Status im internationalen Vergleich (Prädiktorvariable: ESCS)

OECD-Staaten	Naturwissenschafts-kompetenz- *Achsen-abschnitt*	*(SE)*	Steigung des sozialen Gradienten *Steigung*	*(SE)*	Stärke des Zusammenhangs *R^2*	*(SE)*
Island	454	(2.3)	27.6	(2.1)	4.9	(0.8)
Estland	533	(2.0)	32.3	(1.8)	7.8	(0.9)
Norwegen	482	(1.8)	37.5	(2.2)	8.2	(0.9)
Lettland	502	(1.5)	26.4	(1.6)	8.7	(1.0)
Kanada	511	(1.8)	33.7	(1.5)	8.8	(0.7)
Türkei	455	(4.8)	20.4	(2.1)	9.0	(1.9)
Italien	484	(2.4)	29.9	(1.7)	9.6	(1.0)
Finnland	521	(2.1)	40.5	(2.3)	10.0	(1.0)
Japan	547	(2.7)	41.9	(2.2)	10.1	(1.0)
Korea	525	(2.6)	44.3	(2.7)	10.1	(1.3)
Dänemark	483	(2.0)	33.5	(1.7)	10.4	(1.0)
Vereinigtes Königreich	504	(2.0)	37.5	(1.9)	10.5	(1.0)
Mexiko	440	(2.4)	19.4	(1.1)	10.9	(1.3)
Israel	461	(2.8)	41.6	(2.3)	11.2	(1.3)
Vereinigte Staaten	494	(2.5)	33.2	(1.8)	11.4	(1.1)
Australien	500	(1.5)	43.8	(1.5)	11.7	(0.8)
Schweden	481	(2.6)	43.6	(2.2)	12.2	(1.1)
Niederlande	502	(2.2)	46.8	(2.6)	12.5	(1.3)
Griechenland	458	(3.3)	33.8	(2.1)	12.5	(1.3)
Irland	497	(2.2)	37.6	(1.6)	12.7	(1.0)
Polen	518	(2.3)	40.1	(2.0)	13.4	(1.3)
Spanien	507	(1.8)	26.9	(1.1)	13.4	(1.1)
Slowenien	512	(1.3)	42.7	(1.5)	13.5	(0.9)
Neuseeland	508	(2.1)	48.7	(2.6)	13.6	(1.2)
Portugal	514	(2.1)	30.8	(1.5)	14.9	(1.4)
Schweiz	500	(2.5)	42.8	(1.9)	15.6	(1.2)
Deutschland	511	(2.3)	41.7	(1.9)	15.8	(1.2)
Österreich	492	(2.0)	45.4	(2.0)	15.9	(1.3)
Slowakei	467	(2.3)	41.5	(2.3)	16.0	(1.4)
Chile	463	(2.2)	32.3	(1.4)	16.9	(1.3)
Tschechische Republik	505	(2.0)	51.7	(2.1)	18.8	(1.2)
Belgien	496	(1.7)	48.2	(1.8)	19.3	(1.3)
Frankreich	505	(1.7)	57.0	(2.0)	20.3	(1.3)
Luxemburg	481	(1.2)	41.3	(1.1)	20.8	(1.0)
Ungarn	487	(2.1)	46.6	(1.9)	21.4	(1.4)
OECD-Durchschnitt	494	(0.4)	38.4	(0.3)	12.9	(0.2)

Anmerkung: 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

zusammen etwa mit Österreich, der Schweiz, Frankreich oder Belgien signifikant über dem OECD-Durchschnitt.

Wird die Varianzaufklärung des ESCS zusammen mit dem durchschnittlichen naturwissenschaftlichen Kompetenzniveau der Jugendlichen betrachtet, erlaubt dies einen Einblick in die Frage, inwieweit etwa ein hohes Kompetenzniveau mit einem hohen Vor-

© Waxmann Verlag GmbH

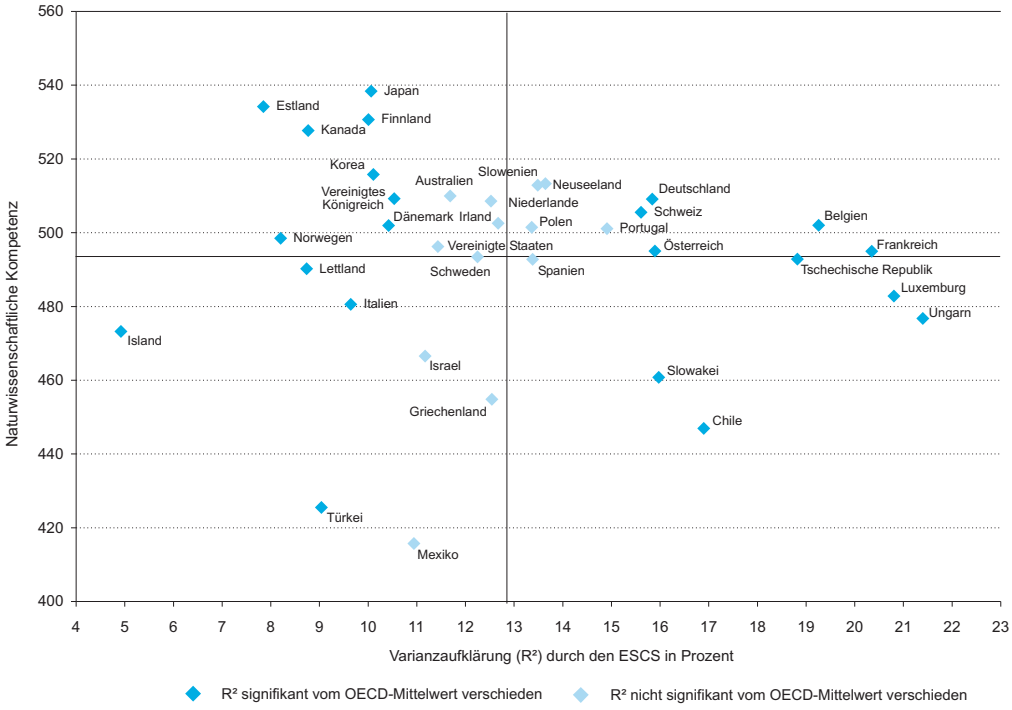
hersagewert durch den ESCS einhergeht. In Abbildung 8.2 sind die Staaten zum einen nach dem Niveau der naturwissenschaftlichen Kompetenz und zum anderen nach dem Varianzanteil, der durch den ESCS aufgeklärt wird, angeordnet. Die Einteilung in die vier Quadranten gibt Auskunft darüber, inwieweit die Staaten bezüglich dieser beiden Merkmale über oder unter dem OECD-Durchschnitt liegen (naturwissenschaftliche Kompetenz M = 493; Varianzaufklärung R^2 = 13 %). Durch die gemeinsame Abbildung ist es etwa möglich, jene Staaten zu identifizieren, in denen es gelingt, Schülerinnen und Schülern ein hohes Kompetenzniveau zu vermitteln, und in denen gleichzeitig der Zusammenhang zwischen sozioökonomischer und soziokultureller Herkunft (Quadrant links oben) gering ist – eine im Hinblick auf den Ausgleich sozialer Ungleichheit insgesamt wünschenswerte Konstellation. Andererseits können jene Staaten identifiziert werden, in denen der Zusammenhang hoch, d.h. die sozialen Disparitäten stark ausgeprägt, und das Kompetenzniveau niedrig ausfällt (Quadrant rechts unten).

Der Quadrant oben rechts repräsentiert jene Staaten, in denen das Kompetenzniveau über dem Durchschnitt liegt und die erreichten Kompetenzen in relativ hohem Maß durch Unterschiede in der sozialen, kulturellen und ökonomischen Herkunft vorhergesagt werden. Der Datenpunkt für Deutschland ist in diesem Quadranten zu finden. Die überdurchschnittlich hohen Werte im Bereich der naturwissenschaftlichen Kompetenz werden stark durch den ESCS vorhergesagt. Ein ähnliches Muster findet sich etwa in Belgien oder der Schweiz.

Ebenfalls überdurchschnittlich hohe Kompetenzniveaus zeigen sich in Finnland, Kanada oder Estland, die sich im Quadranten oben links wiederfinden. Im Vergleich zum Quadranten oben rechts allerdings werden die Kompetenzunterschiede unterdurchschnittlich stark durch die soziale Herkunft aufgeklärt. Diese Staaten zeichnen sich dadurch aus, dass die im Vergleich zum OECD-Mittelwert hohen naturwissenschaftlichen Kompetenzen weniger stark durch die soziale, ökonomische und kulturelle Herkunft vorhergesagt werden. Am Beispiel dieser Staaten wird deutlich, dass ein überdurchschnittliches Kompetenzniveau nicht zwangsläufig mit hohen sozialen Disparitäten einhergehen muss und eine Entkopplung von Kompetenz und sozialer Herkunft mit gleichzeitig hohem Kompetenzniveau möglich ist.

In den beiden unteren Quadranten befinden sich jene Staaten, in denen ein unterdurchschnittliches Kompetenzniveau vorgefunden wurde. In der Slowakei, Chile oder Ungarn (Quadrant unten rechts) zeigte sich, dass die insgesamt gering ausgeprägten Kompetenzniveaus im Bereich der Naturwissenschaften überdurchschnittlich stark durch Unterschiede in den Merkmalen der sozialen Herkunft vorhergesagt werden können, in Island, der Türkei oder Italien (Quadrant unten links) konnten die Unterschiede weniger stark durch diese Merkmale erklärt werden.

© Waxmann Verlag GmbH



Anmerkung: Dunkel gefüllte Datenpunkte unterscheiden sich bezüglich der Varianzaufklärung durch den ESCS signifikant vom OECD-Durchschnitt.

Abbildung 8.2: Naturwissenschaftliche Kompetenz und Varianzaufklärung (R^2) durch den ökonomischen, sozialen und kulturellen Status (ESCS)

Abbildung 8.2 gibt zu erkennen, dass es durchaus Staaten gibt, in denen Schülerinnen und Schüler hohe Kompetenzwerte erlangen, ohne dass diese in überdurchschnittlichem Maße durch die soziale, kulturelle und ökonomische Herkunft vorhergesagt würden. Es gibt demnach keine Indizien dafür, dass ein hohes Kompetenzniveau unbedingt mit starken sozialen Disparitäten einhergehen muss. Für Staaten wie Deutschland heißt dies, dass eine Verringerung sozialer Ungleichheit auch mit einem stabil hohen oder sogar noch gesteigerten Kompetenzniveau möglich wäre.

8.2.3 Veränderungen in den sozialen Disparitäten der naturwissenschaftlichen Kompetenz

Um der Frage nachzugehen, inwieweit sich die sozialen Disparitäten in den erreichten naturwissenschaftlichen Kompetenzen in den vergangenen Jahren verändert haben, werden im Folgenden die aktuellen Kennwerte des sozialen Gradienten mit denen vorangegangener PISA-Erhebungen verglichen. Da die Naturwissenschaftskompetenzen zuletzt in PISA 2006 als Hauptdomäne erhoben wurden, beziehen sich die Angaben auf den Zeitraum der vergangenen neun Jahre von 2006 bis 2015. Als abhängige Variable gehen

© Waxmann Verlag GmbH

Tabelle 8.3: Veränderungen in den sozialen Disparitäten des Kompetenzerwerbs über die Zeit im internationalen Vergleich des Trend-ESCS-Index und seines Zusammenhangs mit der naturwissenschaftlichen Kompetenz zwischen PISA 2006 und PISA 2015

	2006				2015				2006–2015			
OECD-Staaten	*b*	*(SE)*	*R²*	*(SE)*	*b*	*(SE)*	*R²*	*(SE)*	Δ *b*	*(SE)*	Δ *R²*	*(SE)*
Australien	42.0	(1.4)	12.0	(0.7)	43.8	(1.5)	11.7	(0.8)	1.8	(2.0)	-0.4	(1.1)
Belgien	46.4	(1.8)	20.0	(1.3)	48.2	(1.8)	19.3	(1.3)	1.7	(2.5)	-0.7	(1.8)
Chile	38.8	(1.7)	23.3	(1.8)	32.3	(1.4)	16.9	(1.3)	**-6.4**	(2.2)	**-6.4**	(2.3)
Dänemark	40.4	(2.2)	14.0	(1.5)	33.5	(1.7)	10.4	(1.0)	**-6.9**	(2.8)	**-3.6**	(1.8)
Deutschland	46.4	(2.1)	19.8	(1.5)	41.7	(1.9)	15.8	(1.2)	-4.7	(2.8)	**-4.0**	(2.0)
Estland	30.6	(2.0)	8.9	(1.1)	32.3	(1.8)	7.8	(0.9)	1.7	(2.7)	-1.0	(1.4)
Finnland	31.0	(1.6)	8.2	(0.9)	40.5	(2.3)	10.0	(1.0)	**9.5**	(2.8)	1.8	(1.4)
Frankreich	52.4	(2.3)	22.3	(1.8)	57.0	(2.0)	20.3	(1.3)	4.6	(3.1)	-1.9	(2.2)
Griechenland	35.4	(2.2)	14.6	(1.7)	33.8	(2.1)	12.5	(1.3)	-1.5	(3.0)	-2.1	(2.1)
Irland	36.3	(2.1)	13.2	(1.4)	37.6	(1.6)	12.7	(1.0)	1.3	(2.6)	-0.5	(1.7)
Island	30.4	(1.8)	7.5	(0.8)	27.6	(2.1)	4.9	(0.8)	-2.8	(2.8)	**-2.6**	(1.1)
Israel	41.8	(2.5)	10.3	(1.0)	41.6	(2.3)	11.2	(1.3)	-0.1	(3.4)	0.9	(1.6)
Italien	30.7	(1.6)	10.3	(1.0)	29.9	(1.7)	9.6	(1.0)	-0.8	(2.3)	-0.6	(1.4)
Japan	39.7	(2.7)	8.5	(1.0)	41.9	(2.2)	10.1	(1.0)	2.2	(3.5)	1.6	(1.4)
Kanada	32.4	(1.4)	8.5	(0.7)	33.7	(1.5)	8.8	(0.7)	1.2	(2.0)	0.3	(1.0)
Korea	31.4	(3.4)	7.0	(1.4)	44.3	(2.7)	10.1	(1.3)	**12.9**	(4.3)	3.1	(1.9)
Lettland	30.2	(2.4)	9.3	(1.4)	26.4	(1.6)	8.7	(1.0)	-3.8	(2.9)	-0.5	(1.7)
Luxemburg	39.8	(1.1)	22.5	(1.1)	41.3	(1.1)	20.8	(1.0)	1.6	(1.6)	-1.7	(1.5)
Mexiko	24.3	(1.3)	16.2	(1.7)	19.4	(1.1)	10.9	(1.3)	**-4.9**	(1.7)	**-5.2**	(2.1)
Neuseeland	48.7	(1.6)	15.7	(1.0)	48.7	(2.6)	13.6	(1.2)	0.0	(3.0)	-2.0	(1.6)
Niederlande	44.1	(2.3)	16.3	(1.7)	46.8	(2.6)	12.5	(1.3)	2.7	(3.5)	-3.8	(2.1)
Norwegen	36.6	(2.5)	8.6	(1.1)	37.5	(2.2)	8.2	(0.9)	0.9	(3.3)	-0.4	(1.4)
Österreich	45.8	(3.0)	15.8	(2.0)	45.4	(2.0)	15.9	(1.3)	-0.4	(3.6)	0.1	(2.4)
Polen	40.6	(1.8)	14.7	(1.1)	40.1	(2.0)	13.4	(1.3)	-0.5	(2.7)	-1.4	(1.7)
Portugal	27.8	(1.4)	16.3	(1.5)	30.8	(1.5)	14.9	(1.4)	3.0	(2.0)	-1.4	(2.0)
Schweden	37.4	(2.0)	11.0	(1.0)	43.6	(2.2)	12.2	(1.1)	**6.1**	(3.0)	1.2	(1.5)
Schweiz	43.2	(1.7)	16.3	(1.2)	42.8	(1.9)	15.6	(1.2)	-0.4	(2.6)	-0.7	(1.7)
Slowakei	45.9	(2.6)	19.6	(1.9)	41.5	(2.3)	16.0	(1.4)	-4.4	(3.5)	-3.6	(2.4)
Slowenien	47.2	(1.6)	17.5	(1.1)	42.7	(1.5)	13.5	(0.9)	**-4.6**	(2.3)	**-4.0**	(1.4)
Spanien	23.8	(1.0)	12.4	(1.1)	26.9	(1.1)	13.4	(1.1)	**3.2**	(1.5)	0.9	(1.6)
Tschechische Republik	50.3	(2.4)	16.1	(1.3)	51.7	(2.1)	18.8	(1.2)	1.4	(3.2)	2.7	(1.8)
Türkei	27.8	(3.0)	15.1	(2.8)	20.4	(2.1)	9.0	(1.9)	**-7.5**	(3.7)	-6.1	(3.4)
Ungarn	44.9	(1.9)	21.1	(1.6)	46.6	(1.9)	21.4	(1.4)	1.7	(2.7)	0.3	(2.1)
Vereinigte Staaten	45.9	(2.3)	17.4	(1.6)	33.2	(1.8)	11.4	(1.1)	**-12.7**	(2.9)	**-6.0**	(2.0)
Vereinigtes Königreich	45.4	(1.9)	13.4	(1.2)	37.5	(1.9)	10.5	(1.0)	**-7.9**	(2.7)	-2.9	(1.6)
OECD-Durchschnitt	38.7	(0.4)	14.4	(0.2)	38.4	(0.3)	12.9	(0.2)	-0.4	(0.5)	**-1.4**	(0.3)

Anmerkung: Indikator für die soziale Herkunft ist der für Trendberechnungen konzipierte ökonomische, kulturelle und soziale Index (Trend-ESCS) *z*-standardisiert am OECD-Durchschnitt (*b*: sozialer Gradient; R^2: Varianzaufklärung).

Signifikante Unterschiede (p <.05) in den Kennwerten für die Regression zwischen PISA 2006 und PISA 2015 sind fettgedruckt.

die naturwissenschaftlichen Kompetenzen ein, Indikator für die soziale Herkunft ist der für Trendanalysen neu berechnete ökonomische, kulturelle und soziale Index (Trend-ESCS) (OECD, in Vorbereitung).

In Tabelle 8.3 sind für jeden OECD-Staat die sozialen Gradienten der naturwissenschaftlichen Kompetenz für PISA 2006 und PISA 2015 wiedergegeben. Zusätzlich zu den Steigungen des sozialen Gradienten (b) und dem Anteil der durch den ökonomischen, sozialen und kulturellen Status aufgeklärten Varianz (R^2) werden Unterschiede zwischen den Erhebungen in den beiden Koeffizienten einander gegenübergestellt.

© Waxmann Verlag GmbH

Für den OECD-Durchschnitt zeigt sich, dass sowohl die Kennwerte für die Steigung des sozialen Gradienten als auch die der Varianzaufklärung zurückgegangen sind. Von PISA 2006 nach PISA 2015 ist der soziale Gradient der naturwissenschaftlichen Kompetenz leicht zurückgegangen. Die Varianzaufklärung ist im OECD-Mittelwert statistisch signifikant von 14 Prozent auf 13 Prozent gesunken.

Eine statistisch bedeutsame Zunahme des sozialen Gradienten zeigt sich in Finnland, Korea, Schweden und Spanien. In Chile, Dänemark, Mexiko, Slowenien, der Türkei, den Vereinigten Staaten und im Vereinigten Königreich ist der soziale Gradient zwischen PISA 2006 und PISA 2015 statistisch bedeutsam zurückgegangen. Eine signifikante Verringerung der Varianzaufklärung ist in Chile, Dänemark, Island, Mexiko, Slowenien und den Vereinigten Staaten zu finden.

Auch Deutschland gehört zur Gruppe jener Staaten, in denen in PISA 2015 im Vergleich zu 2006 deutlich weniger Unterschiede in der naturwissenschaftlichen Kompetenz durch den ESCS vorhergesagt werden können. Beim sozialen Gradienten hingegen kann die Abnahme von 46 auf 42 Punkte nicht statistisch abgesichert werden. Insgesamt zeichnet sich damit bei den Kennwerten tendenziell eine wünschenswerte Entwicklung ab. Dennoch scheinen soziale Disparitäten in der naturwissenschaftlichen Kompetenz nach wie vor vorhanden zu sein, und es wird deutlich, dass die Herausforderung, den Zusammenhang zwischen naturwissenschaftlicher Kompetenz und sozialer Herkunft zu entkoppeln, immer noch besteht.

8.3 Vertiefende nationale Analysen: die soziale Herkunft der Jugendlichen in Deutschland

Im vorangegangenen Kapitel stand die Kopplung zwischen sozialer Herkunft und naturwissenschaftlicher Kompetenz im internationalen Vergleich der OECD-Staaten im Vordergrund. Zum Einsatz kamen dabei zwei Indikatoren, nämlich zum einen der HISEI, der den sozioökonomischen Status des Elternhauses abbildet, und zum anderen der ESCS, der neben dem sozioökonomischen Status auch den kulturellen und sozialen Status erfasst.

Mit den folgenden Analysen wenden wir den Blick auf die Schülerinnen und Schüler in Deutschland und untersuchen vertieft einzelne Aspekte ihrer sozialen Herkunft. Wir knüpfen dabei zunächst an die vorangegangenen Regressionsanalysen an und untersuchen sukzessive, inwieweit verschiedene bildungsrelevante Merkmale zur Vorhersage der naturwissenschaftlichen Kompetenz beitragen. Verwendet werden dazu neben dem sozioökonomischen Status und dem Bildungsniveau der Eltern auch eine Reihe der im Elternhaus zur Verfügung stehenden Besitztümer – Komponenten also, die auch in den globalen ESCS-Indikator eingehen. Dies erlaubt die Einschätzung der spezifischen Vorhersagekraft der einzelnen Merkmale auf die naturwissenschaftliche Kompetenz.

© Waxmann Verlag GmbH

Wie stark sind diese Merkmale bei Jugendlichen unterschiedlicher sozialer Herkunft ausgeprägt? Die EGP-Klassifikation, die eine Zuordnung der elterlichen Berufe zu soziale Klassen erlaubt, ermöglicht es, die Verteilung dieser Merkmale differenzierter zu beschreiben. Innerhalb der EGP-Klassen werden zunächst Veränderungen der Lesekompetenz und der Bildungsbeteiligung betrachtet. Anschließend wird die soziale Lage in den EGP-Klassen genauer beschrieben. Dabei gehen wir insbesondere auch auf die Ausprägung der Merkmale in der Gruppe der besonders leistungsstarken und leistungsschwachen Schülerinnen und Schüler ein.

8.3.1 Vorhersage der naturwissenschaftlichen Kompetenz durch Merkmale der sozialen Lage und des Bildungsweges

Regressionsanalytisch wird im Folgenden untersucht, wie gut sich durch verschiedene bildungsrelevante Merkmale der sozialen Herkunft die naturwissenschaftliche Kompetenz der Schülerinnen und Schüler in Deutschland vorhersagen lässt. Dabei interessieren neben dem Beitrag des sozioökonomischen Status der Eltern zur Vorhersage der Kompetenz auch deren Bildungsniveau, die Ausstattung mit verschiedenen kulturellen und lernrelevanten Besitztümern sowie Merkmale des Bildungsweges. Um den relativen Beitrag der einzelnen Merkmale zur Vorhersage der Kompetenz statistisch abschätzen zu können, wurden sukzessive vier aufeinander aufbauende Modelle gerechnet, deren Ergebnisse in Tabelle 8.4 abgebildet sind. Schrittweise wurde der sozioökonomische Status der Eltern (Modell I), das Bildungsniveau der Eltern (Modell II), die häuslichen, kulturellen und bildungsbezogenen Besitztümer (Modell III) sowie das Eintrittsalter in den Kindergarten und das Eintrittsalter in die Schule (Modell IV) einbezogen.

Modell I sagt – analog zum Vorgehen im internationalen Vergleich der sozialen Gradienten (vgl. Tabelle 8.1) – die naturwissenschaftliche Kompetenz durch den sozioökonomischen Status (HISEI) vorher. Die Steigung des sozialen Gradienten liegt in der naturwissenschaftlichen Kompetenz bei 36 Punkten, und rund 13 Prozent der Kompetenzunterschiede lassen sich durch den sozioökonomischen Status der Eltern erklären.[1]

In Modell II wird neben dem HISEI das Bildungsniveau der Eltern berücksichtigt. Die Kennwerte zeigen, dass beide Merkmale einen spezifischen Beitrag zur Vorhersage der naturwissenschaftliche Kompetenz haben. Zwar verringert sich gegenüber Modell I der Koeffizient für den sozioökonomischen Status. Dennoch kommt dem sozioökonomischen Status ein eigener Beitrag bei der Vorhersage der naturwissenschaftlichen Kompetenz zu, insbesondere wenn man das Bildungsniveau der Eltern als Kontrollvariable einbezieht.

Das Modell III berücksichtigt zusätzlich die Ausstattung mit häuslichen, kulturellen und bildungsbezogenen Besitztümern. Für die häuslichen Besitztümer kann ein deutlich

1 Abweichungen in den Regressionskoeffizienten zwischen Tabelle 8.1 und Tabelle 8.4, Modell I, sind auf Unterschiede in den Stichprobengrößen zurückzuführen, die durch die verschiedenen Regressionsmodelle bedingt sind.

© Waxmann Verlag GmbH

positiver, für die kulturellen Besitztümern ein negativer und für die bildungsbezogenen Besitztümer kein spezifischer Vorhersagebeitrag festgestellt werden. Insgesamt lassen sich in diesem Modell 19 Prozent der Varianz in der naturwissenschaftlichen Kompetenz aufklären. Vor dem Hintergrund, dass der sozioökonomische Status, das Bildungsniveau der Eltern und die häuslichen Besitztümer auch in den ESCS eingehen, illustrieren diese Befunde im Vergleich etwa zum HISEI (Modell I) auch die breitere Operationalisierung des Konstruktes und die höhere Vorhersagekraft.

In das letzte Modell (IV) gehen zusätzlich zwei Merkmale des früheren Bildungsweges des Schülers bzw. der Schülerin ein. Jugendliche, die den Kindergarten erst ab einem Alter von vier Jahren bzw. später besucht haben, erreichen, wenn die übrigen Merkmale kontrolliert werden, eine um 15 Punkte niedrigere naturwissenschaftliche Kompetenz als Jugendliche, die den Kindergarten mit drei Jahren oder früher besucht haben. Die Höhe der Regressionskoeffizienten für den sozioökonomischen Status, das Bildungsniveau der Eltern sowie die häuslichen und kulturellen Besitztümer verringert sich gegenüber Modell III nur geringfügig. Kein spezifischer Vorhersagebeitrag kommt dem Einschulungsalter zu. Insgesamt klärt dieses Modell 21 Prozent der Varianz in der naturwissenschaftlichen Kompetenz auf und veranschaulicht damit den jeweils unterschiedlichen Beitrag, den verschiedene Merkmale der sozialen Herkunft für die Vorhersage der naturwissenschaftlichen Kompetenz leisten.

Tabelle 8.4: Regressionsmodelle zur Schätzung von sozialen Disparitäten in der naturwissenschaftlichen Kompetenz

	Modell I		**Modell II**		**Modell III**		**Modell IV**	
	b	*(SE)*	b	*(SE)*	b	*(SE)*	b	*(SE)*
Naturwissenschaftliche Kompetenz Achsenabschnitt	517	(2.5)	517	(2.5)	517	(2.3)	513	(3.3)
Sozioökonomischer Status[1]	**36**	(1.7)	**30**	(1.8)	**22**	(1.6)	**22**	(1.6)
Bildungsniveau der Eltern [1]			**11**	(1.9)	**5**	(1.9)	**5**	(1.9)
Häusliche Besitztümer der Eltern [1]					**29**	(3.0)	**28**	(3.0)
Kulturelle Besitztümer[1]					**-6**	(2.4)	**-7**	(2.4)
Bildungsbezogene Besitztümer [1]					2	(1.9)	2	(1.8)
Kindergartenbesuch mit 4 Jahren oder älter [2]							**-15**	(1.9)
Einschulung im Alter von 5 Jahren [3]							-3	(2.7)
N	4595		4595		4595		4595	
R^2	0.13	(0.0)	0.14	(0.0)	0.19	(0.0)	0.21	(0.0)

* Die hier angegebene Regressionskonstante ist der geschätzte Mittelwert der naturwissenschaftlichen Kompetenz mit mittlerer Ausprägung der Prädiktoren HISEI, des Bildungsniveaus der Eltern sowie der Besitztümer der Eltern, falls die Variablen im Modell enthalten sind. Signifikante Partialregressionskoeffizienten ($p < .05$) sind fettgedruckt.
[1] *z*-standardisiert.
[2] Referenzgruppe: Kindergartenbesuch mit 3 Jahren oder jünger.
[3] Referenzgruppe: Einschulung im Alter von 6 Jahren oder älter.

© Waxmann Verlag GmbH

8.3.2 Beschreibung der sozialen Lage der Jugendlichen aus unterschiedlichen EGP-Klassen

Um im Folgenden für Deutschland die kulturelle, soziale und ökonomische Ausstattung der Fünfzehnjährigen unterschiedlicher sozialer Herkunft genauer beschreiben zu können, werden die Berufe der Eltern in verschiedene soziale Klassen eingeteilt. Das dafür verwendete EGP-Klassifikationssystem von Erikson und Kollegen (1979; 2002) berücksichtigt für diese Einteilung Angaben zum Beruf, zur Art des Beschäftigungsverhältnisses sowie zur Weisungsbefugnis und führt zu einer anschaulichen und soziologisch fundierten Typologie. Grundlegend für diese Typologie und die Einteilung der Berufe in Klassen ist die Annahme, dass die Lebensbedingungen in Abhängigkeit des Bildungshintergrunds, der gesellschaftlichen Anerkennung, der beruflichen Gestaltungsmöglichkeiten sowie des Einkommens variieren (Hradil, 2005). In Tabelle 8.5 sind die verwendeten sechs EGP-Klassen beispielhaft beschrieben.

Tabelle 8.5: Beispielhafte Beschreibung der EGP-Klassifikation (Baumert & Schümer, 2001, S. 339)

EGP-Klassen
Obere Dienstklasse (I)
Dazu zählen: Freie akademische Berufe, führende Angestellte, höhere Beamte, selbstständige Unternehmer mit mehr als 10 Mitarbeitern, Hochschul- und Gymnasiallehrer.
Untere Dienstklasse (II)
Dazu zählen: Angehörige von Semiprofessionen, mittleres Management, Beamte im mittleren und gehobenen Dienst, technische Angestellte mit nicht manueller Tätigkeit.
Routinedienstleistungen Handel und Verwaltung (III)
Dazu zählen: Büro- und Verwaltungsberufe mit Routinetätigkeiten, Berufe mit niedrig qualifizierten, nicht manuellen Tätigkeiten, die oftmals auch keine Berufsausbildung erfordern.
Selbstständige (IV)
Dazu zählen: Selbstständige aus manuellen Berufen mit wenigen Mitarbeitern und ohne Mitarbeiter, Freiberufler, sofern sie keinen hoch qualifizierten Beruf ausüben.
Facharbeiter und Arbeiter mit Leitungsfunktion (V, VI)
Dazu zählen: Untere technische Berufe wie Vorarbeiter, Meister, Techniker, die in manuelle Arbeitsprozesse eingebunden sind; Aufsichtskräfte im manuellen Bereich.
Un- und angelernte Arbeiter, Landarbeiter (VII)
Dazu zählen: alle un- und angelernten Berufe aus dem manuellen Bereich, Dienstleistungstätigkeiten mit manuellem Charakter und geringem Anforderungsniveau, Arbeiter in der Land-, Forst- und Fischwirtschaft.

© Waxmann Verlag GmbH

Die für die EGP-Klassifikation erforderlichen Informationen basieren auf Angaben der Schülerinnen und Schüler zur Bezugsperson.[2] Da diese Angaben nicht international erfasst wurden, sind Vergleiche mit anderen Staaten nicht möglich. Fehlende Werte wurden mit der Software Norm 2.03 (Schafer, 2000; Schafer & Graham, 2002) nach dem Multiple-Imputation-Ansatz geschätzt. Um Veränderungen über den gesamten Erhebungszeitraum der vergangenen 15 Jahre in PISA in den Blick nehmen zu können, wird in den folgenden Analysen die Lesekompetenz untersucht.

Die vorangegangen Analysen haben gezeigt, dass für die Kompetenz von Jugendlichen die Ausstattung mit ökonomischen, kulturellen und sozialen Ressourcen von zentraler Bedeutung ist. Wie unterschiedlich die Ausstattung der Schülerinnen und Schüler mit diesen Bedingungsfaktoren schulischer Leistung ist, illustrieren die Ergebnisse in Tabelle 8.6. Differenziert nach EGP-Klassen werden hier Mittelwerte der Merkmale der sozialen Lage aufgeschlüsselt, die in den Regressionsmodellen (Tabelle 8.4) enthalten sind.

Sozioökonomischer Status und Ausbildungsdauer der Bezugspersonen in der oberen Dienstklasse (I) liegen deutlich über dem Durchschnittniveau (M = 50.7 bzw. M = 14.2). Auch die Ausstattung mit häuslichen, kulturellen und lernrelevanten Besitztümern ist in der oberen Dienstklasse (I) signifikant höher als im Durchschnitt. Ähnlich verhält es sich in der Gruppe der unteren Dienstklasse (II). Mit Ausnahme der Ausstattung mit lernrelevanten Besitztümern, in denen sich diese Gruppe nicht vom Durchschnitt unterscheidet, ist auch in dieser Gruppe ein überdurchschnittliches Niveau an Ressourcen feststellbar. Ein entgegengesetztes Muster hingegen findet sich bei den Bezugspersonen, die den Routinedienstleistungen (III), den Facharbeitern (V–VI) sowie den un- und angelernten Arbeitern (VII) zugeordnet werden. Sozioökonomischer Status, Bildungsniveau sowie Ausstattung mit häuslichen und kulturellen Besitztümern liegen in diesen EGP-Klassen signifikant unter dem Durchschnitt. Die Gruppe derer, deren Eltern der Gruppe der Selbstständigen (IV) zugeordnet werden kann, verfügt über eine dem Durchschnitt entsprechende Ausstattung, was die häuslichen, kulturellen und lernrelevanten Besitztümer angeht. Auch die in Jahren gemessene Bildungsdauer liegt im mittleren Bereich. Der sozioökonomische Status hingegen liegt statistisch bedeutsam auf unterdurchschnittlichem Niveau.

Insgesamt haben rund 26 Prozent der Fünfzehnjährigen den Kindergarten erst ab einem Alter von vier Jahren oder später besucht. Aus der Klasse der un- und angelernten Arbeiter (VII) gehen überdurchschnittlich viele Kinder spät in den Kindergarten, in der oberen Dienstklasse (I) ist der Anteil rund 6 Prozent unter dem Durchschnitt und damit signifikant geringer. Hinsichtlich einer frühen Beschulung zeigen die Ergebnisse, dass rund 9 Prozent der Kinder im Alter von fünf Jahren in die Schule gekommen sind. In der oberen Dienstklasse ist dieser Anteil mit rund 12 Prozent im Vergleich zu den übrigen EGP-Klassen überdurchschnittlich hoch.

2 Angaben zur Bezugsperson rekurrieren auf den Vater bzw. wenn diese Angaben fehlen, auf die Mutter.

© Waxmann Verlag GmbH

Tabelle 8.6: Merkmale der sozialen Lage und der Bildungswege differenziert nach EGP-Klassen

	Bezugsperson		Sozioökonomischer Status		Bildungsdauer in Jahren		Häusliche Besitztümer		Kulturelle Besitztümer		Lernrelevante Besitztümer		Kindergartenbesuch mit 4. Jahren oder älter		Einschulung im Alter von 5 Jahren	
EGP-Klassen	%	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	%	*(SE)*	%	*(SE)*
Obere Dienstklasse (I)	24.0	(0.7)	**70.8**	(0.5)	**16.2**	(0.1)	**0.5**	(0.0)	**0.4**	(0.0)	**0.2**	(0.0)	**20.5**	(1.2)	**11.7**	(0.9)
Untere Dienstklasse (II)	17.0	(0.6)	**58.8**	(0.6)	**14.7**	(0.1)	**0.1**	(0.0)	**0.2**	(0.0)	0.0	(0.0)	23.2	(1.6)	9.7	(0.9)
Routinedienstleistungen (III)	7.7	(0.4)	**43.1**	(0.6)	**13.3**	(0.2)	**-0.2**	(0.1)	**-0.2**	(0.0)	-0.1	(0.1)	27.3	(2.5)	9.2	(1.4)
Selbstständige, einschl. Landwirte (IV)	10.2	(0.4)	**42.6**	(0.8)	14.3	(0.1)	0.1	(0.0)	0.0	(0.0)	0.0	(0.0)	29.7	(2.1)	7.0	(1.1)
Facharbeiter und leitende Arbeiter (V-VI)	20.6	(0.6)	**40.9**	(0.5)	**13.2**	(0.1)	**-0.3**	(0.0)	**-0.3**	(0.0)	0.0	(0.0)	26.6	(1.5)	7.0	(0.7)
Un- und angelernte Arbeiter, Landarbeiter (VII)	20.5	(0.6)	**37.0**	(0.6)	**12.7**	(0.1)	**-0.4**	(0.0)	**-0.2**	(0.0)	**-0.2**	(0.0)	**31.3**	(1.8)	7.1	(0.9)
Gesamt	100		50.7	(0.4)	14.2	(0.1)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)	26.4	(0.7)	8.6	(0.4)

Anmerkung: Signifikante Unterschiede (p <.05) in den Kennwerten für die EGP-Klassen zum Gesamtmittelwert sind fettgedruckt.

Mit insgesamt 24 Prozent ist der Anteil der Schülerinnen und Schüler mit Bezugspersonen in der oberen Dienstklasse (I) am höchsten. Die beiden zweitgrößten EGP-Klassen sind mit jeweils rund 21 Prozent die Facharbeiter und leitenden Angestellten (V–VI) und die un- und angelernten Arbeiter (VII). Der Rest verteilt sich auf die untere Dienstklasse (II) mit 17 Prozent, die Gruppe der Selbstständigen (IV) mit 10 Prozent und die Routinedienstleistungen (III) mit rund 8 Prozent. Im Vergleich zu PISA 2000 hat der Anteil der Bezugspersonen in der oberen Dienstklasse zugenommen (+4.9 Prozent), wohingegen geringere Anteile bei den Facharbeitern (-3.6 Prozent) und den Selbstständigen (-2.9 Prozent) zu verzeichnen sind und sich damit der Trend, der bereits bei PISA 2012 gefunden wurde, fortzuschreiben scheint. Die insgesamt relativ geringen Verschiebungen der Anteile in den EGP-Klassen lassen auf eine ausreichende Stabilität in der Klassifikation schließen und sprechen für Trendvergleiche zwischen den Zyklen.

© Waxmann Verlag GmbH

8.3.3 Veränderung der sozialen Disparitäten der Lesekompetenz und der Bildungsbeteiligung in den EGP-Klassen zwischen PISA 2000 und PISA 2015

Aus Tabelle 8.7 kann man entnehmen, wie sich das Leistungsniveau bei den Schülerinnen und Schülern in den verschiedenen EGP-Klassen darstellt und wie sich die Kompetenzwerte im Vergleich zum ersten PISA-Zyklus vor 15 Jahren verändert haben. Um einen Bezug zum ersten PISA-Zyklus herzustellen, wird im Folgenden von der Schwerpunktdomäne Naturwissenschaft auf die Domäne Lesen gewechselt.

Die höchsten Kompetenzwerte finden sich mit 542 Punkten und 529 Punkten bei den Jugendlichen, deren Eltern in den beiden Dienstklassen I und II tätig sind. Die Jugendlichen, deren Bezugspersonen der Klasse der Routinedienstleistungen (III) und der Facharbeiter (V, VI) zugeordnet sind, weisen mit 496 Punkten ein niedrigeres Niveau auf, niedriger auch als das der Jugendlichen, deren Bezugselternteil der Gruppe der Selbstständigen (IV) angehört (M = 501). Die niedrigsten Lesekompetenzwerte finden sich bei Schülerinnen und Schüler, deren Bezugspersonen berufliche Tätigkeiten ausüben, die in die EGP-Klasse der un- und angelernten Arbeiter fallen (M = 476).

Der Trendvergleich verdeutlicht, dass mit Ausnahme der Fünfzehnjährigen aus Elternhäusern, die den beiden Dienstklassen angehören, in allen EGP-Klassen eine deutliche Kompetenzerhöhung konstatiert werden kann. Mit kleinen bzw. mittleren Effekten sind diese Befunde im Bereich Lesen als substanziell zu betrachten. Während sich bei den Routinedienstleistungen (III) sowie den EGP-Klassen V, VI und VII damit ein Trend fortschreibt, der sich bereits im vergangenen PISA-Zyklus 2012 abzeichnete, finden sich nun auch bei den Kindern mit Bezugspersonen im selbstständigen Bereich (IV) deutliche Leistungssteigerungen. Im Mittel liegen deren Werte im Vergleich zur ersten PISA-Erhebung nun 21 Punkte höher.

Rund ein Drittel der Fünfzehnjährigen in PISA 2015 besucht das Gymnasium. Aufgeschlüsselt nach dem ersten und dem aktuellen PISA-Durchgang sind in Tabelle 8.8 die prozentualen Anteile der Gymnasiastinnen und Gymnasiasten nach EGP-Klassen wiedergegeben. Mehr als die Hälfte der Kinder mit Eltern der oberen Dienstklasse (I) besucht ein Gymnasium. Der Anteil der Gymnasiastinnen und Gymnasiasten liegt in der unteren Dienstklasse mit 43 Prozent etwas darunter und jeweils fast ein Drittel der Fünfzehnjährigen mit Eltern der EGP-Klassen III und IV geht ebenfalls ins Gymnasium. Die Anteile in diesen Gruppen haben sich gegenüber der PISA-Erhebung 2000 nur geringfügig verändert. Deutliche Unterschiede in der Bildungsbeteiligung finden sich in den EGP-Klassen V, VI und VII. Inzwischen besuchen rund 9 Prozent mehr von den Kindern, deren Bezugspersonen Tätigkeiten dieses Bereiches ausüben, ein Gymnasium. Die erhöhte Gymnasialbeteiligung – im Vergleich zum Jahr 2000 – ist demnach auf eine deutliche relative Erhöhung der Jugendlichen dieser Gruppen zurückzuführen.

© Waxmann Verlag GmbH

Tabelle 8.7: Mittelwerte und Standardabweichungen der Lesekompetenz differenziert nach EGP-Klassen (Bezugsperson) im Vergleich zwischen PISA 2000 und PISA 2015 in Deutschland

	PISA 2000			PISA 2015				
EGP-Klassen	*M*	*(SE)*	*SD*	*M*	*(SE)*	*SD*	*ΔM*	*d*
Obere Dienstklasse (I)	538	(3.4)	94	542	(3.8)	97	4	0.0
Untere Dienstklasse (II)	531	(4.0)	93	529	(4.5)	87	-2	0.0
Routinedienstleistungen Handel und Verwaltung (III)	470	(6.4)	109	496	(5.9)	88	**26**	0.3
Selbstständige (IV)	480	(5.2)	94	501	(4.7)	94	**21**	0.2
Facharbeiter und Arbeiter mit Leitungsfunktion (V, VI)	459	(4.4)	104	496	(3.9)	89	**37**	0.4
Un- und angelernte Arbeiter, Landarbeiter (VII)	432	(3.9)	111	476	(4.9)	86	**44**	0.4

Anmerkung: Signifikante Unterschiede ($p < .05$) sind fettgedruckt.

Tabelle 8.8: Prozentuale Anteile der Schülerinnen und Schüler in den Schularten, differenziert nach EGP-Klassen im Vergleich zwischen PISA 2000 und PISA 2015 in Deutschland

	PISA 2000				PISA 2015			
	Gymnasium		Andere Schularten		Gymnasium		Andere Schularten	
EGP-Klassen	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
Obere Dienstklasse (I)	51.6	(2.0)	48.4	(2.0)	55.0	(2.5)	45.0	(2.5)
Untere Dienstklasse (II)	44.8	(2.0)	55.2	(2.0)	43.1	(2.6)	56.9	(2.6)
Routinedienstleistungen Handel und Verwaltung (III)	23.7	(2.6)	76.3	(2.6)	28.7	(2.6)	71.3	(3.0)
Selbstständige (IV)	26.2	(1.7)	73.8	(1.7)	29.3	(2.8)	70.7	(2.8)
Facharbeiter und Arbeiter mit Leitungsfunktion (V, VI)	15.6	(1.2)	84.4	(1.2)	**24.4**	(2.3)	75.6	(2.3)
Un- und angelernte Arbeiter, Landarbeiter (VII)	11.0	(1.0)	89.0	(1.0)	**20.1**	(1.8)	79.9	(1.8)
Gesamt	28.3	(1.0)	71.7	(1.0)	33.4	(1.0)	66.6	(1.0)

Anmerkung: Signifikante Unterschiede ($p < .05$) zwischen PISA 2000 und PISA 2015 sind fettgedruckt.
Die Kategorie „Andere Schularten“ umfasst Hauptschulen, Realschulen, integrierte Gesamtschulen, Schulen mit mehreren Bildungsgängen, Sonder- und Förderschulen sowie berufsbildende Schulen.

© Waxmann Verlag GmbH

8.3.4 Beschreibung der sozialen Lage der leistungsstarken und leistungsschwachen Jugendlichen

Im Folgenden soll anhand der Kompetenzstufenverteilung in den EGP-Klassen ein vertiefter Blick auf die besonders leistungsstarken und die besonders leistungsschwachen Schülerinnen und Schüler erfolgen. Beide Gruppen werden analog zu Tabelle 8.6 hinsichtlich des sozioökomischen Status, des Bildungshintergrundes der Eltern sowie der Ausstattung mit lernrelevanten Besitztümern genauer beschrieben.

Der Anteil der besonders leistungsschwachen Schülerinnen und Schüler ist in EGP-Klasse VII am höchsten (vgl. Abbildung 8.3). Rund ein Viertel der Fünfzehnjährigen mit Bezugspersonen dieser Klasse weist ein Kompetenzniveau unter Stufe II auf. Der Anteil an Schülerinnen und Schülern dieses Kompetenzniveaus liegt in den beiden Dienstklassen (I und II) bei jeweils nur bei ca. 10 Prozent. Die Anteile der leistungsschwachen Jugendlichen mit Bezugspersonen der übrigen EGP-Klassen liegen mit rund 18 Prozent (EGP-Klasse V–VI), 17 Prozent (EGP-Klasse IV) und 19 Prozent (EGP-Klasse III) dazwischen.

Bei den besonders leistungsstarken Fünfzehnjährigen zeigt sich eine spiegelbildliche Verteilung. Über ein sehr hohes Kompetenzniveau im Lesen auf Kompetenzstufe V oder höher verfügen 19 Prozent der Jugendlichen mit Eltern der oberen Dienstklasse (I) und 15 Prozent der unteren Dienstklasse (II). Während der Anteil der Jugendlichen aus Elternhäusern der EGP-Klassen III, IV und V–VI bei jeweils rund 9 Prozent liegt, finden sich bei jenen der EGP-Klasse VII nur 6 Prozent besonders leistungsstarke Jugendliche.

Die Gruppe der besonders leistungsstarken Schülerinnen und Schüler unterscheidet sich hinsichtlich des sozioökonomischen Status, des Bildungshintergrundes der Eltern und der Ausstattung mit häuslichen, kulturellen und lernrelevanten Besitztümern signifikant von den Schülerinnen und Schülern, deren Kompetenzniveau unter Stufe II liegt (vgl. Tabelle 8.9). Die Jugendlichen mit sehr hoher Lesekompetenz weisen in allen Merk-

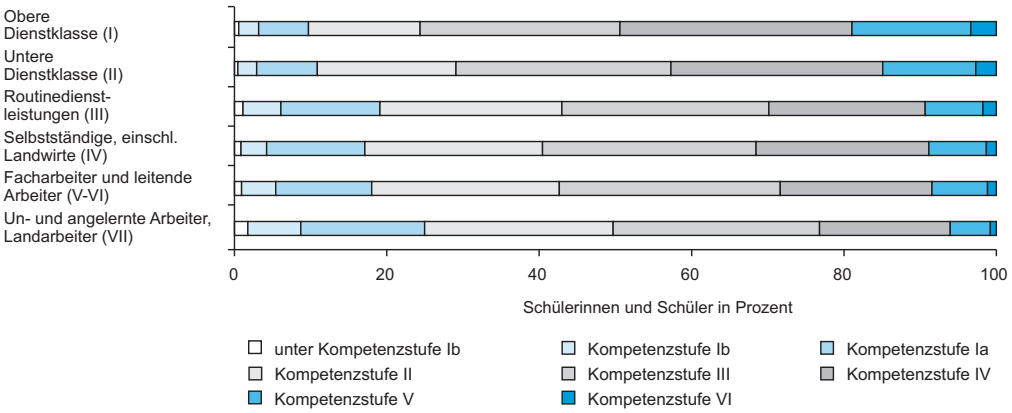


Abbildung 8.3: Verteilung der Kompetenzstufen in Lesen differenziert nach EGP-Klassen

© Waxmann Verlag GmbH

Tabelle 8.9: Merkmale der sozialen Lage differenziert nach EGP-Klassen in der Gruppe der im Lesen leistungstarken (mindestens auf Kompetenzstufe V) und leistungsschwachen (unter Kompetenzstufe II) Schülerinnen und Schüler

	Lesekompetenz mindestens auf Kompetenzstufe V									
	Sozio-ökon. Status		Bildungs-dauer in Jahren		Häusliche Besitz-tümer		Kulturelle Besitz-tümer		Lern-relevante Besitz-tümer	
EGP-Klassen	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
Obere Dienstklasse (I)	**75**	(0.7)	**17**	(0.1)	**0.8**	(0.1)	**0.7**	(0.1)	0.4	(0.1)
Untere Dienstklasse (II)	64	(1.3)	16	(0.2)	0.5	(0.1)	0.5	(0.1)	0.2	(0.1)
Routinedienstleistungen (III)	**49**	(2.7)	**14**	(0.5)	**0.3**	(0.1)	0.2	(0.2)	0.2	(0.2)
Selbstständige, einschl. Landwirte (IV)	**52**	(2.6)	15	(0.4)	0.6	(0.1)	0.3	(0.2)	0.2	(0.2)
Facharbeiter und leitende Arbeiter (V–VI)	**46**	(1.6)	**14**	(0.3)	**0.2**	(0.1)	**0.0**	(0.1)	0.3	(0.1)
Un- und angelernte Arbeiter, Landarbeiter (VII)	**43**	(2.6)	**13**	(0.4)	**0.1**	(0.1)	0.1	(0.2)	0.1	(0.1)
Gesamt	62	(0.9)	16	(0.1)	0.5	(0.0)	0.4	(0.1)	0.3	(0.0)

	Lesekompetenz unter Kompetenzstufe II									
	Sozio-ökon. Status		Bildungs-dauer in Jahren		Häusliche Besitz-tümer		Kulturelle Besitz-tümer		Lern-relevante Besitz-tümer	
EGP-Klassen	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
Obere Dienstklasse (I)	**64**	(1.8)	**15**	(0.3)	**0.0**	(0.2)	**0.1**	(0.1)	**-0.1**	(0.1)
Untere Dienstklasse (II)	**52**	(2.1)	14	(0.5)	-0.4	(0.1)	-0.2	(0.1)	-0.4	(0.2)
Routinedienstleistungen (III)	39	(2.2)	13	(0.5)	-0.7	(0.2)	-0.4	(0.1)	-0.5	(0.2)
Selbstständige, einschl. Landwirte (IV)	**35**	(1.8)	14	(0.4)	-0.3	(0.1)	-0.3	(0.1)	-0.4	(0.1)
Facharbeiter und leitende Arbeiter (V–VI)	**36**	(1.4)	13	(0.3)	-0.6	(0.1)	-0.4	(0.1)	-0.4	(0.1)
Un- und angelernte Arbeiter, Landarbeiter (VII)	**33**	(1.3)	**12**	(0.3)	**-0.8**	(0.1)	-0.4	(0.1)	-0.6	(0.1)
Gesamt	41	(0.8)	13	(0.2)	-0.5	(0.0)	-0.3	(0.0)	-0.4	(0.0)

Anmerkung: Signifikante Unterschiede (p <.05) in den Kennwerten für die EGP-Klassen zum Gruppenmittelwert sind fettgedruckt.

malen sowohl im Gesamtdurchschnitt als auch innerhalb der EGP-Klassen deutlich höhere Werte auf als die Jugendlichen mit niedrigem Kompetenzniveau.

Innerhalb der beiden Gruppen zeigt sich, dass es bei den Jugendlichen mit hoher Lesekompetenz zwischen den EGP-Klassen deutliche Unterschiede im sozioökonomischen Status, der Bildungsdauer und der Ausstattung mit häuslichen Besitztümern gibt. Bei den leistungsstarken Schülerinnen und Schülern mit Bezugspersonen, die den Routinedienstleistungen (III), den Facharbeitern (V–VI) und un- und angelernten Arbeitern (VII) zugeordnet werden können, sind in diesen Merkmalen deutlich unterdurchschnittliche Werte zufinden, wohingegen sich bei jenen mit Eltern in der oberen Dienstklasse (I) Werte über dem Durchschnitt zeigen. Bei den Schülerinnen und Schülern, die

© Waxmann Verlag GmbH

schwache Leistungen im Lesen aufweisen, finden sich primär im sozioökonomischen Status signifikante Unterschiede zwischen den EGP-Klassen. Mit Ausnahme der Jugendlichen, deren Eltern im Bereich der oberen Dienstklasse tätig sind, finden sich kaum vom Gruppendurchschnitt abweichende Mittelwerte zwischen den EGP-Klassen.

Daran wird deutlich, dass sich die Gruppe der leistungsschwachen Schülerinnen und Schüler hinsichtlich dieser Merkmale des sozialen Hintergrundes insgesamt deutlich homogener darstellt als die Gruppe der Jugendlichen, die im Lesen besonders gute Leistungen zeigen. Bei den Jugendlichen, deren Eltern im Bereich der oberen Dienstklasse tätig sind, finden sich in beiden Leistungsextremgruppen überdurchschnittlich hohe Ausprägung in den Merkmalen der sozialen Lage.

8.4 Zusammenfassung und Diskussion

Ziel dieses Kapitels war es, die sozialen Disparitäten des Kompetenzerwerbs und der Bildungsbeteiligung für die Fünfzehnjährigen in Deutschland zu beschreiben und sie im Vergleich zwischen den OECD-Staaten sowie im Vergleich mit den früheren PISA-Erhebungen zu interpretieren. Mithilfe von zusätzlichen Fragebogendaten zur sozialen Herkunft konnten zudem die Ungleichheiten in Bezug auf Bildungsbeteiligung und Kompetenzen der Jugendlichen für Deutschland differenzierter betrachtet werden. Die vorliegenden Analysen stützen sich im internationalen Teil primär auf zwei Indikatoren: Zum einen wird der elterliche sozioökonomische Status (*Highest Socio-Economic Status*, HISEI) verwendet, zum anderen kommt mit dem *Economic, Social and Cultural Status* (ESCS) ein im Vergleich zum HISEI breiter gefasster Index zum Einsatz. Bei der Konstruktion des ESCS werden neben dem HISEI zusätzliche Indikatoren berücksichtigt. Während der HISEI ausschließlich den sozioökonomischen Status des Elternhauses abbildet, werden beim ESCS zudem der elterliche Bildungsabschluss und der Besitz von Kultur- und Wohlstandsgütern und damit soziale und kulturelle Komponenten der sozialen Herkunft integriert.

Zunächst zeigte sich, dass der anhand des HISEI gemessene sozioökonomische Status zwischen den Staaten beträchtlich variiert. Die Schülerinnen und Schüler in Deutschland befinden sich wie bereits bei den vorangegangen PISA-Erhebungen im Bereich des OECD-Durchschnitts. Ferner lässt sich in PISA 2015, wie auch schon in früheren PISA-Erhebungen, in allen OECD-Staaten ein Zusammenhang zwischen dem sozioökonomischen Status (HISEI) und dem ökonomischen, sozialen und kulturellen Status (ESCS) des Elternhauses und den erreichten naturwissenschaftlichen Kompetenzen der Jugendlichen nachweisen. In keinem Staat ist dieser Zusammenhang völlig entkoppelt. Dennoch gibt es eine hohe Variabilität zwischen den Staaten. Um den Zusammenhang genauer zu erfassen, kommen in PISA zwei Kennwerte zum Einsatz, die Steigung des sozialen Gradienten und das Maß der aufgeklärten Varianz. Die Steigung des sozialen Gradienten zeigt an, wie umfänglich sich die Kompetenz verändern würde, wenn sich der Prädiktor um eine Standardabweichung erhöhte. Damit gibt der soziale Gradient

© Waxmann Verlag GmbH

auch Hinweise dazu, wie groß die Wirkung bildungspolitischer Maßnahmen möglicherweise sein könnte, die auf die Verringerung sozialer Disparitäten zielen. Bei einem größeren sozialen Gradienten kommt es zu größeren Kompetenzunterschieden im Zusammenhang mit den erworbenen Kompetenzen. Sozial benachteiligte sowie begünstigte Jugendliche zeigen demnach deutliche Kompetenzunterschiede, die durch gezielte bildungspolitische Maßnahmen verringert werden könnten. Die Varianzaufklärung gibt im Unterschied dazu an, zu welchem Anteil sich die Unterschiede in der Kompetenz durch das Merkmal der sozialen Herkunft vorhersagen lassen. Im Vergleich zum sozialen Gradienten spiegelt die Varianzaufklärung also die Präzision wider, mit der sich die Kompetenzunterschiede durch das Merkmal der sozialen Herkunft erklären lassen.

Eine geringe Kopplung bei gleichzeitig hohem Durchschnittsniveau in der naturwissenschaftlichen Kompetenz findet man etwa in Estland, Dänemark und Kanada. Auch in Japan und Finnland zeigen sich überdurchschnittliche Kompetenzwerte und liegt der soziale Gradient nicht signifikant über dem OECD-Mittelwert. Ein enger Zusammenhang zwischen Herkunft und Kompetenzniveau bei gleichzeitig niedrigen oder durchschnittlichen Kompetenzmittelwerten ergab sich etwa in Ungarn, Luxemburg, Frankreich oder Österreich.

Für die Fünfzehnjährigen in Deutschland liegt sowohl die Steigung des sozialen Gradienten als auch die Varianzaufklärung in PISA 2015 gemessen am sozioökonomischen Status (HISEI) deutlich über dem OECD-Durchschnitt. Wird hingegen statt des sozioökonomischen Status ein breiteres Indikatorensystem verwendet und der Zusammenhang am ESCS gemessen, sind in Deutschland beim sozialen Gradienten keine signifikanten Unterschiede beim Vergleich mit dem OECD-Durchschnitt feststellbar. Die Varianzaufklärung dagegen liegt deutlich über dem OECD-Mittelwert.

Darüber hinaus zeigt sich anhand des reskalierten Trend-ESCS-Indexes zwischen PISA 2006 und PISA 2015 eine teilweise Entkopplung zwischen naturwissenschaftlicher Kompetenz und sozialer Herkunft. Werden nicht nur sozioökonomische, sondern gleichzeitig ökonomische, soziale und kulturelle Ressourcen im Elternhaus in die Analysen einbezogen, so nimmt der Zusammenhang von sozialer Herkunft und naturwissenschaftlicher Kompetenz gemessen am Anteil der aufgeklärten Varianz über die Zeit ab. Offenbar gelingt es in Deutschland inzwischen zumindest in Ansätzen, die neben dem sozioökonomischen Status vorhandenen kulturellen und sozialen Ressourcen der sozialen Herkunft besser zu kompensieren.

Dies bestätigt sich auch in den vertieften Analysen für Deutschland, in denen festgestellt werden konnte, dass sich die Abstände in der Lesekompetenz zwischen den sozialen Lagen über die Zeit reduziert haben. Zwar unterscheidet sich auch in PISA 2015 die mittlere Lesekompetenz von Jugendlichen, deren Eltern der oberen Dienstklasse, von solchen, deren Eltern der Klasse der un- und angelernten Arbeiter zugeordnet werden können, noch deutlich. Die Differenz der Mittelwerte ist jedoch von 106 Punkten in PISA 2000 auf 66 Punkte in PISA 2015 gesunken. Profitiert haben vor allem die Schülerinnen und Schüler, die aus Arbeiterfamilien stammen oder zu den EGP-Klassen der Routinedienstleistungen und der Selbstständigen zählen. Die Lesekompetenz

© Waxmann Verlag GmbH

der Jugendlichen von Familien aus der oberen Dienstklasse ist hingegen unverändert auf hohem Niveau geblieben. Im Hinblick auf die sozialen Disparitäten der Bildungsbeteiligung belegen die Analysen einerseits weiterhin große Unterschiede zwischen den EGP-Klassen im Gymnasialbesuch. Andererseits wurde auch hier ein tendenzieller Rückgang der sozialen Disparitäten festgestellt. Vergleicht man die Daten von PISA 2000 und PISA 2015, erhöhte sich die Gymnasialbeteiligung vor allem bei Schülerinnen und Schülern, deren Eltern den EGP-Klassen der Facharbeiter (16 vs. 24 Prozent) und un- und angelernten Arbeiter (11 vs. 20 Prozent) zuzuordnen sind. Bei den Schülerinnen und Schülern mit Eltern der oberen Dienstklasse blieb die Gymnasialbeteiligung weiterhin auf hohem Niveau bestehen. Insgesamt liegt die Gymnasialbeteiligung mit 33 Prozent etwas unterhalb des Wertes von 36 Prozent in PISA 2012. Dies könnte ein Indiz dafür sein, dass zunehmend mehr nicht gymnasiale Schulformen auch einen Weg zum Abitur anbieten (beispielsweise an integrierten Gesamtschulen). Inwieweit dies tatsächlich ein Trend ist, der sich langfristig bestätigt, wird sich in den kommenden PISA-Zyklen herausstellen.

Detailanalysen zu den besonders leseschwachen und besonders lesestarken Schülerinnen und Schülern haben gezeigt, dass diese in allen EGP-Klassen vertreten sind, allerdings nicht gleich häufig. Erwartungskonform sind die besonders lesestarken Schülerinnen und Schülern in den oberen beiden EGP-Klassen vertreten und die besonders leseschwachen Jugendlichen überproportional häufig in den Familien von un- und angelernten Arbeitern. Der Zusammenhang ist jedoch nicht streng deterministisch, und in den Familien von un- und angelernten Arbeitern finden sich ebenso lesestarke Kinder wie in den Familien der unteren und oberen Dienstklassen leseschwache Schülerinnen und Schüler. Dies kann als Hinweis darauf gewertet werden, dass für den Kompetenzerwerb vor allen die lernrelevanten Prozesse im Elternhaus eine Rolle spielen.

Betrachtet man die sozialen Disparitäten hinsichtlich der Kompetenzen von Fünfzehnjährigen für Deutschland insgesamt, so muss beachtet werden, dass sich die Befunde aus dem internationalen Staatenvergleich nicht in allen Ländern innerhalb von Deutschland gleichermaßen widerspiegeln. Wie die Ergebnisse des Ländervergleichs der Bildungsstandards belegt haben, gibt es auch in den deutschen Bundesländern eine erhebliche Varianz in Bezug auf die Kennwerte sozialer Disparitäten (Knigge & Köller, 2010; Knigge & Leucht, 2010; Kuhl et al., 2016; Kuhl et al., 2013; Richter et al., 2012).

Welche bildungspolitischen Maßnahmen können nun dazu beitragen, soziale Disparitäten des Kompetenzerwerbs und der Bildungsbeteiligung in Deutschland (weiter) zu reduzieren? Nach wie vor gilt für Deutschland, dass konkrete auf den Kompetenzerwerb ausgerichtete Förderangebote darauf abzielen sollten, alle Schülerinnen und Schüler beim Lernen voranzubringen. Eine Fokussierung auf leistungsschwache Jugendliche ist dabei wichtig, um sicherzustellen, dass alle Schülerinnen und Schüler am Ende der Schulzeit über jenes Kompetenzniveau verfügen, das für einen erfolgreichen Übergang ins Berufsleben erforderlich ist. Förderangebote sollte es aber auch für leistungsstarke Jugendliche geben. Insbesondere der internationale Vergleich zeigt, dass die Gruppe der sehr leistungsstarken Schülerinnen und Schüler in Deutschland kleiner ist als in ande-

© Waxmann Verlag GmbH

ren Staaten. Parallele Förderangebote im unteren und oberen Leistungsbereich können über die Zeit dazu führen, dass sich Kopplungsmaße zwischen sozialer Herkunft und erreichtem Kompetenzniveau gleichbleibend entwickeln. Solange dabei die Gruppe der leistungsschwachen Schülerinnen und Schüler kontinuierlich abnimmt, wäre diese Entwicklung aus bildungspolitischer Perspektive als positiv zu bewerten.

Literatur

Baumert, J. & Maaz, K. (2010). Bildungsungleichheit und Bildungsarmut – Der Beitrag von Large-Scale-Assessments. In G. Quenzel & K. Hurrelmann (Hrsg.), *Bildungsverlierer. Neue Ungleichheiten* (S. 159–179). Wiesbaden: VS-Verlag für Sozialwissenschaft.

Baumert, J., Maaz, K. & Trautwein, U. (Hrsg.) (2009). Bildungsentscheidungen. *Zeitschrift für Erziehungswissenschaft*, Sonderheft Nr. 12. Wiesbaden: VS Verlag für Sozialwissenschaften.

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.

Bourdieu, P. (1982). *Die feinen Unterschiede: Kritik der gesellschaftlichen Urteilskraft*. Frankfurt a.M.: Suhrkamp.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (S. 183–198). Göttingen: Schwartz.

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.

Coleman, J. S. (1988). Social capital and the creation of human capital. *American Journal of Sociology*, *94*, 95–120.

Coleman, J. S. (1996). Der Verlust sozialen Kapitals und seine Auswirkungen auf die Schule. In A. Leschinsky (Hrsg.), *Die Institutionalisierung von Lehren und Lernen* (S. 99–105). Weinheim: Beltz.

Dumont, H., Maaz, K., Neumann, M. & Becker, M. (2014). Soziale Ungleichheiten beim Übergang von der Grundschule in die Sekundarstufe I: Theorie, Forschungsstand, Interventions- und Fördermöglichkeiten. *Zeitschrift für Erziehungswissenschaft*, Sonderheft 24, 141–165.

Ehmke, T. & Baumert, J. (2007). Soziale Herkunft und Kompetenzerwerb. Vergleiche zwischen PISA 2000, 2003 und 2006. 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. 309–335). Münster: Waxmann.

Ehmke, T. & Baumert, J. (2008). Soziale Disparitäten des Kompetenzerwerbs und der Bildungsbeteiligung in den Ländern: Vergleiche zwischen PISA 2000 und 2006. In M. Prenzel, C. Artelt, J. Baumert, W. Blum, M. Hammann, E. Klieme & R. Pekrun (Hrsg.), *PISA 2006 in Deutschland. Die Kompetenzen der Jugendlichen im dritten Ländervergleich* (S. 319–342). Münster: Waxmann.

Ehmke, T., Hohensee, F., Heidemeier, H. & Prenzel, M. (2004). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, H.-G. Rolff, J. Rost & U. Schiefele (Hrsg.),

© Waxmann Verlag GmbH

PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs (S. 225–254). Münster: Waxmann.

Ehmke, T. & Jude, N. (2010). Soziale Herkunft und Kompetenzerwerb. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 231–254). Münster: Waxmann

Ehmke, T. & Siegle, T. (2005). ISEI, ISCED, HOMEPOS, ESCS. *Zeitschrift für Erziehungswissenschaft, 8* (4), 521–540.

Ehmke, T., Siegle, T. & Hohensee, F. (2005). Soziale Herkunft im Ländervergleich. In M. Prenzel, J. Baumert, W. Blum, R. Lehmann, D. Leutner, M. Neubrand, R. Pekrun, J. Rost & U. Schiefele (Hrsg.), *PISA 2003. Der zweite Vergleich der Länder in Deutschland – Was wissen und können Jugendliche?* (S. 235–268). Münster: Waxmann.

Erikson, R. & Goldthorpe, J. H. (2002). Intergenerational inequality: A sociological perspective. *Journal of Economic Perspectives, 16* (3), 31–44.

Erikson, R., Goldthorpe, J. H. & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology, 30*, 341–415.

Faust, G. & Roßbach, H.-G. (2014). Herkunft und Bildungserfolg von der frühen Kindheit bis ins Erwachsenenalter: Forschungsstand und Interventionsmöglichkeiten aus interdisziplinärer Perspektive. *Zeitschrift für Erziehungswissenschaft*, Sonderheft 24, 119–140.

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), 1–56.

Ganzeboom, H. B. G., de Graaf, P. M., Treiman, D. J. & de Leeuw, J. (1992). *A standard international socioeconomic index of occupational status* (WORC Reprint). Tilburg: WORC, Work and Organization Research Centre.

Ganzeboom, H. B. G. & Treiman, D. J. (1996). Internationally comparable measures of occupational Status for the 1988 International Standard Classification of Occupations. *Social Science Research 25*, 201–239.

Ganzeboom, H. B. G & Treiman, D. (2012). *International stratification and mobility file: Conversion tools.* Amsterdam: Department of Social Research Methodology. Zugriff am 01.10.2016. Verfügbar unter http://www.harryganzeboom.nl/ismf/index.htm.

Ganzeboom, H. B. G. & Treiman, D. J. (2003). Three internationally standardised measures for comparative research on occupational status. In J. H. P. Hoffmeyer-Zlotnik & C. Wolf (Hrsg.), *Advances in cross-national comparison* (S. 159–193). Berlin: Springer.

Granato, M. & Ulrich, J. G. (2014). Soziale Ungleichheit beim Zugang in eine Berufsausbildung: Welche Bedeutung haben die Institutionen? *Zeitschrift für Erziehungswissenschaft*, Sonderheft 24, 205–232.

Hasselhorn, M. & Kuger, S. (2014). Wirksamkeit schulrelevanter Förderung in Kindertagesstätten. *Zeitschrift für Erziehungswissenschaft*, Sonderheft 24, 299–314.

Hradil, S. (2005). *Soziale Ungleichheit in Deutschland* (8. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.

ILO (1969) = International Labour Office. (1969). *International Standard Classification of Occupations* (2. überarb. Aufl.). Genf: ILO.

ILO (1990) = International Labour Office. (1990). *International Standard Classification of Occupations* ISCO-88. Genf: ILO.

ILO (2012) = International Labour Office. (2012). ISCO-08 *International Standard Classification of Occupations. Structure, group definitions and correspondence tables*. Vol I. ILO: Genf.

© Waxmann Verlag GmbH

Knigge, M. & Köller, O. (2010). Effekte der sozialen Zusammensetzung der Schülerschaft. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 227–244). Münster: Waxmann.

Knigge, M. & Leucht, M. (2010). Soziale Disparitäten im Spracherwerb. In O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 185–201). Münster: Waxmann.

Kramer, R.-T. (2011). Abs*chied von Bourdieu? Perspektiven ungleichheitsbezogener Bildungsforschung*. Wiesbaden: VS-Verlag für Sozialwissenschaften.

Kuhl, P., Haag, N., Federlein, F., Weirich, S. & Schipolowksi, S. (2016). Soziale Disparitäten. In P. Stanat, K. Böhme, S. Schipolowski & N. Haag (Hrsg.), *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich* (S. 409–430). Münster: Waxmann.

Kuhl, P., Siegle, T. & Lenski, A. E. (2013). Soziale Disparitäten. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 175–296). Münster: Waxmann.

Maaz, K., Baumert, J. & Cortina, K. S. (2008). Soziale und regionale Ungleichheit im deutschen Bildungssystem. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), Das *Bildungswesen in der Bundesrepublik Deutschland: Strukturen und Entwicklungen im Überblick* (S. 205–243). Reinbek bei Hamburg: Rowohlt.

Maaz, K., Baumert, J. & Trautwein, U. (2009). Genese sozialer Ungleichheit im institutionellen Kontext der Schule: Wo entsteht und vergrößert sich soziale Ungleichheit? *Zeitschrift für Erziehungswissenschaft*, Sonderheft 12, 11–46.

Maaz, K., Hausen, C., McElvany, N. & Baumert, J. (2006). Stichwort: Übergänge im Bildungssystem. Theoretische Konzepte und ihre Anwendung in der empirischen Forschung beim Übergang in die Sekundarstufe. *Zeitschrift für Erziehungswissenschaft, 9* (3), 299–327.

Maaz, K., Neumann, M. & Baumert, J. (Hrsg.). (2014). Herkunft und Bildungserfolg von der frühen Kindheit bis ins Erwachsenenalter: Forschungsstand und Interventionsmöglichkeiten aus interdisziplinärer Perspektive. *Zeitschrift für Erziehungswissenschaft*, Sonderheft 24. Wiesbaden: Springer VS für Sozialwissenschaften.

Müller, K. & Ehmke, T. (2013). Soziale Herkunft als Bedingung der Kompetenzentwicklung. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (S. 245–275). Münster: Waxmann

OECD. (2001). *Knowledge and skills for life. First results from PISA 2000*. Paris: OECD.

OECD. (2004). *Learning for tomorrow's world. First results from PISA 2003*. Paris: OECD.

OECD. (2007). *PISA 2006: Science competencies for tomorrow's world*. Paris: OECD.

OECD. (2010). *PISA 2009. Was Schülerinnen und Schüler wissen und können. Schülerleistungen in Lesekompetenz, Mathematik und Naturwissenschaften*. Paris: OECD.

OECD. (2013), *PISA 2012 results: Excellence through equity: giving every student the chance to succeed* (Volume II). Paris: OECD.

OECD. (2016). *PISA 2015 results: Excellence and equity in education* (Volume I). Paris: OECD.

OECD. (in Vorbereitung). *PISA 2015 technical report*. Paris: OECD.

Paetsch, J., Wolf, K. M., Stanat, P. & Darsow, A. (2014). Sprachförderung von Kindern und Jugendlichen aus Zuwandererfamilien. *Zeitschrift für Erziehungswissenschaft*, Sonderheft 24, 315–349.

© Waxmann Verlag GmbH

Richter, D., Kuhl, P. & Pant, H. A. (2012). Soziale Disparitäten. In P. Stanat, H. A. Pant, K. Böhme, D. Richter (Hrsg.). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011* (S. 191–208). Münster: Waxmann.

Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7* (2), 147–177.

Schafer, J. L. (2000). *Norm for Windows 95/98/NT* (Version 2.03) – Pennsylvania.

Schimpl-Neimanns, B., (2004). Zur Umsetzung des Internationalen Sozioökonomischen Index des beruflichen Status (ISEI). *ZUMA-Nachrichten*, *54*, 154–170

Seyda, S. (2009), Kindergartenbesuch und späterer Bildungserfolg. Eine bildungsökonomische Analyse anhand des Sozio-ökonomischen Panels. *Zeitschrift für Erziehungswissenschaft, 12* (2), 233–251.

Strietholt, R., Manitius, V., Berkemeyer, N. & Bos, W. (2015). Bildung und Bildungsungleichheit an Halb- und Ganztagsschulen. *Zeitschrift für Erziehungswissenschaft, 4* (18), 737–716.

Watermann, R., Daniel, A. & Maaz, K. (2014). Primäre und sekundäre Disparitäten des Hochschulzugangs: Erklärungsmodelle, Datengrundlagen und Entwicklungen. *Zeitschrift für Erziehungswissenschaft,* Sonderheft 24, 233–261.

Züchner, I. & Fischer, N. (2014). Kompensatorische Wirkungen von Ganztagsschulen – Ist die Ganztagsschule ein Instrument zur Entkopplung des Zusammenhangs von sozialer Herkunft und Bildungserfolg? *Zeitschrift für Erziehungswissenschaft*, Sonderheft 24, 349–368.

© Waxmann Verlag GmbH

9 Naturwissenschaftliche Kompetenz von Schülerinnen und Schülern mit Zuwanderungshintergrund

Dominique Rauch, Julia Mang, Hendrik Härtig & Nicole Haag

Der Anteil Jugendlicher mit Zuwanderungshintergrund in Deutschland ist seit dem Jahr 2006 gestiegen und beträgt mittlerweile etwa 28 Prozent. Gleichzeitig hat die Heterogenität innerhalb der Jugendlichen mit Zuwanderungshintergrund in Bezug auf ihre Herkunftsländer zugenommen. Nach wie vor verfügen Familien von Jugendlichen mit Zuwanderungshintergrund in Deutschland über geringere sozioökonomische und kulturelle Ressourcen. Ebenso wie in den meisten europäischen Nachbarstaaten erreichen Jugendliche mit Zuwanderungshintergrund in Deutschland im Jahr 2015 geringere mittlere naturwissenschaftliche Kompetenzen, wobei für Jugendliche der ersten Generation besonders hohe Disparitäten bestehen. Im Gegensatz zu den Bereichen Lesen in PISA 2009 und Mathematik in PISA 2012 konnte für die Naturwissenschaften keine Reduktion zuwanderungsbezogener Disparitäten seit dem Bezugsjahr 2006 – in dem sie ebenfalls Hauptdomäne waren – festgestellt werden. Die angemessene Unterstützung des naturwissenschaftlichen Kompetenzerwerbs von Jugendlichen aus zugewanderten Familien stellt demnach weiterhin eine Herausforderung für das Bildungssystem dar.

9.1 Einleitung

Während der Mikrozensus erst ab 2005 das Geburtsland zur Definition des Zuwanderungsstatus heranzog (Statistisches Bundesamt, 2007), wurde im Rahmen von PISA bereits in der ersten Erhebung im Jahr 2000 das Geburtsland der Jugendlichen und ihrer Eltern als Indikator verwendet. Dabei zeigte sich, dass der Anteil von Fünfzehnjährigen mit mindestens einem im Ausland geborenen Elternteil bei rund 22 Prozent lag und damit deutlich höher war als der von der amtlichen Statistik für das Jahr 2000 ermittelte Anteil ausländischer Schülerinnen und Schüler, der dort etwa 8 Prozent betrug (Baumert & Schümer, 2001; Statistisches Bundesamt, 2001). Seitdem ist der Anteil der Jugendlichen mit Zuwanderungshintergrund an der Gesamtpopulation fünfzehnjähriger Schülerinnen und Schüler in jeder PISA-Erhebung gestiegen, wobei in den Erhebungen der Jahre 2006, 2009 und 2012 der Anteil selbst zugewanderter Jugendlicher zurückging. In der nationalen PISA-Berichterstattung akzentuiert seit dem Jahr 2012

© Waxmann Verlag GmbH

der Begriff *Zuwanderung* (analog zu *immigration* im Englischen) stärker den Aspekt des Zugewandertseins, im Sinne von angekommen sein. Mit der erhöhten Aufmerksamkeit, die durch PISA der Differenzkategorie „Zuwanderung ja/nein" zuteilgeworden ist, ist ein „Differenzdilemma" (Kiesel, 1996) verbunden: Einerseits ist es sinnvoll und notwendig, Ungleichheiten zwischen zugewanderten und nicht zugewanderten Jugendlichen in Schulsystemen zu erkennen. Andererseits besteht durch die vielseitig rezipierte Darstellung von Unterschieden zwischen Jugendlichen ohne und mit Zuwanderungshintergrund die Gefahr der Stigmatisierung. Dieses Dilemma kann nicht aufgelöst werden, sollte jedoch reflektiert (vgl. hierzu Stošić, 2016) und mitbedacht werden.

Trotz oder gerade wegen der Aktualität der verstärkten Ankunft geflohener Menschen aus vielen Teilen der Welt im Jahr 2015 konnten für den vorliegenden Bericht keine Daten von Jugendlichen aus dieser Gruppe mit einbezogen werden. An den PISA-Erhebungen darf sinnvollerweise nur teilnehmen, wer zum Zeitpunkt der Erhebung bereits mindestens ein Jahr eine Schule in Deutschland besucht hat und über die Minimalanforderungen (das heißt die Fähigkeit, laut Schulkoordinatorin oder Schulkoordinator selbstständig und ohne fremde Hilfe am PISA-Test teilnehmen zu können) an schriftsprachlichen Deutschkenntnissen verfügt. Bereits vor dem aktuellen Anstieg der Zuwanderung – zum Zeitpunkt der Erhebung zu PISA 2012 – hatten rund ein Viertel aller Fünfzehnjährigen an allgemeinbildenden Schulen einen Zuwanderungshintergrund. Diese Zahlen verdeutlichen, dass kulturelle und sprachliche Heterogenität in deutschen Schulen durchaus die Regel war und ist (vgl. auch Gogolin, 1994).

Für die Interpretation zuwanderungsbezogener Disparitäten ist relevant, ob Schülerinnen und Schüler selbst zugewandert sind oder anders als ihre zugewanderten Eltern bereits im Inland geboren wurden. Familien, die ihr Herkunftsland verlassen haben, müssen sich im Kontext des Landes neu orientieren, in welches sie zugewandert sind. Hierzu müssen sie häufig zunächst die Verkehrssprache des Landes erlernen und sich in Institutionen zurechtfinden, die anders strukturiert sind als in ihrem Herkunftsland. Auch für das Schulsystem stellen neu zugewanderte Kinder und Jugendliche (die sogenannte erste Generation) eine andere Herausforderung dar als solche, deren Eltern zugewandert sind, die selbst aber bereits in Deutschland geboren wurden (die sogenannte zweite Generation). Aufgrund zusätzlicher individueller, institutioneller und systemseitiger Herausforderungen ist zu erwarten, dass der Bildungserfolg von Personen, die selbst zugewandert sind, im Durchschnitt geringer sein wird als der von Personen ohne eigene Zuwanderungserfahrungen. Unter Berücksichtigung der Zusammensetzung der Gruppe von Schülerinnen und Schülern mit eigener Zuwanderungserfahrung, etwa im Hinblick auf den Bildungshintergrund der Eltern, können Unterschiede in den Disparitäten zwischen Bildungssystemen als Hinweis darauf gewertet werden, dass die Integration zugewanderter Schülerinnen und Schüler durch das Schulsystem unterschiedlich gut gelingt. Inwieweit das der Fall ist, wird im Rahmen der PISA-Studien auch am Bildungserfolg von Kindern der zweiten Generation gemessen. Zumindest bei vergleichbaren Lebens- und Lernumwelten sollten diese Schülerinnen und Schüler ähnliche Kompetenzen erwerben wie Schülerinnen und Schüler ohne Zuwanderungshintergrund. Bestehen

© Waxmann Verlag GmbH

auch für diese Gruppe von Schülerinnen und Schülern substanzielle zuwanderungsbezogene Disparitäten, sollte dies Anlass sein, über Konsequenzen auf allen Ebenen des Bildungssystems nachzudenken in Bezug auf die Gestaltung von individueller Förderung, Unterricht und schulischen und außerschulischen Rahmenbedingungen.

Die Analyse zuwanderungsbezogener Disparitäten für die erste und zweite Generation vermittelt einen Eindruck davon, inwieweit es im Bildungssystem insgesamt gelingt, den Kompetenzerwerb aller Schülerinnen und Schüler zu fördern. Dabei werden im Rahmen der PISA-Studien die Daten auch differenziert für verschiedene Herkunftsländer der zugewanderten Familien berichtet. So kann genauer bestimmt werden, wie gut die Kompetenzförderung im Schulsystem für Schülerinnen und Schüler unterschiedlicher Herkunftsgruppen gelingt. Betrachtet werden diejenigen Herkunftsländer, aus denen die meisten fünfzehnjährigen Schülerinnen und Schüler in Deutschland beziehungsweise deren Eltern stammen.

PISA hat in verschiedener Hinsicht dazu beigetragen, den Blick auf die Integrationskraft des deutschen Bildungssystems für Schülerinnen und Schüler mit Zuwanderungshintergrund zu schärfen: Schon vor PISA 2000 war bekannt, dass Jugendliche mit Zuwanderungshintergrund eine geringere Bildungsbeteiligung aufweisen und niedrigere Bildungsabschlüsse erzielen als Jugendliche ohne Zuwanderungshintergrund (Alba, Handl & Müller, 1994; Esser, 1990; Nauck, Diefenbach & Petri, 1998). Dass es im internationalen Vergleich besonders große Disparitäten in den Kompetenzen zwischen Jugendlichen mit und ohne Zuwanderungshintergrund gibt, konnte jedoch erst anhand der in PISA 2000 erhobenen Daten gezeigt werden. Aus ethischer Sicht verletzen diese Disparitäten eine spezifische Form von Bildungsgerechtigkeit, nämlich das Prinzip des „*luck egalitarianism*“ (Stojanov, 2015), das besagt, dass eine Ungleichverteilung von Bildungserträgen nur dann akzeptabel ist, wenn sie aus vorsätzlichen Handlungen der Individuen hervorgeht (Rawls, 1999; Dworkin, 2000; Brighouse, 2003). Vieluf (im Druck) argumentiert, dass demnach der systematische Einfluss des Zuwanderungshintergrunds auf Bildungserfolg als ungerecht angesehen werden muss, da Jugendliche nicht dafür verantwortlich gemacht werden können, aus einer zugewanderten Familie zu stammen. Auch der *International Migration Outlook* der OECD (2013) sieht eine vergleichbar erfolgreiche Unterstützung von Schülerinnen und Schülern mit Zuwanderungshintergrund sowie Schülerinnen und Schülern ohne Zuwanderungshintergrund als eine wichtige Herausforderung für Bildungssysteme an.

Seit PISA 2000 wurden in Deutschland verschiedene Fördermaßnahmen auf den Weg gebracht, die sich gezielt an Kinder und Jugendliche aus zugewanderten Familien richten (z. B. Becker & Beck, 2009; Stanat, Becker, Baumert, Lüdtke & Eckhardt, 2012). Gleichzeitig gab es auf Schulsystemebene Veränderungen, beispielsweise die Erhöhung der Gymnasialquote und zahlreiche Reformen in einzelnen Ländern, die beispielsweise die Anzahl angebotener Schulformen reduzierten. Alle diese Veränderungen könnten mit einer Verringerung der zuwanderungsbezogenen Disparitäten einhergehen. Für die bereits zweimal als Schwerpunktdomäne in PISA erfassten Domänen Lesen und Mathematik kann anhand der Trendanalysen nachvollzogen werden, inwieweit seit dem Jahr

© Waxmann Verlag GmbH

2000 eine solche Reduzierung der Disparitäten stattgefunden hat. Auf Basis des Vergleichs der PISA-Erhebungen der Jahre 2000 und 2009 ließ sich ein positiver Trend in der Lesekompetenz von Jugendlichen mit Zuwanderungshintergrund nachweisen. Dieser ging vor allem auf eine Verbesserung der Testleistungen der selbst mit ihren Eltern nach Deutschland zugewanderten Jugendlichen zurück. Für die mathematische Kompetenz ließ sich zwischen den PISA-Erhebungen der Jahre 2003 und 2012 ebenfalls ein positiver Trend für Jugendliche mit Zuwanderungshintergrund nachweisen. Dieser ging aber auf die Verbesserung der mathematischen Kompetenzen von Jugendlichen zurück, deren Eltern zwar im Ausland geboren wurden, die selbst jedoch bereits in Deutschland zur Welt kamen. Jugendliche ohne Zuwanderungshintergrund konnten weder im Trend der Lesekompetenz (2000–2009) noch im Trend der Mathematikkompetenz (2003–2012) signifikante Verbesserungen erzielen, sodass sich die Disparitäten zwischen Jugendlichen mit und Jugendlichen ohne Zuwanderungshintergrund reduzierten. Auf Basis der Daten der Erhebung des Jahres 2015 stellt sich die Frage, ob es für Jugendliche mit Zuwanderungshintergrund ebenfalls einen positiven Trend für die naturwissenschaftliche Kompetenz gibt.

Naturwissenschaften und schulische Förderung

In PISA 2015 steht die naturwissenschaftliche Kompetenz nach PISA 2006 zum zweiten Mal als Hauptdomäne im Mittelpunkt. Dieses Kapitel geht deshalb zentral der Frage nach, ob sich für diesen Kompetenzbereich seit PISA 2006 Veränderungen der zuwanderungsbezogenen Disparitäten nachweisen lassen. Der Erwerb naturwissenschaftlicher Kompetenzen ist eng an den naturwissenschaftlichen Unterricht geknüpft. Die Enkulturation in mathematisch-naturwissenschaftlichen Fächern im Rahmen formaler Bildungsprozesse könnte so auch Schülerinnen und Schülern mit Zuwanderungshintergrund helfen, Barrieren bei der beruflichen Entwicklung zu überwinden. Für die Teilhabe am Arbeitsmarkt sind Naturwissenschaften von besonderer Bedeutung. Entsprechend soll der Unterricht in diesen Schulfächern auch auf eine anschließende Tätigkeit vorbereiten (vgl. Kapitel 6 sowie KMK, 2005a; 2005b; 2005c). Frank, Härtig und Neumann (im Druck) stellen für einzelne Ausbildungsberufe Evidenz dafür vor, dass in der Mittelstufe erworbene Kompetenzen in Physik tatsächlich den Erfolg in der dualen Ausbildung partiell erklären können.

Die Bedingungen für einen erfolgreichen Einstieg in naturwissenschaftlich-technische Berufe sind für Schülerinnen und Schüler mit Zuwanderungshintergrund günstig. Aufgrund des demografischen Wandels sinkt die Zahl potenzieller Arbeitnehmer und Arbeitnehmerinnen systematisch (vgl. Autorengruppe Bildungsberichterstattung, 2016). In Teilen zeigt sich, dass Zuwanderung den demografischen Wandel kompensieren kann. Schülerinnen und Schüler mit Zuwanderungshintergrund könnten insbesondere durch eine erfolgreiche Teilnahme am mathematisch-naturwissenschaftlichen Unterricht ihre Chancen am Arbeitsmarkt deutlich steigern. Bislang scheint sich dieser Effekt jedoch nicht einzustellen: Damelang und Haas (2006) wiesen bereits vor zehn Jahren

© Waxmann Verlag GmbH

darauf hin, dass von der zunehmenden Technisierung der Ausbildungsberufe Migranten nur unterproportional profitieren. Auch heute ist die Erwerbstätigkeit bei Personen mit Zuwanderungshintergrund insgesamt niedriger ausgeprägt (Bundesagentur für Arbeit, 2016). Dabei kommt es zu differenziellen Effekten: Während es manchen Herkunftsgruppen gelingt, zu partizipieren, erfahren andere eher Nachteile (Seebaß & Siegert, 2011; Segeritz, Walter & Stanat, 2010).

Es stellt sich die Frage, warum mathematisch-naturwissenschaftlicher Unterricht für Jugendliche mit Zuwanderungshintergrund im Durchschnitt nicht zu vergleichbarer Kompetenz führt wie für Jugendliche ohne Zuwanderungshintergrund. Im Sinne eines Angebots-Nutzungs-Modells formaler Bildungsprozesse sind individuelle und familiäre Ursachen genauso möglich wie systemseitige Ursachen. Bezogen auf letztere scheint es für eine Schülerin oder einen Schüler mit Zuwanderungshintergrund mehr Barrieren auf dem Weg zum Bildungserfolg zu geben. Jugendliche mit Zuwanderungshintergrund besuchen nach wie vor seltener ein Gymnasium als Jugendliche ohne Zuwanderungshintergrund (Arbeitsgruppe Bildungsberichterstattung, 2016). Somit muss man von einem systematischen, nachteiligen Effekt auf die erworbenen Kompetenzen der Schülerinnen und Schüler mit Zuwanderungshintergrund ausgehen. Üblicherweise werden auf der individuellen bzw. familiären Seite zur Erklärung der sozioökonomische Status der Eltern sowie die Bildungsgeschichte und die Sprachbiografie der Lernenden selbst herangezogen. Der sozioökonomische Status und die elterliche Bildung dienen als Indikatoren dafür, ob die Schülerinnen und Schüler zu Hause Unterstützung beim Lernen erfahren und somit eventuelle Probleme im Unterricht kompensieren können.

Die Bedeutung der zu Hause gesprochenen Sprache

Bis zur PISA-Erhebung des Jahres 2006 konnten enge Zusammenhänge der gemessenen Kompetenz mit der zu Hause gesprochenen Sprache festgestellt werden (vgl. z. B. Stanat & Christensen, 2006). So gab es in der PISA-Erhebung des Jahres 2000 für Jugendliche mit Zuwanderungshintergrund, die zu Hause hauptsächlich ihre Herkunftssprache mit ihrer Familie sprachen, einen Lesekompetenznachteil von 60 Punkten gegenüber denjenigen mit Zuwanderungshintergrund, die hauptsächlich Deutsch mit ihrer Familie sprachen. Bereits in der PISA-Erhebung des Jahres 2009 wurde die Stärke dieses Zusammenhangs deutlich geringer. Hier reduzierte sich dieser Nachteil auf 24 Punkte. Für die mathematische Kompetenz konnte in der PISA-Erhebung des Jahres 2012 kein signifikanter Nachteil für Jugendliche mit Zuwanderungshintergrund, die zu Hause hauptsächlich ihre Herkunftssprache mit ihrer Familie sprachen, nachgewiesen werden. In aktuellen Studien zeigt sich die Kompetenz in der Unterrichtssprache konsistent als wichtiger Prädiktor für den Erfolg auch in mathematisch-naturwissenschaftlichen Fächern (vgl. Härtig, Bernholt et al., 2015; Härtig et al., 2016). Anhand der PISA-2015-Daten soll deshalb geprüft werden, ob sich die zu Hause gesprochene Sprache als signifikanter Prädiktor der naturwissenschaftlichen Kompetenz erweist. Einschränkend muss jedoch festgestellt werden, dass die im Folgenden vorgestellten Analysen auf der in PISA genutzten

© Waxmann Verlag GmbH

globalen Frage nach der Sprachnutzung zu Hause: „Welche Sprache sprichst du am häufigsten zu Hause?“ beruhen. Als Antwortkategorien stehen die zehn in früheren PISA-Erhebungsrunden am häufigsten angegebenen Sprachen zur Auswahl sowie eine offene Antwortkategorie für Ergänzungen. Neuere und nach Sprechpartnern differenzierte Reanalysen der internationalen PISA-2012-Daten zeigen, dass die Nutzung der Herkunftssprache im Familienkontext und mit Freunden unter Kontrolle weiterer Merkmale keinen einheitlichen Zusammenhang mit der mathematischen und der Lesekompetenz aufweist (Agirdag & Vanlaar, 2016). Differenzierte Analysen nach Gesprächspartner oder Kommunikationszweck oder gar Analysen auf Basis von Messungen tatsächlicher sprachlicher Kompetenzen können auf Basis der PISA-2015-Daten nicht vorgenommen werden.

Im Folgenden werden in der Tradition der bisherigen nationalen PISA-Berichterstattung die Lebens- und Lernumwelten sowie die Naturwissenschaftskompetenz von Jugendlichen mit Zuwanderungshintergrund beschrieben. Dabei werden sowohl eine internationale wie auch eine auf Deutschland zugeschnittene, vertiefende Perspektive eingenommen.

9.1.1 Methodische Vorbemerkungen

Im Folgenden sollen die Definition von Zuwanderungshintergrund sowie die Auswahl der Herkunftsländer näher beschrieben werden. Des Weiteren wird kurz auf den Umgang mit fehlenden Werten sowie Besonderheiten in der Darstellung der Schülerinnen und Schüler mit Zuwanderungshintergrund und ihre Einteilung in die Schularten des deutschen Bildungssystems eingegangen.

Definition des Zuwanderungshintergrunds

PISA 2015 schließt an die nationale Definition des Zuwanderungsstatus der vorhergegangenen Erhebungsrunden seit dem Jahr 2006 an (vgl. Gebhardt, Rauch, Mang, Sälzer & Stanat, 2013 sowie Stanat et al., 2010). Grundlage ist die Unterscheidung des Zuwanderungsstatus in die erste und zweite Generation (Stanat & Christensen, 2006). Des Weiteren werden in diesem Bericht – ebenfalls analog zum nationalen Vorgehen seit PISA 2003 – Schülerinnen und Schüler mit nur einem in Deutschland geborenen Elternteil extra ausgewiesen (entgegen dem internationalen Vorgehen der OECD, diese Jugendliche den Schülerinnen und Schülern ohne Zuwanderungshintergrund zuzuweisen). Es zeigen sich für diese Jugendlichen weiterhin Disparitäten im Vergleich zu Jugendlichen ohne Zuwanderungshintergrund (Gebhardt et al., 2013). Konkret wird in diesem Kapitel folgende Einteilung genutzt:

© Waxmann Verlag GmbH

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, Jugendliche/r in Deutschland (bzw. im jeweiligen OECD-Teilnehmerstaat) geboren.
Erste Generation:	Beide Elternteile und Jugendliche/r im Ausland geboren.

Differenzierung verschiedener Herkunftsgruppen

Untergliedert man die Schülerinnen und Schüler mit Zuwanderungshintergrund entsprechend ihren Herkunftsländern finden sich über alle PISA-Erhebungswellen hinweg deutliche Unterschiede zwischen den Gruppen im Bezug auf die im Mittel erreichten Kompetenzen, die zu Hause gesprochene Sprache sowie die soziale Herkunft (vgl. Gebhardt et al., 2013). Ebenso unterscheiden sich die Gruppen im Anteil der Jugendlichen, die den verschiedenen Generationen zugeordnet werden.

In diesem Berichtsband werden – analog zum Vorgehen bei früheren PISA-Erhebungen – die folgenden Herkunftsländer unterschieden:

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 sehr geringer Anteil an Jugendlichen stammt.

Aus Italien, Bosnien und Herzegowina sowie Griechenland stammen nach der ehemaligen UdSSR, der Türkei und Polen die meisten Familien von Jugendlichen mit Zuwanderungshintergrund. Nach eingehender Analyse wurden diese Jugendlichen jedoch aufgrund geringer Anteile der Kategorie „Anderes Land“ zugeschrieben. Inkonsistenzen aufgrund unterschiedlicher Abstammung des Vaters und der Mutter wurden aus den Analysen im Sinne einer eindeutigen Datengrundlage ausgeschlossen. Schülerinnen und Schüler, welche einen im Ausland und einen in Deutschland geborenen Elternteil haben, wurden dem Herkunftsland des zugewanderten Elternteiles zugeordnet.

Im Gegensatz zu früheren PISA-Erhebungen wurde für PISA 2015 darauf verzichtet, Analysen zu Jugendlichen mit Zuwanderungshintergrund zugleich unter Berücksichtigung des Herkunftslandes und des Generationsstatus vorzunehmen. Dies ist durch die Zunahme der Anteile der Gruppe aus einem anderen Herkunftsland und die damit einhergehenden größeren Standardfehler der Generationsstatusgruppen für die Herkunftsländer ehemalige UdSSR, Türkei und Polen begründet.

© Waxmann Verlag GmbH

Umgang mit fehlenden Werten und Betrachtung von Schülerinnen und Schülern in Förderschulen sowie beruflichen Schulen

Wie in den vorhergehenden Berichterstattungen zu PISA findet sich auch in dieser Erhebungsrunde eine Gruppe von Schülerinnen und Schülern, für welche aufgrund fehlender Werte keine eindeutige Zuordnung des Zuwanderungshintergrundes möglich ist (vgl. Gebhardt et al., 2013, sowie Stanat et al., 2010). National wird daher die Ausweisung des Zuwanderungsstatus um die Kategorie „nicht zuzuordnen" erweitert. International werden diese Schülerinnen und Schüler zum Teil aus den Analysen ausgeschlossen und zu anderen Teilen den anderen Zuwanderungskategorien zugeordnet, wodurch eine Vergleichbarkeit mit den Ausweisungen der OECD erschwert wird (OECD, 2016).

Ebenfalls werden analog zu bisherigen PISA-Erhebungsrunden Schülerinnen und Schüler aus Sonder- und Förderschulen sowie beruflichen Schulen in allen Analysen dieses Berichtsbands integriert, jedoch in schulartspezifischen Ausweisungen nicht mit dargestellt (vgl. Kapitel 1).

Die Lernumwelten von Schülerinnen und Schülern unterscheiden sich an verschiedenen Schularten erheblich. Daher wird auch eine vergleichende Analyse der zuwanderungsbezogenen Disparitäten in verschiedenen Schularten in Deutschland vorgenommen. Aufgrund der Heterogenität der Schullandschaft in Deutschland sind die Analysen für einzelne Schularten jedoch nicht mehr möglich. Daher wurden die Schularten Hauptschule, Integrierte Gesamtschule, Realschule sowie Schule mit mehreren Bildungsgängen zu einer gemeinsamen Kategorie „andere Schularten" zusammengefasst (vgl. Kapitel 1 dieses Berichtsbandes). Die Darstellung der erreichten Kompetenzsstufen in Naturwissenschaft, Mathematik und Lesen findet sich in Anhang A (Abbildung A1-3). Dort werden auch die mittleren Kompetenzen dieser drei Domänen für Schülerinnen und Schüler mit und ohne Zuwanderungshintergrund ausgewiesen (Tabelle B11).

9.2 Jugendliche mit Zuwanderungshintergrund im europäischen Vergleich

Im Folgenden werden zunächst die Anteile von Jugendlichen mit Zuwanderungshintergrund in ausgewählten europäischen Staaten dargestellt, bevor dann auf die Merkmale des familiären Hintergrunds der Jugendlichen in diesen Staaten und ihre naturwissenschaftlichen Kompetenzen eingegangen wird. Die für die internationale Übersicht ausgewählten Staaten beschränken sich auf europäische Staaten, was an aktuelle Publikationen zu Migrationsbewegungen angelehnt ist. Das Gutachten des Sachverständigenrats deutscher Stiftungen für Integration und Migration (2015) konzentriert sich beispielsweise hauptsächlich auf die europäischen Staaten. Die im vorliegenden Bericht gewählte Einschränkung auf Europa lässt sich im Wesentlichen durch zwei Argumente stützen: Zum einen gibt es eine zunehmende Zentralisierung der politischen Rahmenvorgaben

© Waxmann Verlag GmbH

durch europäische Institutionen, sodass sich streng genommen nur europäische Staaten sinnvoll vergleichen lassen. Zum anderen ist dieser Fokus auf Europa auch interessant, weil die demografische Entwicklung in den europäischen Ländern ähnlich verläuft und Arbeitsmigration aus den östlichen in die mitteleuropäischen Länder nachgelassen hat. Bei der Interpretation der Ergebnisse ist zu berücksichtigen, dass einige europäische Staaten (z. B. Polen) aufgrund zu geringer Anteile von Schülerinnen und Schülern aus zugewanderten Familien nicht für vergleichende Analysen berücksichtigt werden konnten.

Anteile von Jugendlichen mit Zuwanderungshintergrund im europäischen Vergleich

Die west-, süd-, und nordeuropäischen Staaten weisen einen substanziellen Anteil an Schülerinnen und Schülern mit Zuwanderungshintergrund auf (mehr als 10 Prozent).

Der Anteil der Jugendlichen mit Zuwanderungshintergrund liegt in Westeuropa zwischen 20 (Niederlande) und 52 Prozent (Schweiz). Belgien, Deutschland, Österreich, die Schweiz und das Vereinigte Königreich verzeichnen eine im Vergleich zum Zeitpunkt der PISA-Erhebung des Jahres 2006 erhöhte Zuwanderung, wobei diese nur in Belgien und dem Vereinigten Königreich mit einem erhöhten Anteil von Jugendlichen der ersten Generation einhergeht. In Südeuropa liegt der Anteil der Fünfzehnjährigen mit Zuwanderungshintergrund zwischen 16 und 27 Prozent. In Griechenland, Italien, Portugal und Spanien hat sich der Anteil der Jugendlichen mit Zuwanderungshintergrund seit 2006 signifikant erhöht, wobei in Italien und Spanien der Anteil der Jugendlichen der ersten Generation anstieg. In den nordeuropäischen Staaten Dänemark, Finnland, Norwegen und Schweden stieg seit der PISA-2006-Erhebung der Anteil der Jugendlichen mit Zuwanderungshintergrund ebenfalls signifikant an. Finnland, Norwegen und Schweden haben im gleichen Zeitraum auch einen signifikanten Anstieg von Jugendlichen der ersten Generation zu verzeichnen. In einigen osteuropäischen Staaten (Bulgarien, Slowakische Republik und Ungarn) ist der Anteil der Jugendlichen mit Zuwanderungshintergrund zum Teil deutlich geringer als in den west-, süd- und nordeuropäischen Staaten. In Bulgarien, der Slowakischen Republik und Slowenien hat sich seit PISA 2006 der Anteil der Jugendlichen der ersten Generation signifikant erhöht.

Für Deutschland ergibt sich für den Zeitpunkt der PISA-2015-Datenerhebung ein Anteil von knapp 28 Prozent Schülerinnen und Schüler mit Zuwanderungshintergrund. Verglichen mit dem Mikrozensus von 2015 scheint dies eine leichte Unterschätzung darzustellen. Je nachdem welche Population dort herangezogen wird, ergeben sich 33 Prozent (für die 10- bis 15-Jährigen) bzw. 29 Prozent (für die 15- bis 20-Jährigen). Für die Zukunft zeichnet sich ab, dass der Anteil der Schülerinnen und Schüler mit Zuwanderungshintergrund, auch unabhängig von der verstärkten Zuwanderung ab dem Jahr 2015, weiter steigen wird (vgl. Mikrozensus: jüngere Gruppen). Bereits jetzt findet sich im Vergleich von PISA 2006 und 2015 ein signifikanter Anstieg des Anteils der Jugendlichen mit Zuwanderungshintergrund in Deutschland um knapp 8 Prozentpunkte. Dieses

© Waxmann Verlag GmbH

Tabelle 9.1: Prozentuale Anteile von fünfzehnjährigen Schülerinnen und Schülern mit Zuwanderungshintergrund in ausgewählten OECD-Staaten

OECD-Staaten	Ohne Zuwanderungshintergrund	Mit Zuwanderungshintergrund															
		Insgesamt Zuwanderung				Ein Elternteil im Ausland geboren			Zweite Generation			Erste Generation					
	*gültige%**	(*SE*)	*gültige%**	(*SE*)	*+/-*	*gültige%**	(*SE*)	*+/-*	*gültige%**	(*SE*)	*+/-*	*gültige%**	(*SE*)	*+/-*	zuzuordnen	nicht zuzuordnen
Westeuropa																	
Belgien	67.0	(1.1)	33.0	(1.1)	**7.9**	15.4	(0.4)	**3.3**	8.9	(0.6)	**2.1**	8.6	(0.6)	**2.5**	95.9	4.1
Deutschland	72.2	(1.1)	27.8	(1.1)	**8.0**	11.0	(0.5)	**5.4**	13.1	(0.7)	**5.5**	3.7	(0.4)	**-2.9**	85.7	14.3
Frankreich	74.4	(1.3)	25.6	(1.3)	0.3	12.5	(0.5)	0.0	8.7	(0.8)	-0.8	4.4	(0.4)	**1.1**	96.6	3.4
Niederlande	80.2	(1.2)	19.8	(1.2)	0.3	9.2	(0.5)	0.9	8.4	(0.8)	0.8	2.1	(0.2)	**-1.4**	97.2	2.8
Österreich	69.4	(1.0)	30.6	(1.0)	**10.3**	10.4	(0.4)	**3.2**	12.6	(0.7)	**7.3**	7.6	(0.6)	-0.3	98.3	1.7
Schweiz	48.4	(1.4)	51.6	(1.4)	**11.2**	20.6	(0.8)	**2.6**	20.7	(1.0)	**8.9**	10.2	(0.5)	-0.3	97.3	2.7
Vereinigtes Königreich	71.9	(1.1)	28.1	(1.1)	**10.6**	11.6	(0.5)	**2.6**	7.9	(0.7)	**3.0**	8.6	(0.7)	**5.0**	94.8	5.2
Südeuropa																	
Griechenland	79.5	(0.8)	20.5	(0.8)	**5.7**	9.9	(0.5)	**2.6**	6.8	(0.5)	**5.7**	3.8	(0.4)	**-2.6**	97.6	2.4
Italien	84.2	(0.6)	15.8	(0.6)	**6.5**	8.1	(0.5)	**2.5**	3.1	(0.3)	**2.5**	4.6	(0.4)	**1.5**	96.4	3.6
Kroatien	73.1	(0.8)	26.9	(0.8)	-0.4	16.2	(0.6)	0.9	8.9	(0.5)	**4.1**	1.7	(0.2)	**-5.4**	97.1	2.9
Portugal	76.7	(0.7)	23.3	(0.7)	**7.6**	16.1	(0.6)	**6.1**	3.1	(0.2)	0.9	4.1	(0.3)	0.6	97.5	2.5
Spanien	82.2	(0.8)	17.8	(0.8)	**6.3**	6.8	(0.4)	**2.2**	1.9	(0.2)	**1.1**	9.0	(0.7)	**3.0**	97.4	2.6
Nordeuropa																	
Dänemark	78.7	(0.7)	21.3	(0.7)	**6.8**	10.7	(0.4)	**3.6**	7.9	(0.5)	**3.7**	2.8	(0.2)	-0.6	97.5	2.5
Finnland	89.3	(0.7)	10.7	(0.7)	**5.7**	6.8	(0.5)	**3.2**	1.8	(0.3)	**1.6**	2.2	(0.3)	**0.9**	97.9	2.1
Norwegen	77.8	(1.2)	22.2	(1.2)	**7.9**	10.2	(0.5)	**2.0**	6.0	(0.7)	**3.0**	6.0	(0.4)	**2.9**	96.0	4.0
Schweden	70.6	(1.3)	29.4	(1.3)	**8.1**	12.4	(0.5)	**1.8**	9.7	(0.8)	**3.6**	7.3	(0.7)	**2.7**	97.0	3.0
Osteuropa																	
Bulgarien	95.4	(0.3)	4.6	(0.3)	**1.9**	3.7	(0.3)	**1.1**	0.5	(0.1)	**0.4**	0.5	(0.1)	**0.4**	95.5	4.5
Russland	83.1	(0.8)	16.9	(0.8)	**-2.8**	10.2	(0.6)	-0.9	3.6	(0.3)	-0.1	3.0	(0.3)	**-1.7**	94.0	6.0
Slowakische Republik	93.2	(0.3)	6.8	(0.3)	0.8	5.7	(0.3)	0.1	0.5	(0.1)	**0.3**	0.5	(0.1)	**0.4**	96.6	3.4
Slowenien	84.0	(0.6)	16.0	(0.6)	**-2.7**	8.3	(0.4)	-0.3	4.5	(0.3)	**-4.0**	3.3	(0.3)	**1.5**	98.2	1.8
Tschechische Republik	89.1	(0.5)	10.9	(0.5)	1.3	7.6	(0.4)	-0.1	1.6	(0.2)	**0.9**	1.7	(0.2)	0.5	97.7	2.3
Ungarn	92.4	(0.4)	7.6	(0.4)	**3.9**	5.1	(0.3)	**3.0**	1.4	(0.2)	**1.1**	1.1	(0.2)	-0.2	97.6	2.4

Anmerkung: In der Tabelle werden gerundete Werte angegeben. Es können auf die Rundungen zurückzuführende vermeintliche Inkonsistenzen vorkommen.
+/-: Veränderung gegenüber PISA 2006
fett: signifikante Veränderungen gegenüber PISA 2006 ($p < .05$)
* gültige % = Prozentangaben beruhen nur auf Angaben der Schülerinnen und Schüler. die eindeutig zuzuordnen sind.

© Waxmann Verlag GmbH

Wachstum setzt den Trend fort, der unter anderem auf Basis von PISA-Daten seit der ersten Erhebung im Jahr 2000 dokumentiert wird.

Merkmale des familiären Hintergrundes von Jugendlichen mit Zuwanderungshintergrund im europäischen Vergleich

Der Unterricht an den Schulen findet mit wenigen Ausnahmen in der jeweiligen Verkehrssprache des Landes statt, die bei Jugendlichen mit Zuwanderungshintergrund häufig nicht als Erstsprache erworben wurde. In PISA wird als Sprachindikator die zu Hause am häufigsten genutzte Sprache erhoben. Geben Jugendliche hier an, zu Hause hauptsächlich eine andere Sprache als die jeweilige Landessprache zu sprechen, so wurde ein negativer Zusammenhang mit der naturwissenschaftlichen Kompetenz vor allem auf mangelnde Lerngelegenheiten für den Erwerb der Instruktionssprache Deutsch zurückgeführt (Stanat, 2008; Kempert et al., 2015). Es wird erwartet, dass Jugendliche der zweiten Generation in ihren Familien mehr auf die landesübliche Sprache zurückgreifen als Jugendliche der ersten Generation, da sie selbst bereits im Inland geboren wurden, die Eltern in den meisten Fällen länger im Inland leben und unter Umständen bereits selber das Bildungssystem durchlaufen haben.

Tabelle 9.2 zeigt, dass die Anteile fünfzehnjähriger Schülerinnen und Schüler mit Zuwanderungshintergrund, die zu Hause die Sprache des Einwanderungslandes nutzen, im Vergleich der Staaten unterschiedlich groß sind. Die Nutzung der deutschen Sprache in Familien von Jugendlichen mit Zuwanderungshintergrund in Deutschland liegt mit knapp 61 Prozent im Mittelfeld. In vielen westeuropäischen Staaten nimmt die Nutzung der Sprache des Einwanderungslands von der ersten zur zweiten Generation deutlich zu (etwa in Deutschland, Frankreich, Niederlande). In Österreich und der Schweiz, aber auch in Spanien, Finnland und Russland zeigt sich eine stabile Sprachnutzung über die Generationen. Dabei ist der Anteil der Jugendlichen, die zu Hause häufiger die Sprache des Einwanderungslandes sprechen, in Finnland mit 21 Prozent (erste Generation) und 25 Prozent (zweite Generation) niedrig, während dieser Anteil in Kroatien mit 85 Prozent (erste Generation) und 93 Prozent (zweite Generation) sehr hoch ist. Die großen Unterschiede in der Sprachnutzung in den Familien korrespondieren vermutlich mit dem Herkunftsland der Zuwanderer. Handelt es sich wie etwa in Kroatien um Jugendliche, die mit ihren Familien aus anderen Ländern zurückgekehrt sind, so erklärt dies die hohe Nutzung des Kroatischen in der ersten Generation. Ein ähnlicher Erklärungsmechanismus greift auch für die erste Generation in Frankreich, dem Vereinigten Königreich, Spanien und Belgien, wo die Nutzung der Sprache des Einwanderungslandes in der ersten Generation bei über 40 Prozent liegt. Diese Jugendlichen haben die Sprache des Einwanderungslandes vermutlich schon in ihrem Herkunftsland – einer ehemaligen Kolonie des Aufnahmelandes – erworben und nutzen sie im Einwanderungsland weiter zur Familienkommunikation. Gesamteuropäische oder auf einzelne Regionen begrenzte Entwicklungen lassen sich in Bezug auf die Sprachnutzung in den Familien mit Zuwanderungshintergrund nicht feststellen.

© Waxmann Verlag GmbH

Tabelle 9.2: Prozentuale Anteile von fünfzehnjährigen Schülerinnen und Schülern mit Zuwanderungshintergrund, die zu Hause die Sprache des Einwanderungslandes sprechen

OECD-Staaten	Mit Zuwanderungshintergrund							
	Insgesamt Zuwanderung		Ein Elternteil im Ausland geboren		Zweite Generation		Erste Generation	
	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
Westeuropa								
Belgien	59.9	(1.5)	75.1	(1.5)	50.2	(2.0)	42.7	(3.1)
Deutschland	60.6	(1.6)	83.6	(1.5)	52.1	(2.1)	22.7	(2.8)
Frankreich	73.7	(1.5)	90.0	(1.3)	67.0	(2.4)	40.4	(3.4)
Niederlande	68.9	(1.7)	88.6	(1.5)	56.0	(2.9)	34.3	(4.1)
Österreich	42.9	(1.7)	76.6	(1.8)	26.9	(1.8)	23.2	(2.5)
Schweiz	53.9	(1.4)	82.7	(1.6)	35.3	(1.6)	33.5	(2.2)
Vereinigtes Königreich	71.3	(1.8)	93.7	(1.0)	72.3	(2.2)	40.2	(3.0)
Südeuropa								
Griechenland	78.4	(1.7)	93.6	(1.3)	76.8	(2.5)	41.2	(3.9)
Italien	62.1	(2.2)	83.6	(1.7)	55.8	(4.4)	28.3	(3.2)
Kroatien	94.1	(0.8)	95.4	(0.9)	93.4	(1.5)	85.3	(4.4)
Portugal	89.8	(1.1)	96.8	(0.8)	90.4	(2.4)	61.7	(3.0)
Spanien	58.7	(2.3)	76.1	(2.1)	48.2	(4.8)	47.7	(3.0)
Nordeuropa								
Dänemark	70.4	(1.1)	91.6	(1.1)	55.6	(2.3)	31.0	(3.1)
Finnland	63.5	(2.5)	87.4	(1.6)	25.1	(5.0)	21.0	(4.0)
Norwegen	63.9	(1.6)	90.9	(1.2)	56.7	(2.9)	25.0	(2.6)
Schweden	54.3	(1.6)	85.2	(1.3)	41.6	(2.4)	18.9	(2.5)
Osteuropa								
Bulgarien	77.3	(2.9)	86.5	(2.6)	62.3	(9.6)	22.3	(8.1)
Russland	93.1	(1.1)	98.7	(0.6)	87.4	(3.1)	81.4	(3.4)
Slowakische Republik	82.4	(2.0)	89.9	(1.7)	62.3	(9.0)	23.7	(9.1)
Slowenien	58.0	(2.0)	87.7	(1.5)	39.8	(3.6)	7.3	(1.9)
Tschechische Republik	71.7	(2.0)	88.7	(1.5)	44.7	(5.7)	21.2	(4.6)
Ungarn	88.1	(1.8)	92.8	(1.6)	74.0	(4.9)	84.9	(5.0)

Als Indikator für den sozioökonomischen Hintergrund wird in PISA unter anderem der Highest International Socio-Economic Index of Occupational Status (HISEI, Ganzeboom, de Graaf, Treiman & de Leeuw, 1992) verwendet, der auf der Basis der sogenannten International Standard Classification of Occupations (ISCO) gebildet wird (vgl. Kapitel 8). In Tabelle 9.3 sind die Unterschiede zwischen Jugendlichen ohne Zuwanderungshintergrund und Jugendlichen aus zugewanderten Familien in unterschiedlichen Staaten dargestellt.

© Waxmann Verlag GmbH

Tabelle 9.3: Disparitäten im sozioökonomischen Status zwischen Familien ohne Zuwanderungshintergrund und Familien mit Zuwanderungshintergrund in ausgewählten OECD-Staaten

OECD-Staaten	Ohne Zuwanderungshintergrund		Mit Zuwanderungshintergrund							
			Insgesamt Zuwanderung		Ein Elternteil im Ausland geboren		Zweite Generation		Erste Generation	
	M	*(SE)*	*DIFF*	*(SE)*	*DIFF*	*(SE)*	*DIFF*	*(SE)*	*DIFF*	*(SE)*
Westeuropa										
Belgien	55.2	(0.5)	**-7.4**	(0.9)	**-4.4**	(0.9)	**-11.0**	(1.2)	**-9.7**	(1.4)
Deutschland	53.7	(0.5)	**-7.9**	(0.8)	**-2.4**	(1.0)	**-12.2**	(0.8)	**-9.8**	(2.2)
Frankreich	52.6	(0.6)	**-4.8**	(0.8)	**2.7**	(1.0)	**-12.9**	(1.4)	**-12.5**	(1.7)
Niederlande	55.5	(0.5)	**-4.8**	(0.9)	1.7	(1.2)	**-11.5**	(1.2)	**-10.8**	(2.0)
Österreich	53.1	(0.5)	**-5.9**	(0.8)	**2.6**	(0.9)	**-11.6**	(1.0)	**-8.9**	(1.6)
Schweiz	56.4	(0.6)	**-6.4**	(0.8)	**3.0**	(0.8)	**-14.6**	(0.9)	**-9.5**	(1.5)
Vereinigtes Königreich	56.6	(0.5)	0.0	(1.0)	1.7	(1.1)	**-2.7**	(1.4)	0.0	(1.9)
Südeuropa										
Griechenland	51.6	(0.7)	**-7.5**	(0.9)	1.8	(1.1)	**-15.3**	(1.5)	**-18.6**	(2.0)
Italien	50.8	(0.4)	**-7.5**	(0.9)	-1.3	(1.2)	**-11.8**	(2.1)	**-16.3**	(1.2)
Kroatien	46.9	(0.5)	-3.7	(0.7)	-1.3	(0.9)	**-7.5**	(0.9)	**-7.4**	(2.2)
Portugal	47.3	(0.6)	**6.1**	(0.8)	**9.5**	(0.8)	**4.2**	(1.7)	**-5.7**	(1.5)
Spanien	49.0	(0.7)	**-5.1**	(0.8)	**2.9**	(1.3)	**-6.2**	(2.3)	**-11.3**	(1.1)
Nordeuropa										
Dänemark	56.9	(0.5)	**-3.3**	(0.8)	**3.1**	(1.1)	**-12.1**	(1.4)	**-8.3**	(1.8)
Finnland	53.3	(0.6)	**-2.4**	(1.1)	1.1	(1.3)	**-9.1**	(2.5)	**-9.4**	(2.4)
Norwegen	63.0	(0.4)	**-5.5**	(0.8)	0.9	(0.9)	**-10.6**	(1.3)	**-12.1**	(1.8)
Schweden	59.1	(0.5)	**-4.7**	(0.9)	-0.3	(1.0)	**-8.8**	(1.5)	**-7.6**	(1.4)
Osteuropa										
Bulgarien	50.3	(0.6)	**-2.7**	(1.3)	-1.4	(1.4)	**-11.2**	(3.6)	-3.2	(5.8)
Russland	55.1	(0.6)	0.1	(0.9)	1.3	(1.0)	-1.4	(1.6)	-2.0	(2.1)
Slowakische Republik	48.1	(0.5)	**-2.9**	(1.3)	**-3.2**	(1.4)	-3.4	(4.2)	2.0	(5.3)
Slowenien	53.4	(0.3)	**-6.6**	(0.9)	0.6	(1.2)	**-13.5**	(1.4)	**-16.3**	(1.9)
Tschechische Republik	47.8	(0.3)	-1.6	(0.9)	-0.7	(1.1)	**-5.7**	(2.3)	-2.2	(2.3)
Ungarn	47.3	(0.5)	**4.6**	(1.3)	**4.2**	(1.5)	**9.1**	(2.1)	-0.1	(3.6)

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

© Waxmann Verlag GmbH

In Deutschland sowie in nahezu allen anderen betrachteten europäischen Staaten weisen zugewanderte Familien einen geringeren sozioökonomischen Status auf, wobei die Unterschiede nicht für alle Staaten signifikant sind. Signifikante Nachteile für zugewanderte Familien zeigen sich insbesondere für Staaten mit einem eher hohen sozioökonomischen Status der Familien ohne Zuwanderungshintergrund. Eine Ausnahme bildet das Vereinigte Königreich, wo trotz insgesamt hohem sozioökonomischen Status kein signifikanter Nachteil für zugewanderte Familien besteht. Es gibt auch Staaten, etwa Portugal und Ungarn, in denen zugewanderte Familien im Durchschnitt einen höheren sozioökonomischen Status als Familien ohne Zuwanderungsgeschichte haben.

Eine differenzierte Betrachtung der einzelnen Generationen zeigt Unterschiede im Hinblick auf die erste und die zweite Generation von Zuwanderern. In der Mehrzahl der Staaten sind mit einer Differenz von etwa zehn HISEI-Punkten die Familien am stärksten benachteiligt, bei denen die Jugendlichen im Aufnahmeland und beide Eltern im Ausland geboren sind (Zweite Generation). Jugendliche mit einem im Aufnahmeland und einem im Ausland geborenen Elternteil haben in den meisten Staaten einen ähnlichen oder sogar einen höheren sozioökonomischen Status als Jugendliche ohne Zuwanderungshintergrund. Auch für Jugendliche der ersten Generation bestehen in den meisten Ländern substanzielle und signifikante Disparitäten im sozioökonomischen Status.

In Deutschland zeigen sich für Jugendliche aus zugewanderten Familien im internationalen Vergleich besonders starke Disparitäten im sozioökonomischen Status, die für alle Zuwanderergenerationen statistisch signifikant sind. Für Jugendliche der ersten Generation sind substanzielle Disparitäten von etwa 10 HISEI-Punkten und für Jugendliche der zweiten Generation von etwa 12 HISEI-Punkten festgestellt werden.

Naturwissenschaftliche Kompetenz von Jugendlichen mit Zuwanderungshintergrund im europäischen Vergleich

In Tabelle 9.4 sind die Kompetenzunterschiede in den Naturwissenschaften zwischen Jugendlichen aus zugewanderten Familien und ihren Mitschülerinnen und Mitschülern ohne Zuwanderungshintergrund sowie die Veränderungen seit PISA 2006 im internationalen Vergleich dargestellt.

Schülerinnen und Schüler ohne Zuwanderungshintergrund erreichen in Deutschland im Mittel 532 Punkte, bei Jugendlichen aus zugewanderten Familien sind es 471 Punkte. Die geringsten Kompetenzwerte werden von Jugendlichen der ersten Generation erreicht (433 Punkte), aber auch für Jugendliche der zweiten Generation (461 Punkte) und Jugendliche mit einem im Ausland geborenen Elternteil (497 Punkte) bestehen substanzielle Disparitäten, die einem Leistungsabstand von etwa einer ganzen bzw. einer halben Kompetenzstufe entsprechen. Dieses Muster der Disparitäten findet sich in den meisten westeuropäischen Staaten. Auch hier bildet das Vereinigte Königreich eine Ausnahme; dort liegen nur für Schülerinnen und Schüler der ersten Generation zuwanderungsbezogene Disparitäten vor. Abweichende Muster finden sich für die Gruppe der süd- und nordeuropäischen Staaten, in denen zumeist keine signifikanten Kompetenznachteile

© Waxmann Verlag GmbH

Tabelle 9.4: Mittlere naturwissenschaftliche Kompetenz von Jugendlichen mit Zuwanderungshintergrund in ausgewählten OECD-Staaten

OECD-Staaten	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	522	(2.0)	-6.0	470[a]	(4.2)	5.5	491[a]	(4.2)	-2.2	455[a]	(5.8)	10.4	448[a]	(5.9)	17.7
Deutschland	532	(2.7)	-2.5	471[a]	(5.3)	9.8	497[a]	(5.1)	-1.5	461[a]	(5.9)	21.5	433[a]	(10.7)	-23.0
Frankreich	509	(2.4)	4.3	468[a]	(4.8)	-11.8	493[a]	(4.5)	-15.7	457[a]	(8.4)	-0.4	420[a]	(7.9)	-16.0
Niederlande	518	(2.5)	**-17.6**	480[a]	(5.8)	-7.6	504[a]	(5.4)	**-21.7**	462[a]	(8.2)	8.1	442[a]	(12.5)	-25.2
Österreich	513	(2.5)	-10.7	458[a]	(3.7)	-7.8	494[a]	(4.6)	**-32.1**	447[a]	(4.3)	16.6	428[a]	(6.9)	-7.1
Schweiz	533	(2.8)	-1.0	483[a]	(3.8)	2.3	513[a]	(4.0)	-7.3	462[a]	(4.6)	0.6	467[a]	(6.0)	**30.7**
Vereinigtes Königreich	516	(2.4)	-3.3	503[a]	(4.4)	-5.3	515	(5.0)	-12.1	503	(6.3)	9.4	486[a]	(7.7)	5.9
Südeuropa															
Griechenland	461	(4.2)	**-16.2**	438[a]	(4.8)	**-18.8**	461	(5.8)	**-19.8**	424[a]	(7.8)	**-39.0**	404[a]	(7.7)	-24.0
Italien	484	(2.6)	5.0	471[a]	(3.8)	14.7	489	(4.5)	10.2	464[a]	(6.7)	20.1	445[a]	(6.0)	**26.4**
Kroatien	479	(2.5)	**-17.0**	469[a]	(3.4)	**-19.4**	479	(3.9)	**-18.5**	454[a]	(5.1)	**-26.7**	456[a]	(10.2)	-18.7
Portugal	500	(2.5)	**22.9**	508	(3.8)	**43.6**	517[a]	(3.9)	**28.9**	504	(9.4)	**62.8**	475[a]	(6.3)	**63.2**
Spanien	498	(2.1)	5.3	476[a]	(3.9)	11.1	506	(4.0)	-0.1	471[a]	(9.8)	-16.0	454[a]	(4.7)	**23.9**
Nordeuropa															
Dänemark	511	(2.6)	6.9	475[a]	(3.9)	**21.8**	509	(5.0)	16.6	441[a]	(4.8)	23.5	441[a]	(7.8)	**26.4**
Finnland	535	(2.2)	**-30.4**	499[a]	(6.6)	**-33.1**	527	(6.4)	**-30.3**	462[a]	(10.3)	-45.6	443[a]	(10.6)	-24.0
Norwegen	507	(2.4)	**14.8**	478[a]	(3.3)	6.2	505	(4.6)	6.7	463[a]	(5.7)	26.8	447[a]	(6.5)	11.8
Schweden	508	(3.2)	-5.1	466[a]	(6.2)	-10.0	503	(6.7)	1.1	454[a]	(8.1)	-8.9	419[a]	(9.7)	-14.6
Osteuropa															
Bulgarien	451	(4.2)	15.6	420[a]	(10.5)	-38.7	430	(11.5)	-33.6	390[a]	(19.9)	30.8	376[a]	(22.1)	-17.7
Russland	489	(3.2)	9.6	489	(4.1)	5.5	494	(5.1)	-0.7	483	(7.3)	13.8	478	(10.0)	10.4
Slowakische Republik	466	(2.5)	**-23.8**	445[a]	(6.2)	**-39.4**	454	(6.7)	**-33.7**	404[a]	(21.6)	-34.0	395[a]	(16.9)	-59.3
Slowenien	522	(1.3)	-5.1	476[a]	(3.8)	**-15.8**	501[a]	(4.8)	**-16.2**	465[a]	(7.4)	-4.6	428[a]	(9.2)	**-48.5**
Tschechische Republik	496	(2.2)	**-21.2**	477[a]	(5.5)	-4.7	482[a]	(5.7)	-5.6	481	(13.7)	**60.4**	452[a]	(11.7)	-26.7
Ungarn	476	(2.5)	**-28.8**	497[a]	(6.3)	-9.2	496[a]	(7.2)	-16.0	520[a]	(12.8)	49.8	478	(16.1)	-30.9

fett = signifikante Unterschiede zwischen 2006 und 2015 ($p < .05$)
[a] signifikante Unterschiede zu Schülerinnen und Schülern ohne Zuwanderungshintergrund ($p < .05$)
+/- Mittelwertsdifferenz 2015–2006

© Waxmann Verlag GmbH

von Jugendlichen mit einem im Ausland geborenen Elternteil bestehen. Im Vergleich mit anderen europäischen Staaten ist der Kompetenzunterschied zwischen Jugendlichen aus zugewanderten Familien und Jugendlichen ohne Zuwanderungshintergrund in Deutschland am größten ausgeprägt. Dies gilt sowohl allgemein für Jugendliche aus zugewanderten Familien als auch innerhalb der einzelnen Generationen. Um etwa fünf bis zehn Punkte geringere, aber immer noch substanzielle Disparitäten finden sich in Belgien, der Schweiz und Österreich. In Tabelle 9.4 werden außerdem die bei PISA 2006 und PISA 2015 im Mittel erreichten naturwissenschaftlichen Kompetenzen getrennt für die einzelnen Gruppen verglichen. Hier zeigen sich in Deutschland für keine der Gruppen statistisch signifikante Veränderungen.

9.3 Jugendliche mit Zuwanderungshintergrund in Deutschland

Im Folgenden wird die Situation von Schülerinnen und Schülern mit Zuwanderungshintergrund für Deutschland differenzierter beschrieben. Dabei wird zunächst nach Generationsstatus und Herkunftsland gegliedert der Anteil von Jugendlichen mit Zuwanderungshintergrund dargestellt, bevor wichtige Merkmale des familiären Hintergrundes dieser Schülerinnen und Schüler vorgestellt werden. Für eine Einschätzung der aktuellen Bildungsbeteiligung von Jugendlichen mit Zuwanderungshintergrund wird ihr Anteil in den verschiedenen Schularten genutzt, soweit dies auf Basis von PISA-Daten möglich ist. In modellbasierten Analysen werden die Zusammenhänge zwischen Zuwanderungshintergrund, naturwissenschaftlicher Kompetenz und Merkmalen des familiären Hintergrunds untersucht.

Prozentuale Anteile von Jugendlichen mit Zuwanderungshintergrund in PISA 2006 und 2015

Wie in früheren PISA-Erhebungen (z. B. Gebhardt et al., 2013) gibt es auch in PISA 2015 einen Anteil von Fünfzehnjährigen, für die aufgrund fehlender Werte keine Angaben zum Status der Zuwanderung gemacht werden können. Wie in Tabelle 9.5 zu sehen ist, hat sich dieser Anteil mit 14 Prozent im Vergleich zu PISA 2006 verdoppelt, ist jedoch im Vergleich zu PISA 2012 (damals 18 Prozent) rückläufig. Somit kann für etwa 86 Prozent der fünfzehnjährigen Schülerinnen und Schüler der Zuwanderungshintergrund beschrieben werden.

Die Gruppe der Schülerinnen und Schüler mit Zuwanderungshintergrund ist im Vergleich zur PISA-Erhebung im Jahr 2006 um knapp 8 Prozent angewachsen. Diese Steigerung findet man in den Anteilen der Jugendlichen, bei welchen nur ein Elternteil im Ausland geboren worden ist (Steigerung um gut 5 Prozent) sowie bei den Schülerinnen und Schülern der zweiten Generation (Steigerung um ebenfalls gut 5 Prozent). Der Anteil von Jugendlichen, die bereits in Deutschland geboren wurden, hat sich seit PISA

© Waxmann Verlag GmbH

Tabelle 9.5: Prozentuale Anteile fünfzehnjähriger Schülerinnen und Schüler mit Zuwanderungshintergrund in Deutschland

Zuwanderungsstatus	**2006**		**2015**	
	%	*(SE)*	*%*	*(SE)*
Zuwanderungsstatus eindeutig zuzuordnen?				
Nicht zuzuordnen	7.7	(0.5)	**14.3**	(0.8)
Zuzuordnen	92.3	(0.5)	**85.7**	(0.8)
	*gültige %**	*(SE)*	*gültige %**	*(SE)*
Ohne Zuwanderungshintergrund	80.1	(1.2)	**72.2**	(1.1)
Mit Zuwanderungshintergrund	19.9	(1.2)	**27.8**	(1.1)
Generationsstatus				
Ein Elternteil im Ausland geboren	5.6	(0.4)	**11.0**	(0.5)
Zweite Generation	7.6	(0.7)	**13.1**	(0.7)
Erste Generation	6.6	(0.5)	**3.7**	(0.4)
Herkunftsländer				
Ehemalige UdSSR	4.2	(0.4)	4.8	(0.4)
Türkei	5.0	(0.6)	5.5	(0.5)
Polen	2.1	(0.3)	2.4	(0.2)
Anderes Land	8.6	(0.6)	**15.2**	(0.7)

Anmerkung: In der Tabelle werden gerundete Werte angegeben. Es können auf die Rundungen zurückzuführende vermeintliche Inkonsistenzen vorkommen.
Daten beruhen auf Schülerangaben.
fett: signifikante Unterschiede zwischen 2006 und 2015 ($p < .05$)
* gültige % = Prozentangaben beruhen nur auf Angaben der Schülerinnen und Schüler, die eindeutig zuzuordnen sind.

2006 fast verdoppelt. Dagegen findet sich eine Halbierung des Anteils von Fünfzehnjährigen der ersten Generation im gleichen Zeitraum.

Mehr als die Hälfte aller Jugendlichen mit Zuwanderungshintergrund und 15 Prozent aller Fünfzehnjährigen in Deutschland stammt aus einem anderen Land als der Türkei, der ehemaligen UdSSR und Polen. Anteilig sind in dieser Gruppe besonders häufig Jugendliche vertreten, deren Familien aus Italien, Bosnien und Herzegowina oder Griechenland stammen. Allerdings können diese Herkunftsländer aufgrund der in der Stichprobe geringen Fallzahlen nicht separat untersucht werden.

Merkmale des familiären Hintergrundes von Jugendlichen mit Zuwanderungshintergrund

Die Analysen zum familiären Hintergrund von Jugendlichen aus zugewanderten Familien in Deutschland werden in den folgenden Abschnitten analog zum internationalen Vergleich in Abschnitt 9.3 dargestellt. Tabelle 9.6 stellt neben der Familiensprache auch die Gruppenunterschiede für den sozioökonomischen Status, die kulturellen Ressourcen und den Bildungshintergrund der Eltern dar.

© Waxmann Verlag GmbH

PISA 2015 zeigt, dass etwa 61 Prozent der Schülerinnen und Schüler mit Zuwanderungshintergrund zu Hause Deutsch als Sprache der Familienkommunikation nutzen (vgl. Tabelle 9.6). Dieser Anteil ist im Vergleich zur PISA-Erhebung des Jahres 2006 stabil. Innerhalb der Gruppe der Schülerinnen und Schüler mit Zuwanderungshintergrund unterscheidet sich die Sprachpraxis in der Familie nach Generationsstatus. Während ein über die Zeit stabiler Anteil von 86 Prozent (PISA 2006) bzw. 84 Prozent (PISA 2015) der Jugendlichen mit einem im Ausland geborenen Elternteil zu Hause hauptsächlich das Deutsche nutzt, ist dieser Anteil für die Jugendlichen der zweiten Generation mit 45 Prozent (PISA 2006) bzw. 52 Prozent (PISA 2015) deutlich geringer. Ein signifikantes Absinken des Anteils von Deutsch als Familiensprache ist für die Jugendlichen der ersten Generation festzustellen. Während bei PISA 2006 noch knapp 41 Prozent der selbst zugewanderten Jugendlichen mit ihren Familien hauptsächlich Deutsch sprachen, sind es in 2015 nur noch knapp 23 Prozent. Darüber hinaus variiert der Anteil derjenigen, die zu Hause Deutsch sprechen, über die Herkunftsgruppen. In Familien, die aus der Türkei stammen, ist dieser Anteil geringer als in den anderen Herkunftsgruppen.

Tabelle 9.6 verdeutlicht neben der zu Hause gesprochenen Sprache die Unterschiede zwischen Jugendlichen aus zugewanderten Familien und Jugendlichen ohne Zuwanderungshintergrund in Bezug auf den sozioökonomischen Status. Sowohl die Unterschiede zwischen Jugendlichen mit nur einem im Ausland geborenen Elternteil als auch zwischen Jugendlichen der ersten und zweiten Generation zu Jugendlichen ohne Zuwanderungshintergrund sind signifikant. Besonders deutlich weichen die sozioökonomischen Bedingungen, unter denen Jugendliche mit dem Herkunftsland Türkei aufwachsen, von den Bedingungen der Jugendlichen ohne Zuwanderungshintergrund ab. Da die Vergleichbarkeit des sozioökonomischen Status zwischen PISA 2006 und PISA 2015 durch die Überarbeitung der Berufsklassifizierung nicht eindeutig ist, wurde auf eine Analyse der Veränderung im HISEI seit PISA 2006 verzichtet.

Der Indikator für die kulturellen Ressourcen basiert in PISA 2015 auf Angaben der Schülerinnen und Schüler zum Besitz von klassischer Literatur, Gedichtbänden und Kunstwerken. Auf Basis der PISA-2015-Daten ist festzustellen, dass Schülerinnen und Schüler der ersten und zweiten Generation in Familien leben, die über niedrigere kulturelle Ressourcen im oben definierten Sinne verfügen als Schülerinnen und Schüler ohne Zuwanderungshintergrund. Die kulturellen Ressourcen aller Gruppen von Schülerinnen und Schülern sind gegenüber PISA 2006 unverändert. In Bezug auf das Herkunftsland sind es in PISA 2006 und 2015 die Schülerinnen und Schüler mit aus der Türkei stammenden Familien, die zu Hause besonders wenig klassische Literatur, Gedichtbände oder Kunstwerke zur Verfügung haben.

Das Bildungsniveau der Eltern von Fünfzehnjährigen mit Zuwanderungshintergrund zeigt sich bei PISA 2015 signifikant niedriger als das der Eltern ihrer Mitschülerinnen und Mitschüler ohne Zuwanderungshintergrund. Allerdings ist das Bildungsniveau der Zuwandererfamilien bei Jugendlichen der zweiten Generation in PISA 2015 höher als in PISA 2006. Gegenüber PISA 2006 ist insbesondere das Bildungsniveau der Eltern von Jugendlichen aus der Türkei signifikant höher.

© Waxmann Verlag GmbH

Tabelle 9.6: Unterschiede in der gesprochenen Sprache sowie in der sozialen Herkunft zwischen Jugendlichen mit und Jugendlichen ohne Zuwanderungshintergrund

Zuwanderungsstatus	Deutsch als gesprochene Sprache zu Hause				Sozioökonomischer Status		Kulturelle Ressourcen [1]				Bildungsniveau der Eltern			
	2006		2015		2015		2006		2015		2006		2015	
	*gültige %**	*(SE)*	*gültige %**	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*	*M*	*(SE)*
Ohne Zuwanderung	99.5	(0.2)	99.1	(0.1)	53.3	(0.5)	0.0	(0.0)	0.1	(0.0)	14.4	(0.1)	14.5	(0.1)
Mit Zuwanderung	55.3[a]	(2.3)	60.7[a]	(1.6)	45.5[a]	(0.6)	-0.1[a]	(0.0)	-0.1[a]	(0.0)	13.6[a]	(0.2)	13.7[a]	(0.1)
Generationsstatus														
Ein Elternteil im Ausland geboren	86.0[a]	(2.4)	83.6[a]	(1.5)	50.6[a]	(0.9)	0.1	(0.1)	0.0	(0.0)	14.4	(0.2)	14.3	(0.2)
Zweite Generation	45.0[a]	(4.4)	52.3[a]	(2.1)	41.5[a]	(0.7)	-0.2[a]	(0.1)	-0.2[a]	(0.0)	12.5[a]	(0.3)	**13.1[a]**	(0.2)
Erste Generation	40.5[a]	(3.0)	**22.7[a]**	(2.8)	44.4[a]	(1.8)	-0.2[a]	(0.1)	-0.2[a]	(0.1)	14.1	(0.3)	14.1	(0.2)
Herkunftsländer														
Ehem. UdSSR	55.1[a]	(4.9)	57.0[a]	(3.3)	44.4[a]	(1.4)	-0.1	(0.1)	-0.1	(0.3)	14.1	(0.3)	13.9[a]	(0.2)
Türkei	40.0[a]	(4.3)	46.0[a]	(2.9)	39.2[a]	(1.1)	-0.3[a]	(0.1)	-0.3	(0.3)	11.6[a]	(0.5)	**12.6[a]**	(0.2)
Polen	60.3[a]	(5.5)	67.3[a]	(4.3)	47.1[a]	(1.6)	-0.1	(0.1)	0.0	(0.3)	14.7	(0.3)	14.4	(0.3)
Anderes Herkunftsland	62.1[a]	(3.6)	66.1[a]	(1.9)	47.7[a]	(0.9)	0.0	(0.1)	-0.1	(0.3)	14.0	(0.3)	13.9[a]	(0.2)

Anmerkung: Die Angaben beruhen auf Schülerangaben. **fett**: signifikante Unterschiede zwischen 2006 und 2015 ($p < .05$)
[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.

Bildungsbeteiligung von Jugendlichen mit Zuwanderungshintergrund

Im deutschen Schulsystem ist die Bildungsbeteiligung von Schülerinnen und Schülern an den verschiedenen weiterführenden Schularten für die im späteren Leben erreichbaren beruflichen, sozialen und ökonomischen Positionen von erheblicher Bedeutung (Becker, 2009). Tabelle 9.7 stellt die Bildungsbeteiligung der fünfzehnjährigen Schülerinnen und Schüler mit und ohne Zuwanderungshintergrund getrennt nach Schulart dar. Da die Vielfalt der Schulformen in Deutschland momentan über alle Bundesländer im Wandel ist und nur das Gymnasium als eigentlich konstante Größe manifestiert ist (Autorengruppe Bildungsberichterstattung, 2016), werden in diesem Berichtsband nur die Schularten Gymnasium vs. nicht gymnasiale Schularten miteinander verglichen, wobei sich Letztere aus Hauptschulen, Integrierten Gesamtschulen, Realschulen sowie Schulen mit mehreren Bildungsgängen zusammensetzen (vgl. Kapitel 1).

Insgesamt ist die Gymnasialbeteiligung in der Gruppe der Schülerinnen und Schüler ohne Zuwanderungshintergrund deutlich höher als in der Gruppe der Schülerinnen und Schüler mit Zuwanderungshintergrund. Der Anteil der Fünfzehnjährigen ohne Zuwanderungshintergrund, die ein Gymnasium besuchen, ist um 11 Prozent höher als für Schülerinnen und Schüler mit Zuwanderungshintergrund. Innerhalb der Gruppe der Jugendlichen mit Zuwanderungshintergrund sind mehr mit nur einem im Ausland geborenen Elternteil am Gymnasium als Jugendliche der zweiten und vor allem der ersten Generation. Für die Bildungsbeteiligung zeigt sich seit PISA 2006 weder für alle Jugendlichen mit Zuwanderungshintergrund noch differenziert nach Generationsstatus eine Veränderung.

Tabelle 9.7: Prozentuale Anteile fünfzehnjähriger Schülerinnen und Schüler mit Zuwanderungshintergrund an Gymnasien und nicht gymnasialen Schularten (Hauptschule, Integrierte Gesamtschule, Realschule und Schule mit mehreren Bildungsgängen)

	Gymnasium		**nicht gymnasiale Schularten***	
	%	*(SE)*	*%*	*(SE)*
Ohne Zuwanderungshintergrund	38.7	(2.4)	57.6	(2.4)
Mit Zuwanderungshintergrund	26.7	(2.2)	69.0	(2.3)
Generationsstatus				
Ein Elternteil im Ausland geboren	31.0	(2.5)	65.2	(2.5)
Zweite Generation	25.7	(2.5)	70.4	(2.7)
Erste Generation	17.7	(3.2)	75.9	(4.8)
Nicht zuzuordnen	21.1	(2.7)	67.1	(3.7)

* nicht gymnasiale Schularten setzten sich zusammen aus Hauptschule, Integrierte Gesamtschule, Realschule und Schule mit mehreren Bildungsgängen

© Waxmann Verlag GmbH

Disparitäten in der naturwissenschaftlichen Kompetenz

Mithilfe von multivariaten Regressionsanalysen wurde geprüft, ob die Kompetenzunterschiede in Naturwissenschaften zwischen Jugendlichen mit und ohne Zuwanderungshintergrund auch unter Kontrolle weiterer Hintergrundmerkmale bestehen bleiben. Um die Veränderungen seit PISA 2006 darstellen zu können, wurden parallele Modelle für PISA 2006 und PISA 2015 spezifiziert. Die Ergebnisse der Analysen sind differenziert nach Generationsstatus in Tabelle 9.8 dargestellt.

In die Regressionsmodelle wurden schrittweise zunächst der Generationsstatus (ein Elternteil im Ausland geboren, zweite Generation, erste Generation und nicht zuzuordnen) aufgenommen (Modell I), im zweiten Schritt der sozioökonomische Status, die Kulturgüter und das Bildungsniveau der Eltern (Modell II) und schließlich die zu Hause mit der Familie gesprochene Sprache (Modell III). In der ersten Zeile wird jeweils der Achsenabschnitt ausgegeben, der in Modell I dem Mittelwert der Jugendlichen ohne Zuwanderungshintergrund entspricht. Die Regressionsgewichte ab Zeile 2 geben die Punktedifferenz der einzelnen Generationsstatusgruppe zu den Jugendlichen ohne Zuwanderungshintergrund wieder.

In Modell I zeigen sich sowohl in PISA 2006 als auch in PISA 2015 signifikante Unterschiede in der naturwissenschaftlichen Kompetenz zwischen Jugendlichen mit Zuwanderungshintergrund und solchen ohne Zuwanderungshintergrund. In PISA 2015 wie in PISA 2006 sind die Disparitäten zwischen Jugendlichen mit Zuwanderungshintergrund aller Generationsstatusgruppen und Jugendlichen ohne Zuwanderungshintergrund signifikant. Während bei PISA 2006 die Jugendlichen der zweiten Generation deutlich die niedrigsten naturwissenschaftlichen Kompetenzen aufwiesen, sind dies bei PISA 2015 die Jugendlichen der ersten Generation. In Modell I erweist sich der Rückgang der naturwissenschaftlichen Kompetenz in der ersten Generation zwischen PISA 2006 und 2015 um 28 Punkte als signifikant. Dieses Ergebnis muss allerdings aufgrund des großen Standardfehlers und der für den Modellvergleich reduzierten Stichprobe mit Vorsicht betrachtet werden. Eine Reduzierung der Stichprobe um die Personen, zu denen Angaben zur gesprochenen Sprache fehlten, ist im Modellvergleich notwendig, um die Regressionsmodelle in ihrer Aussage miteinander vergleichen zu können. Betrachtet man die Ausprägungen der mittleren naturwissenschaftlichen Kompetenz bei Jugendlichen der ersten Generation in Deutschland in Tabelle 9.4, so zeigt sich hier im Trend kein signifikanter Rückgang der naturwissenschaftlichen Kompetenz. Diese Analysen wurden auf Basis der gesamten Stichprobe vorgenommen und sind deshalb zur Einschätzung des tatsächlichen Trends besser geeignet als das Basismodell der Regressionsanalysen (Modell I, Tabelle 9.8). In der Zusammenschau der Analysen in Tabelle 9.4 und 9.8 lässt sich feststellen, dass es im Trend für keine der Generationsstatusgruppen einen signifikanten Zuwachs der naturwissenschaftlichen Kompetenzen gibt.

In Modell II zeigt sich sowohl für PISA 2015 als auch für PISA 2006, dass der sozioökonomische Status, die Kulturgüter und das Bildungsniveau der Eltern signifikante Prädiktoren der naturwissenschaftlichen Kompetenz sind. Je höher also der sozioökonomi-

© Waxmann Verlag GmbH

Tabelle 9.8: Regressionsmodelle zur Schätzung von Disparitäten in der naturwissenschaftlichen Kompetenz zwischen Jugendlichen ohne Zuwanderungshintergrund und Jugendlichen mit Zuwanderungshintergrund

	Modell I		**Modell II**		**Modell III**	
2015	b	*(SE)*	b	*(SE)*	b	*(SE)*
Ohne Zuwanderung	532*	(2.7)	528*	(2.3)	528*	(2.3)
Ein Elternteil im Ausland geboren	-35[a]	(5.3)	-32[a]	(4.7)	-25[a]	(4.6)
Zweite Generation	-71[a]	(6.2)	-51[a]	(6.0)	-30[a]	(5.9)
Erste Generation	**-99[a]**	(10.9)	**-85[a]**	(10.4)	**-52[a]**	(9.9)
Nicht zuzuordnen	-63[a]	(12.1)	-53[a]	(11.2)	-44[a]	(10.4)
HISEI [1]			23[a]	(1.5)	23[a]	(1.5)
Kulturgüter [1]			12[a]	(1.4)	12[a]	(1.4)
Bildungsniveau der Eltern [1]			**8[a]**	(1.6)	**8[a]**	(1.7)
Sprachgebrauch [2]					-43[a]	(5.1)
N	5794		5794		5794	
R^2	0.09		0.21		0.22	
	Modell I		**Modell II**		**Modell III**	
2006	b	*(SE)*	b	*(SE)*	b	*(SE)*
Ohne Zuwanderung	537*	(3.2)	532[a]	(2.7)	533[a]	(2.6)
Ein Elternteil im Ausland geboren	-30[a]	(7.8)	-31[a]	(6.9)	-25[a]	(6.9)
Zweite Generation	-82[a]	(8.5)	-54[a]	(7.8)	-29[a]	(7.7)
Erste Generation	-71[a]	(8.5)	-53[a]	(8.3)	-25[a]	(9.3)
Nicht zuzuordnen	-55[a]	(9.2)	-43[a]	(8.4)	-33[a]	(8.4)
HISEI [1]			21[a]	(1.7)	21[a]	(1.7)
Kulturgüter [1]			14[a]	(1.4)	14[a]	(1.4)
Bildungsniveau der Eltern [1]			16[a]	(1.6)	16[a]	(1.5)
Sprachgebrauch [2]					-48[a]	(8.1)
N	4314		4314		4314	
R^2	0.08		0.23		0.24	

Anmerkung: 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 2006 statistisch signifikant ($p < .05$)
[a] signifikante Partialregressionskoeffizienten ($p < .05$)
[1] z-standardisiert
[2] Referenzgruppe: Deutsch als Familiensprache

© Waxmann Verlag GmbH

sche Status, der Besitz an Kulturgütern und das Bildungsniveau der Eltern sind, desto höher ist auch die naturwissenschaftliche Kompetenz. Der Zusammenhang zwischen Bildungsniveau der Eltern und naturwissenschaftlicher Kompetenz der Jugendlichen hat sich seit PISA 2006 signifikant auf nur noch acht Punkte verringert. Nimmt man alle drei Indikatoren in das Regressionsmodell auf, dann sinken die Regressionsgewichte der Variablen für die Generationsstatusgruppen deutlich ab. Eine Verringerung von Regressionsgewichten bei Aufnahme weiterer signifikanter Prädiktoren in das Regressionsmodell kann für das Vorliegen einer (teilweisen) Mediation sprechen. Demnach könnte ein Teil des Zusammenhangs zwischen Zuwanderungsstatus und naturwissenschaftlicher Kompetenz durch den sozioökonomischen Status, die Kulturgüter und das Bildungsniveau der Eltern vermittelt werden. Eine Ausnahme bildet das Regressionsgewicht für die Jugendlichen mit nur einem im Ausland geborenen Elternteil, dieses verändert sich durch Hinzunahme der Indikatoren der sozialen Herkunft kaum.

In Modell III wird zusätzlich die zu Hause gesprochene Sprache aufgenommen. Diese erweist sich wie bei PISA 2006 auch bei PISA 2015 als signifikanter Prädiktor der naturwissenschaftlichen Kompetenz. Wenn zu Hause eine andere Familiensprache als Deutsch gesprochen wird, ist die naturwissenschaftliche Kompetenz bei PISA 2015 im Mittel um 43 Punkte niedriger, als wenn zu Hause hauptsächlich Deutsch gesprochen wird. Der Zusammenhang zwischen zu Hause gesprochener Sprache und naturwissenschaftlicher Kompetenz verändert sich im Trend nicht. In PISA 2015 reduziert die zusätzliche Aufnahme der zu Hause gesprochenen Sprache die Regressionsgewichte der zweiten Generation um 21 und der ersten Generation um 33 Punkte, was ebenfalls für eine teilweise Mediation des Zusammenhangs zwischen Zuwanderungsstatus und naturwissenschaftlicher Kompetenz durch die zu Hause gesprochene Sprache sprechen kann.

9.4 Zusammenfassung und Diskussion

Die Daten aus PISA 2015 zeigen, dass fast 28 Prozent der fünfzehnjährigen Schülerinnen und Schüler in Deutschland aus zugewanderten Familien stammen. Im Vergleich zu westeuropäischen Nachbarstaaten liegt dieser Anteil in Deutschland im mittleren Bereich, während er vor allem in den osteuropäischen Staaten niedriger ist. Seit PISA 2006 hat der Anteil Jugendlicher mit Zuwanderungshintergrund in Deutschland insgesamt um rund 8 Prozent zugenommen. In Süd-, West- und Nordeuropa gibt es in vielen Staaten eine vergleichbare Steigerung des Anteils. Dies belegt, dass viele europäische Bildungssysteme von einer heterogenen Schülerschaft geprägt sind. In großen Teilen Europas hat zwischen den PISA-Erhebungen der Jahre 2006 und 2015 vor allem der Anteil von Jugendlichen mit einem im Ausland geborenen Elternteil bzw. der zweiten Generation zugenommen. In vielen europäischen Staaten stagnierte hingegen der Anteil der Jugendlichen der ersten Generation oder ging sogar zurück.

In der eingehenderen Betrachtung der Situation von Fünfzehnjährigen mit Zuwanderungshintergrund im deutschen Bildungssystem zeigt sich, dass sich bei PISA 2015

© Waxmann Verlag GmbH

die Anteile von Jugendlichen der zweiten Generation (13 Prozent) und der Jugendlichen mit einem im Ausland geborenen Elternteil (11 Prozent) seit PISA 2006 fast verdoppelt haben. Fast ein Viertel aller Fünfzehnjährigen in Deutschland hat inzwischen einen Zuwanderungshintergrund, ist aber in Deutschland geboren und hat somit das deutsche Bildungssystem von Anfang an durchlaufen. Dagegen findet sich eine Halbierung des Anteils der Fünfzehnjährigen der ersten Generation (knapp 4 Prozent) im gleichen Zeitraum. Damit setzt sich die Abnahme der Jugendlichen der ersten Generation fort, die bereits in den vergangenen PISA-Erhebungen festgestellt wurde. Die Situation hat sich in Deutschland durch die vermehrte Zuwanderung im Jahr 2015 bereits wieder verändert. Dieser aktuellsten Entwicklung konnte im Rahmen von PISA 2015 jedoch nicht Rechnung getragen werden, da zugewanderte Schülerinnen und Schüler in der Regel erst nach einjährigem Schulbesuch in Deutschland zur Teilnahme an PISA verpflichtet sind.

Mit Blick auf die Herkunftsländer der Jugendlichen mit Zuwanderungshintergrund in Deutschland zeigen sich keine signifikanten Veränderungen für die zahlenmäßig größten Gruppen ehemalige UdSSR, Türkei sowie Polen. Dagegen verdoppelt sich der Anteil der Jugendlichen, deren Familien aus einem anderen als den drei genannten Herkunftsländern stammen, auf 15 Prozent. Die Heterogenität in Bezug auf die Herkunftsländer nimmt also im hier betrachteten Zeitraum zu. Das deutliche Anwachsen der Kategorie „Anderes Land“ führt dazu, dass eine Interpretation der Befunde über den Einbezug der Herkunftsländer, wie sie in der Tradition der bisherigen PISA-Berichterstattung auch im vorliegenden Kapitel für Jugendliche mit Herkunftsland Türkei, ehemalige UdSSR und Polen dargestellt wird, eine geringere Reichweite hat. Das zentrale Phänomen sind weniger die Veränderungen in den Gruppen als vielmehr die Zunahme an herkunftslandbezogener Heterogenität.

Was die zu Hause gesprochene Sprache der Jugendlichen mit Zuwanderungshintergrund im europäischen Vergleich angeht, so zeigt sich, dass die Varianz in der Sprachnutzung groß ist. Die Jugendlichen mit Zuwanderungshintergrund in Deutschland sprechen zu Hause weder besonders häufig noch besonders selten die Sprache des Einwanderungslandes. Immerhin knapp zwei Drittel geben an, zu Hause hauptsächlich Deutsch zu sprechen. Gesamteuropäische oder auf einzelne Regionen begrenzte Entwicklungen lassen sich in Bezug auf die Familiensprache nicht feststellen. Insbesondere in Frankreich, Belgien und dem Vereinigten Königreich (Staaten mit ehemaligen Kolonien), aber auch in Kroatien (Staat mit zurückkehrenden Familien) zeigt sich bereits in der ersten Generation eine hohe Nutzung der Sprache des Einwanderungslandes. In Deutschland – wie in vielen anderen europäischen Staaten – nimmt die Nutzung der Sprache des Einwanderungslands von der ersten zur zweiten Generation deutlich zu. In einigen anderen Staaten wie beispielsweise Finnland und Österreich stagniert die Nutzung der Sprache des Einwanderungslandes in den zugewanderten Familien im Vergleich der ersten und zweiten Generation.

In Deutschland unterscheidet sich der Anteil der Jugendlichen, die zu Hause Deutsch sprechen, nach Generationsstatus. Am häufigsten nutzen Jugendliche mit einem im Ausland geborenen Elternteil Deutsch als vorrangige Sprache in der Familie, den nächst-

© Waxmann Verlag GmbH

höheren Anteil hat Deutsch als Sprache in Familien von Jugendlichen der zweiten Generation und am seltensten wird Deutsch in Familien von Jugendlichen der ersten Generation gesprochen. Die Jugendlichen der ersten Generation nutzen bei PISA 2015 nur noch halb so häufig Deutsch als Familiensprache wie bei PISA 2006.

In den meisten europäischen Staaten ist der sozioökonomische Status zugewanderter Familien niedriger als der von nicht zugewanderten Familien. In Deutschland werden im europäischen Vergleich besonders starke Disparitäten im sozioökonomischen Status sichtbar, die für alle Gruppen der Jugendlichen aus zugewanderten Familien statistisch signifikant sind. In Bezug auf den sozioökonomischen Status müssen Jugendliche mit einem Zuwanderungshintergrund in Deutschland also nach wie vor als systematisch benachteiligt gelten. Diese Benachteiligung von Jugendlichen mit Zuwanderungshintergrund kann auch anhand des sozialen Gradienten nachvollzogen werden, der den kontinuierlich nachgewiesenen hohen Zusammenhang zwischen sozioökonomischem Status und der (naturwissenschaftlichen) Kompetenz erneut bestätigt (vgl. auch Kapitel 8). Der sozioökonomische Status der Jugendlichen aus zugewanderten Familien konnte nicht im Trend analysiert werden, da sich die entsprechenden Indikatoren zwischen PISA 2006 und PISA 2015 unterscheiden. Der Zugang zu häuslichen Kulturgütern sowie das Bildungsniveau der Eltern als weitere Merkmale der familiären Situation, in welcher die Jugendlichen mit bzw. ohne Zuwanderungshintergrund aufwachsen, zeigen sich in Deutschland über die Zeit recht stabil.

Disparitäten in der naturwissenschaftlichen Kompetenz

In PISA 2015 betragen die Disparitäten in der naturwissenschaftlichen Kompetenz zwischen Schülerinnen und Schüler mit und ohne Zuwanderungshintergrund 61 Punkte. Die geringsten Kompetenzwerte werden von Jugendlichen der ersten Generation erreicht (433 Punkte), aber auch Jugendliche der zweiten Generation (461 Punkte) und Jugendliche mit einem im Ausland geborenen Elternteil (497 Punkte) weisen substanzielle Disparitäten gegenüber Jugendlichen ohne Zuwanderungshintergrund auf. Dieses Muster der Disparitäten findet sich in den meisten europäischen Staaten. Für Deutschland zeigt sich im Vergleich der Jahre 2006 und 2015 weder für die Jugendlichen ohne Zuwanderungshintergrund noch für die Jugendlichen mit Zuwanderungshintergrund insgesamt noch für eine der Generationsstatusgruppen eine signifikante Verbesserung der naturwissenschaftlichen Kompetenz. Im europäischen Vergleich zeichnet sich ein sehr uneinheitliches Bild ab: Es gibt Staaten, in denen Jugendliche mit Zuwanderungshintergrund sehr große Kompetenzzuwächse erzielen, während in anderen die naturwissenschaftliche Kompetenz der Jugendlichen mit Zuwanderungshintergrund abnimmt.

In den für Deutschland durchgeführten multivariaten Regressionsanalysen wurde geprüft, ob sich zuwanderungsbezogene Disparitäten in der naturwissenschaftlichen Kompetenz teilweise auf die Merkmale des familiären Hintergrunds zurückführen lassen. Die Ergebnisse legen nahe, dass die betrachteten Merkmale des familiären Hintergrunds die Disparitäten nicht vollständig erklären können. Der Zuwanderungshinter-

© Waxmann Verlag GmbH

grund, die sozioökonomische Situation der Familien und die Familiensprache erklären gemeinsam nur 22 Prozent der Varianz in der naturwissenschaftlichen Kompetenz. Für die Naturwissenschaften scheint es einen deutlicheren Zusammenhang zwischen der zu Hause gesprochenen Sprache und den erworbenen Kompetenzen zu geben als für Lesen (PISA 2009) und Mathematik (PISA 2012). Dies spricht für die zentrale Rolle von Kompetenzen der Unterrichtssprache für den Kompetenzerwerb in naturwissenschaftlichen Fächern (vgl. Härtig, Bernholt et al., 2015). Die Ergebnisse sind jedoch mit Vorsicht zu interpretieren, da die in PISA gestellte Frage nach der hauptsächlich zu Hause genutzten Sprache den Jugendlichen nur die Entscheidung für *eine* Sprache ermöglicht. Daher können mehrsprachige Kommunikationspraktiken, wie sie in vielen Zuwandererfamilien üblich sind, nicht erfasst werden. Viele Jugendliche werden beim Ausfüllen dieser Frage eine nach Kommunikationspartnern, Kommunikationsanlässen und Kommunikationsbedürfnissen wechselnde mehrsprachige Sprachnutzungspraxis mehr schlecht als recht auf eine hauptsächlich genutzte Sprache reduzieren. Die Ergebnisse von Agirdag und Vanlaar (2016) zeigen aber, dass es für die in PISA gemessenen Kompetenzen durchaus einen Unterschied macht, mit wem Jugendliche welche Sprache sprechen.

Besondere Aufmerksamkeit soll zum Ende dieser Zusammenschau auf die Situation der Jugendlichen der ersten Generation gelenkt werden. Die Disparitäten in den naturwissenschaftlichen Kompetenzen der ersten Generation von zugewanderten Jugendlichen im Vergleich zu Jugendlichen ohne Zuwanderungshintergrund sind in PISA 2015 mit 99 Punkten sehr hoch. Seit PISA 2006 ist die Gruppe der Jugendlichen der ersten Generation kleiner geworden, sie macht statt knapp sieben nur noch knapp 4 Prozent aller Fünfzehnjährigen aus. Nur noch knapp 23 Prozent der ersten Generation nutzen im Jahr 2015 hauptsächlich Deutsch zur Familienkommunikation. Doch auch unter Kontrolle der zu Hause gesprochenen Sprache und des sozioökonomischen und bildungsbezogenen Hintergrunds der Familien bleiben die Disparitäten der ersten Generation signifikant und mit 52 Punkten hoch.

Handlungsbedarf

Die weiterhin bestehenden Disparitäten zwischen Jugendlichen mit und ohne Zuwanderungshintergrund weisen auf Handlungsbedarf hin. Dieser besteht auf unterschiedlichen Ebenen.

Auf individueller Ebene gibt es bereits im vorschulischen Bereich Unterschiede zwischen schulisch relevanten Kompetenzen von Kindern ohne und mit Zuwanderungshintergrund. Dabei sind insbesondere im Vergleich geringere sprachliche Kompetenzen hervorzuheben. Diese auf sprachliche Kompetenzen bezogenen Nachteile bleiben in allen Bildungsetappen bestehen und erweisen sich als nachteilig für schulisches Lernen, Kompetenzerwerb und Bildungsbeteiligung (Kempert et al., 2016). Da Kinder mit Zuwanderungshintergrund, die das Deutsche als frühe Zweitsprache erwerben, einige sprachliche Phänomene deutlich später erwerben als einsprachige Kinder, argumentiert Schulz (2013), dass sprachliche Anforderungen erstens nicht an einsprachigen Kin-

© Waxmann Verlag GmbH

dern ausgerichtet werden sollten. Zweitens sollte ein früher systematischer Sprachkontakt hergestellt werden, um sprachbiografisch bedingte sprachliche Rückstände in der Unterrichtssprache bis zum Schuleintritt so weit wie möglich aufzuholen (Grimm & Schulz, 2014). Frühe individuelle Förderung der Sprache des Einwanderungslandes ist deshalb ein zentrales Desiderat, um eine Benachteiligung von Kindern und Jugendlichen mit Zuwanderungshintergrund zu reduzieren. Dabei sind sowohl ein früher Beginn wie auch die Dauer des Umgangs mit der Unterrichtssprache Deutsch entscheidend (Grimm & Schulz, 2014). Auch die Dauer des Besuchs eines Kindergartens ist ein wesentlicher Faktor. So konnten Becker und Biedinger (2006) zeigen, dass diese Dauer in einem signifikanten Zusammenhang zur Schulfähigkeit von Kindern mit Zuwanderungshintergrund steht. Zwar unterscheiden sich Kinder mit und ohne Zuwanderungshintergrund inzwischen nicht mehr, was die Frage angeht, ob überhaupt ein Kindergarten besucht wird, jedoch sind Kinder mit Zuwanderungshintergrund seltener in Einrichtungen mit guter Qualität zu finden. Zudem finden sich auch im Kindergarten schon Anzeichen für Segregation: Kinder mit Zuwanderungshintergrund besuchen häufiger Einrichtungen mit einem hohen Zuwandereranteil (vgl. Becker & Biedinger, 2016).

In Hinblick auf die Zusammensetzung von Klassen stellt vor allem die starke Segregation ein Problem dar. Dabei geht es nach der gegenwärtigen Forschungslage weniger um die Komposition hinsichtlich des Anteils von Schülerinnen und Schülern mit Zuwanderungshintergrund als vielmehr um die soziale Zusammensetzung der Schülerschaft und die mittlere Leistung (Kristen, 2008; Portes & Hao, 2004; Rjosk et al., 2014; Walter & Stanat, 2008). In längsschnittlich angelegten Studien zeigte sich kein eigenständiger Einfluss des Zuwandereranteils auf den Kompetenzerwerb (Stanat, Schwippert & Gröhlich, 2010). Eine stärkere Durchmischung der Schülerschaft in Hinblick auf soziale Herkunft könnte demnach zu einem besseren Lernen für alle führen.

Auf Ebene des Bildungssystems ist eine relevante Frage, ob Kinder gemeinsame Schulen besuchen sollten oder nach Leistung in verschiedene Schulformen eingeteilt werden. Der internationale Forschungsstand zum Effekt dieser Stratifikation auf schulisch erworbene Kompetenzen ist uneinheitlich. Der sozioökonomische Status der Eltern, der bei zugewanderten Familien in Deutschland deutlich geringer ist, scheint insbesondere in Schulsystemen zum Tragen zu kommen, in denen Kinder vergleichsweise früh auf unterschiedliche Schulformen verteilt werden (Hanushek & Wößmann, 2006). Für Kinder mit Zuwanderungshintergrund ist der Zusammenhang zwischen Stratifikation und Kompetenzerwerb besonders komplex (vgl. Teltemann, 2015). So zeigten Dronkers, Levels und de Heus (2014) auf Basis von internationalen PISA-Daten der Erhebung 2006, dass Kinder mit Zuwanderungshintergrund, deren Eltern über eine geringe formale Bildung verfügen, eher von Systemen mit einer mittleren Stratifizierung profitieren, während solche, deren Eltern über eine hohe formale Bildung verfügen, in stratifizierten Systemen schlechter abschneiden.

Nicht nur die Stratifikation, sondern auch der Übergang auf weiterführende Schulen stellt ein weiteres entscheidendes Element dar. Forschungssynthesen zeigen, dass der Übergang von der Grundschule auf ein Gymnasium Kindern mit Zuwanderungs-

© Waxmann Verlag GmbH

hintergrund unter Kontrolle der Leistungen und der sozialen Herkunft durchaus gelingt (Diehl, Hunkler & Kristen, 2016). Disparitäten in allen Kompetenzmaßen, insbesondere in sprachlichen Kompetenzen, bestehen aber schon vor dem Übergang auf die weiterführende Schule, sodass ein Übertritt auf das Gymnasium dadurch unwahrscheinlicher wird. Hier schließt sich der Kreis zur Notwendigkeit individueller Förderung in den vorgeschalteten Bildungsinstitutionen. Es ist notwendig dass alle, die mit Kindern zu tun haben, sprachliche Bildung als eine Selbstverständlichkeit annehmen, damit langfristig Disparitäten in allen Kompetenzbereichen reduziert werden können.

Literatur

Agirdag O. & Vanlaar G. (2017). Does more exposure to the language of instruction lead to higher academic achievement? A cross-national examination. *International Journal of Bilingualism*, http://dx.doi.org/10.1177/1367006916658711

Alba, R. D., Handl, J. & Müller, W. (1994). Ethnische Ungleichheit im deutschen Bildungssystem. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 46* (2), 209–237.

Autorengruppe Bildungsberichterstattung. (2016). *Bildung 2016 in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. Bielefeld: Bertelsmann. http://dx.doi.org/10.3278/6001820ew

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.

Becker, R. (2009). Entstehung und Reproduktion dauerhafter Bildungsungleichheiten, In R. Becker: *Lehrbuch der Bildungssoziologie* (S. 85-129). Wiesbaden: VS Verlag für Sozialwissenschaften.

Becker, R. & Beck, M. (2009). Migration, Sprachförderung und soziale Integration. Eine Evaluation der Sprachförderung von Berliner Schulkindern mit Migrationshintergrund. In R. Becker (Hrsg.), *Lehrbuch der Bildungssoziologie* (S. 121–138). Wiesbaden: VS Verlag für Sozialwissenschaften.

Becker, B. & Biedinger, N. (2006). Ethnische Bildungsungleichheit zu Schulbeginn. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 58* (4), 660–684.

Becker, B. & Biedinger, N. (2016). Ethnische Ungleichheiten in der vorschulischen Bildung. In C. Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten* (S. 433–474). Wiesbaden: Springer VS.

Brighouse, H. (2003). Educational equality and justice. In R. Curren (Hrsg.), *A companion to the philosophy of education* (S. 471–486). Oxford: Blackwell.

Bundesagentur für Arbeit. (2016). *Aktueller Bericht.* Zugriff am 17.10.2016. Verfügbar unter https://statistik.arbeitsagentur.de/Statischer-Content/Statistische-Analysen/Statistische-Sonderberichte/Generische-Publikationen/Auswirkungen-der-Migration-auf-den-Arbeitsmarkt.pdf

Damelang, A. & Haas, A. (2006). Arbeitsmarkteinstieg nach dualer Berufsausbildung – Migranten und Deutsche im Vergleich. *IAB Forschungsbericht, 17* (2006), 1–48. Zugriff am 17.10.2016. Verfügbar unter http://doku.iab.de/forschungsbericht/2006/fb1706.pdf

© Waxmann Verlag GmbH

Diehl, C., Hunkler, C. & Kristen, C. (2016). Ethnische Bildungsungleichheiten im Bildungsverlauf. Eine Einführung. In C. Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf. Mechanismen, Befunde, Debatten* (S. 3–31). Wiesbaden: Springer VS.

Dronkers, J., Levels, M. & de Heus, M. (2014). Migrant pupils' scientific performance: the influence of educational system features of origin and destination countries. *Large-scale Assessments in Education, 2.* http://dx.doi.org/10.1186/2196-0739-2-3

Dworkin, R. (2000). *Sovereign virtue: The theory and practice of equality*. Cambridge: Harvard University Press.

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

Frank, C., Härtig, H. & Neumann, K. (im Druck). Schulisch erworbene Kompetenzen als Voraussetzung für berufliches Wissen gewerblich-technischer Auszubildender. *Unterrichtswissenschaft*.

Ganzeboom, H. B. G., de Graaf, P. M., Treiman, D. J. & de Leeuw, J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 21* (1), 1–56.

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.

Gogolin, I. (1994). *Der monolinguale Habitus der multilingualen Schule*. Münster: Waxmann.

Gomolla, M. & Radtke, F.-O. (2002). *Institutionelle Diskriminierung. Die Herstellung ethnischer Differenz in der Schule.* Opladen: Leske + Budrich.

Grimm, A. & Schulz, P. (2014). Sprachfähigkeiten von Kindern mit DaZ bei Schuleintritt. In B. Lütke & I. Petersen (Hrsg.), *Deutsch als Zweitsprache: erwerben, lernen und lehren. Beiträge aus dem 9. Workshop ‚Kinder mit Migrationshintergrund'* (S. 35–50). Stuttgart: Klett Fillibach.

Hanushek, E. A. & Wössmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *Economic Journal*, *116* (510), C63-C76.

Härtig, H., Bernholt, S., Prechtl, H. & Retelsdorf, J. (2015). Unterrichtssprache im Fachunterricht – Stand der Forschung und Forschungsperspektiven am Beispiel des Textverständnisses. *Zeitschrift für Didaktik der Naturwissenschaften, 21,* 55-67. http://dx.doi.org/10.1007/s40573-015-0027-7

Härtig, H., Heitmann, P. & Retelsdorf, J. (2015). Analyse der Aufgaben zur Evaluation der Bildungsstandards in Physik – Differenzierung von schriftsprachlichen Fähigkeiten und Fachlichkeit. *Zeitschrift für Erziehungswissenschaft, 18* (4), 763–779. http://dx.doi.org/10.1007/s11618-015-0646-2

Kempert, S., Edele, A., Rauch, D., Wolf, K., M., Paetsch, J., Darsow, A., Maluch, J. & Stanat, P. (2015). Die Rolle der Sprache bei der Entstehung, Aufrechterhaltung und Kompensation ethnischer Bildungsungleichheiten. In C. Diehl, C., Hunkler, C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten* (S. 157–241). Wiesbaden: Springer VS.

© Waxmann Verlag GmbH

Kiesel, D. (1996). *Das Dilemma der Differenz. Zur Kritik des Kulturalismus in der Interkulturellen Pädagogik*. Frankfurt a.M.: Cooperative Verlag.

Kristen, C. (2008). Schulische Leistungen von Kindern aus türkischen Familien am Ende der Grundschulzeit. Befunde aus der IGLU-Studie. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 48*, 230–251.

KMK (2005a) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005a). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004.* München: Luchterhand.

KMK (2005b) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005b). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004.* München: Luchterhand.

KMK (2005c) = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005c). *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss. Beschluss vom 16.12.2004.* München: Luchterhand.

Maehler, D. B., Teltemann, J., Rauch, D. P., Hachfeld, A., Kroth, A. & Stanat, P. (2015). Interdisziplinäre Operationalisierung des Migrationshintergrunds. In H. U. Brinkmann & D. B. Maehler (Hrsg.), *Methoden der Migrations- und Integrationsforschung – Lehrbuch und Forschungsleitfaden* (S. 263–282). Heidelberg: Springer. http://dx.doi.org/10.1007/978-3-658-10394-1_9

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* (5), 701–722.

OECD. (2013). *International migration outlook*. Paris: OECD.

OECD. (2016). *PISA 2015 results: Student performance in science, reading and mathematics. Volume 1.* Paris: OECD.

Portes, A. & Hao, L. (2004). The schooling of children of immigrants. Contextual effects on the educational attainment of the second generation. *Proceedings of the National Academy of Science, 101*, 11920–11927.

Rawls, J. (1999). *A theory of justice* (überarb. Aufl.). Cambridge: Belknap Press of Harvard University Press.

Rjosk, C., Richter, D., Hochweber, J., Lüdtke, O., Klieme, E. & Stanat, P. (2014). Socioeconomic and language minority classroom composition and individual reading achievement: The mediating role of instructional quality. *Learning and Instruction, 32,* 63–72.

Sachverständigenrat deutscher Stiftungen für Integration und Migration. (2015). *Jahresgutachten 2015. Unter Einwanderungsländern: Deutschland im internationalen Vergleich.* Berlin: Königsdruck.

Schulz, P. (2013). Wer versteht wann was? Sprachverstehen im frühen Zweitspracherwerb des Deutschen am Beispiel der w-Fragen. In A. Deppermann (Hrsg.), *Das Deutsch der Migranten. Jahrbuch 2012 des Instituts für deutsche Sprache Mannheim* (S. 313–338). Mannheim: IDS.

Seebaß, K. & Siegert, M. (2011). *Migranten am Arbeitsmarkt in Deutschland.* Nürnberg: Bundesamt für Migration und Flüchtlinge. Zugriff am 17.10.2016. Verfügbar unter https://www.bamf.de/SharedDocs/Anlagen/DE/Publikationen/WorkingPapers/wp36-migranten-am-arbeitsmarkt-in-deutschland.pdf?__blob=publicationFile

Segeritz, M., Walter, O. & Stanat, P. (2010). Muster des schulischen Erfolgs von jugendlichen Migranten in Deutschland: Evidenz für segmentierte Assimilation? *Kölner Zeitschrift für*

© Waxmann Verlag GmbH

Soziologie und Sozialpsychologie, 62 (1), 113–138. http://dx.doi.org/10.1007/s11577-010-0094-1

Stanat, P. (2008). Heranwachsende mit Migrationshintergrund im deutschen Bildungswesen. In K. S. Cortina, J. Baumert, A. Leschinsky, K. U. Mayer & L. Trommer (Hrsg.), *Das Bildungswesen in der Bundesrepublik Deutschland* (S. 683–743). Reinbek: Rowohlt.

Stanat, P., Becker, M., Baumert, J., Lüdtke, O. & Eckhardt, A. G. (2012). Improving second language skills of immigrant students: A field trial study evaluating the effects of a summer learning program. *Learning and Instruction, 22*, 159–170.

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, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 200–228). Münster: Waxmann.

Stanat, P., Schwippert, K. & Gröhlich, C. (2010). Der Einfluss des Migrantenanteils in Schulklassen auf den Kompetenzerwerb: Längsschnittliche Überprüfung eines umstrittenen Effekts. In C. P. Allemann-Ghionda, P. Stanat, K. Göbel & C. Röhner (Hrsg.), *Migration, Sprache, Identität. 55. Beiheft der Zeitschrift für Pädagogik*, 147–164.

Statistisches Bundesamt. (2001). *Bildung und Kultur* (Fachserie 11, Reihe 1, Allgemeinbildende Schulen, Schuljahr 1999/2000). Stuttgart: Metzler-Poeschel.

Statistisches Bundesamt. (2007). *Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund* (Fachserie 1, Reihe 2.2, Ergebnisse des Mikrozensus 2005).

Stojanov, K. (2015): Educational justice as respect egalitarianism. In *Annual conference 2015 of the Society for Philosophy of Education of Great Britain*, 26–29 March 2015, Oxford.

Stošić, P. (2016). Kinder mit ‚Migrationshintergrund' – Reflexionen einer (erziehungs-)wissenschaftlichen Differenzkategorie. In I. Diehm, M. Kuhn & C. Machold (Hrsg.), *Differenz – Ungleichheit – Erziehungswissenschaft. Verhältnisbestimmungen im (Inter-)Disziplinären* (S. 81–102). Wiesbaden: Springer VS.

Teltemann, J. (2015): *Ungleichheit als System? Der Schulerfolg von Migranten im internationalen Vergleich*. Frankfurt a.M.: Campus.

Vieluf, S. (im Druck). Ethnicity and Migration. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing contexts of learning: An international perspective*. Berlin: Springer.

Walter, O. & Stanat, P. (2008). Der Zusammenhang des Migrantenanteils in Schulen mit der Lesekompetenz. Differenzierte Analyse der erweiterten Migrantenstichprobe von PISA 2003. *Zeitschrift für Erziehungswissenschaft, 11*, 84–105.

© Waxmann Verlag GmbH

© Waxmann Verlag GmbH

10 Die Lernumgebung in der Familie und die elterliche Unterstützung

Nina Jude, Silke Hertel, Susanne Kuger & Christine Sälzer

Merkmale der häuslichen Lernumgebung, der frühen Förderung sowie der Zusammenarbeit der Eltern mit der Schule beschreiben Bedingungen, unter denen Fünfzehnjährige in Deutschland und anderen Staaten ihre Kompetenzen entwickeln können. PISA 2015 kann aufzeigen, dass die Voraussetzungen der Förderung im Elternhaus und die Zusammenarbeit zwischen Eltern und Schule in Deutschland durchaus positiv zu beurteilen sind.

Ein Großteil der fünfzehnjährigen Schülerinnen und Schüler wächst in einer Familie auf, in der gemeinsame Aktivitäten sowie Gespräche mit den Eltern täglich stattfinden und in der die Eltern ihr Kind emotional unterstützen. Eine spezifische fachliche Unterstützung – in PISA 2015 bezogen auf die Naturwissenschaften – findet hingegen vergleichsweise selten statt. Hinsichtlich der beruflichen Erwartungen formulieren die Eltern lediglich jeder und jedes zehnten Fünfzehnjährigen einen naturwissenschaftlichen Berufswunsch. Die Eltern der meisten Schülerinnen und Schüler sind motiviert, sich an der Schule ihres Kindes aktiv zu engagieren, es beim Lernen zu unterstützen und zu fördern. Sie nehmen an vielen Schulen eine hohe Schulqualität und eine zur Zusammenarbeit einladende Atmosphäre wahr. Dabei sehen die Eltern in der Zusammenarbeit mit der Schule einen sehr wichtigen und relevanten Aspekt für die Entwicklung ihres Kindes. Konkrete Hinweise zur Unterstützung der familiären Förderung durch die Schule erfolgen allerdings vergleichsweise selten. Hier wird Optimierungspotenzial deutlich: Schulen könnten noch aktiver auf Eltern zugehen, sie zur Zusammenarbeit einladen und neben umfangreichen Informationen auch konkrete Unterstützungsmöglichkeiten thematisieren.

Im Rahmen von PISA werden die Ergebnisse kumulierter Bildungsprozesse bei fünfzehnjährigen Schülerinnen und Schülern erfasst. Bildungsprozesse dieser Art werden wesentlich durch den institutionellen Kontext von Schule geprägt, aber auch dem außerschulischen Umfeld kommt eine bedeutende Rolle zu. Insbesondere leistet das häusliche Umfeld einen wichtigen Beitrag zum Bildungserfolg von Kindern und Jugendlichen (z. B. Maaz, Neumann & Baumert, 2014). PISA nimmt Indikatoren in den Blick, die diese Rolle beschreiben können. Aus theoretischen Rahmenmodellen zur elterlichen Beteiligung an schulischen Bildungsprozessen (vgl. Green, Walker, Hoover-Dempsey & Sandler, 2007; Hoover-Dempsey et al., 2005) wurden Indikatoren zu Rahmenbedin-

© Waxmann Verlag GmbH

gungen und Ansatzpunkten elterlicher Beteiligung abgeleitet. In Form von Kontextfragebögen greift PISA solche Indikatoren auf, um beispielsweise die Lernumgebung im Elternhaus und die Unterstützung der Eltern in Bezug auf schulische Lern- und Bildungsprozesse zu erfassen. Neben allgemeinen Indikatoren werden auch spezifische Unterstützungsangebote im Bereich der Naturwissenschaften sowie der frühkindlichen Förderung adressiert.

Der PISA-Elternfragebogen richtet sich an die Eltern und Erziehungsberechtigten aller Schülerinnen und Schüler, die an der Studie teilnehmen. In Deutschland ist die Befragung der Eltern seit 2006 ohne Unterbrechung Bestandteil der PISA-Erhebungen. Als internationale Option gibt es dieses Instrument ebenfalls seit 2006, bei PISA 2015 wurde es in 18 Teilnehmerstaaten eingesetzt.

In diesem Kapitel werden zunächst die theoretischen Grundlagen sowie wichtige Befunde zum elterlichen Engagement und zur Gestaltung der häuslichen Lernumgebung beschrieben. Darauf aufbauend werden die in PISA 2015 ausgewählten Indikatoren zur Erfassung dieser beiden Aspekte dargestellt. Nach einer Beschreibung der häuslichen Lernumgebungen sowie des elterlichen Engagements wird das Zusammenspiel dieser Aspekte mit der Kompetenz der Fünfzehnjährigen betrachtet. Abschließend wird diskutiert, welche Merkmale des Elternhauses bei der Interpretation der PISA-Ergebnisse von besonderer Relevanz sind und welches Potenzial sich für eine Unterstützung des Bildungserfolgs der Schülerinnen und Schüler ergibt.

10.1 Theoretische Rahmung und Indikatoren in PISA 2015

Elterliches Engagement in Bildungsprozessen

Eltern nehmen bei den Bildungsprozessen ihrer Kinder einen großen Stellenwert ein – sowohl bezogen auf die schulische Laufbahn als auch auf das außerschulische Lernen. Sie begleiten in der Regel ihr Kind bei den ersten Lernerfahrungen, sind maßgeblich an der Gestaltung der Lernumgebung im Elternhaus beteiligt, kooperieren mit pädagogischem Fachpersonal in der Kinderbetreuungseinrichtung sowie später mit Lehrkräften an Schulen und sind mit ihren Einstellungen, Überzeugungen und Interessen wichtige Vorbilder für ihr Kind. In diesem Sinne meint elterliches Engagement ein selbstmotiviertes, nicht durch akute Probleme angestoßenes Sich-Einbringen in den schulischen Bereich.

Bei einer Beschreibung des elterlichen Engagements für (schulische) Bildungsprozesse kann zunächst zwischen einer Beteiligung im Elternhaus selbst und einer Beteiligung an der Schule des Kindes unterschieden werden (vgl. Green et al., 2007; Hoover-Dempsey et al., 2005). Die *Beteiligung im Elternhaus* umfasst insbesondere Interaktionen, die sich auf lernrelevante Verhaltensweisen des Kindes, seine Einstellungen zum Lernen sowie Lern- und Arbeitsstrategien beziehen. Auch Hilfe bei den Hausaufgaben, die Unterstützung beim Lernen sowie die Beobachtung des Lernfortschrittes sind

© Waxmann Verlag GmbH

wichtige Bestandteile der häuslichen Lernumgebung. Die *Beteiligung und das Engagement an der Schule des Kindes* kann unterschiedliche Formen annehmen: Sie kann sich auf das Kind (z. B. Teilnahme an einem Elternabend oder einem Elterngespräch) oder auf eher allgemeine Aktivitäten an der Schule (z. B. Unterstützung bei einem Projekttag) beziehen (vgl. Green et al., 2007; Jude & Hertel, 2014).

Inwieweit, in welcher Form und mit welcher Qualität Eltern sich an diesen Bildungsprozessen beteiligen, hängt von unterschiedlichen Faktoren ab. Diese können Einstellungen, Überzeugungen oder auch die Motivation der Eltern (z. B. Rollenverständnis, Selbstwirksamkeit), der wahrgenommene Lebenskontext der Familie (u. a. Fähigkeiten und Wissen, Zeit und Energie) sowie die Wahrnehmung der Schule (bspw. Einladung zu Beteiligung durch Schulleitung oder Lehrkräfte) sein (vgl. Green et al., 2007). Im Folgenden werden die einzelnen Aspekte differenziert betrachtet, zudem werden die in PISA 2015 aufgenommenen Indikatoren dargestellt.

Häusliche Lernumgebung, Rahmenbedingungen und Ressourcen der Familien

Kinder verbringen neben der Zeit im Kindergarten und in der Schule einen Großteil ihrer Zeit in der Familie, weshalb diese auch als primäre Sozialisationsinstanz bezeichnet wird (Tillmann, 2010). Wissen, Gewohnheiten, Einstellungen und Werte werden wesentlich von den Eltern und Geschwistern vermittelt. Entsprechende Lernprozesse werden durch die Eltern teils bewusst und teils auch ohne spezifische Absicht angeregt. Die Familie ist damit der früheste und wichtigste Einflussfaktor in der kindlichen Entwicklung (McElvany & Becker, 2009).

In der einschlägigen Fachliteratur zur Bedeutung der familiären Lernumwelt wird zwischen sogenannten Struktur- und Prozessmerkmalen unterschieden. Zu den *Strukturmerkmalen* zählen insbesondere der sozioökonomische Status der Familie, das Bildungsniveau der Eltern sowie ein etwaiger Zuwanderungshintergrund. Diese Merkmale weisen Zusammenhänge zu den im Elternhaus verfügbaren Ressourcen für das Lernen auf (z. B. Lexika, Nachschlagewerke, vgl. Szczesny & Watermann, 2011). Bei der Beschreibung ihrer Wirkung beispielsweise auf den Bildungserfolg, die Schulwahl oder die Berufswahl lassen sich sogenannte primäre und sekundäre Herkunftseffekte unterscheiden. Primäre Herkunftseffekte (Boudon, 1974) bezeichnen Effekte der sozialen Schicht, die sich direkt auf die Förderung und Anregung durch die Eltern auswirken und in der Wertschätzung von Bildung ausdrücken können. Einflüsse der sozialen Herkunft, die sich auch unter Kontrolle der primären Herkunftsmerkmale zeigen, werden als sekundäre Herkunftsmerkmale bezeichnet (Maaz et al., 2014). Sie zeigen sich zum Beispiel in der Bedeutung der Herkunft für bildungsrelevante Entscheidungen wie den Übergang in die weiterführende Schule oder auch die individuelle Benotung. Hinsichtlich der *Prozessmerkmale* lassen sich unterstützende Aktivitäten der Eltern, welche die Emotionsregulation sowie die Motivation fördern, von kognitiv anregenden, lernprozessbezogenen Unterstützungsangeboten unterscheiden. Bezüglich der emotionalen und motivationalen Unterstützung lassen sich insbesondere die Unterstützung der Selbstwirksamkeit der Kinder sowie die

© Waxmann Verlag GmbH

Förderung eines konstruktiven Umgangs mit Erfolg und Misserfolg anführen. Hierfür ist die Fähigkeit, Emotionen wahrzunehmen, zu spiegeln und gegebenenfalls auch aufzunehmen und dem Kind in einer für es gut zu verarbeitenden Form zu präsentieren, von großer Bedeutung (Ainsworth, Bell & Stayton, 1974; Grossmann & Grossmann, 2012). Unterstützungsangebote, die sich auf die Erarbeitung und die Verarbeitung von Lerninhalten beziehen (vgl. Szczesny & Watermann, 2011), umfassen insbesondere elterliche Aktivitäten im Bereich des *Scaffoldings*. Zentrale Wirkmechanismen sind das Fokussieren (von Aufmerksamkeit auf zentrale Aspekte) und das Modellieren (eines Lösungswegs oder einer Lernstrategie). Eltern können ihrem Kind beim Lernen Hilfestellungen geben, indem sie beispielsweise Fragen stellen, Hinweise geben, Handlungen anleiten, Rückmeldungen geben, Erklärungen anbieten und Transfer anregen (Department of Health, 2000; van de Pol, Volmann & Beishuizen, 2010).

Im bildungspolitischen Diskurs wird gegenwärtig ein starker Fokus auf die ersten Lernerfahrungen von Kindern gelegt. In diesem Zusammenhang werden Lernumgebungen in der frühen Kindheit differenziert betrachtet, sowohl in der Familie als auch in formalen Settings der Kinderbetreuung (z. B. Kinderkrippe oder Kindergarten). Die frühen Bildungserfahrungen haben sich dabei als wichtiger Baustein späterer Kompetenzen in der Schule erwiesen (Lehrl, Kluczniok & Roßbach, 2016; Sylva et al., 2012). In PISA 2009 konnten positive Zusammenhänge der Lesekompetenz der fünfzehnjährigen Schülerinnen und Schüler mit einer frühen Förderung in der Familie gezeigt werden (Hertel, Jude & Naumann, 2010).

In den letzten Jahren wurde in Deutschland eine Zunahme an nicht familialen Betreuungsformen in der frühen Kindheit beobachtet (Autorengruppe Bildungsberichterstattung, 2016; Berngruber, Alt & Hubert, 2014; Schober & Spieß, 2014). Für die Beschreibung nicht familialer Lernumgebungen in der frühen Kindheit werden in der Literatur unterschiedliche Indikatoren angeführt wie etwa der Besuch einer entsprechenden Einrichtung, die Dauer und Intensität des Besuchs einer Kinderbetreuungseinrichtung, das Curriculum oder der pädagogische Ansatz der Einrichtung. Eine wichtige, aber sehr schwer zu erfassende Einflussgröße auf die kognitive, soziale und emotionale Entwicklung der Kinder ist die Qualität der Betreuung in der Einrichtung (Anders, Roßbach & Kuger, im Druck; Kluczniok & Roßbach, 2014).

Engagement der Eltern an der Schule ihres Kindes

Die Zusammenarbeit zwischen Eltern und Schule ist im Prinzip ein fester Bestandteil des Schulalltags (Baumert & Kunter, 2011; Bruder, Hertel, Gerich & Schmitz, 2014; Freyaldenhofen, 2005). Ob und in welchem Umfang Eltern sich an der Schule ihres Kindes engagieren, hängt allerdings von ihren Einstellungen, Überzeugungen und von ihrer Motivation ab. Darüber hinaus können auch eine vertrauensvolle Atmosphäre an der Schule oder direkte Einladungen der Eltern zur Beteiligung dazu beitragen, dass sie aktiv werden. Eine wichtige Rolle nehmen dabei die Lehrerinnen und Lehrer ein. Sie können Eltern dazu motivieren, sich an der Schule ihres Kindes zu engagieren und

© Waxmann Verlag GmbH

mit der Schule zusammenzuarbeiten (Corno, 2000; Hoover-Dempsey, Bassler & Burow, 1995; Hoover-Dempsey & Sandler, 1997; Kohl, Lengua & McMahon, 2000). Auch Reformen des Bildungssystems, beispielsweise die Einführung von Ganztagsschulen oder die Umsetzung inklusiver Schulkonzepte, können zu neuen Möglichkeiten der Beteiligung von Eltern an der Schule ihres Kindes führen, aber auch Herausforderungen darstellen. Es gibt Belege, dass Eltern heute stärker als noch vor einigen Jahren aktiv in die Entscheidungen zur schulischen Laufbahn ihrer Kinder einbezogen sind (vgl. Hertel, 2016).

Eine gelingende Zusammenarbeit zwischen Schule und Familie geht mit vielfältigen positiven Effekten einher, etwa mit einer Verbesserung der schulischen Leistungen der Kinder oder einer Reduktion von Verhaltensauffälligkeiten (z. B. Catsambis, 1998; Cox, 2005; Hill et al., 2004; Jeynes, 2007). So gibt es Hinweise darauf, dass der sozioökonomische und der kulturelle Hintergrund der Familien einen Einfluss auf die Wirkung des elterlichen Engagements an der Schule ihres Kindes nehmen. Die elterliche Beteiligung an den schulischen Bildungsprozessen ihres Kindes kann insbesondere in Familien mit niedrigem sozioökonomischen Hintergrund sowie in Familien mit Zuwanderungshintergrund besonders wirkungsvoll sein, sie stellt aber gleichzeitig für diese Elterngruppe eine große Herausforderung dar (vgl. Hill et al., 2004; Hertel, 2016; Hertel, Bruder, Jude & Steinert, 2013).

10.2 Erfassung der häuslichen Lernumgebung sowie der elterlichen Unterstützung in PISA 2015

Erfassung zentraler Indikatoren mit dem Elternfragebogen

Die Fragebögen in PISA 2015 basieren auf der internationalen Rahmenkonzeption, die in Kapitel 13 ausführlich dargestellt wird. Von den 19 dort definierten, bildungspolitisch relevanten Aspekten werden im Elternfragebogen sechs spezifisch in den Blick genommen: die sozioökonomische und die kulturelle Herkunft der Familie, die Lernumgebung in der Familie allgemein und mit einem Fokus auf die Naturwissenschaften, die elterliche Zusammenarbeit mit der Schule sowie die frühkindliche Bildung.

Die *häusliche Lernumgebung* der Fünfzehnjährigen wurde in PISA 2015 durch Indikatoren erfasst, die *Strukturmerkmale* der Unterstützung und Rahmenbedingungen in den Familien abbilden. Hierzu gehören die Breite des Angebots an Schulen, Möglichkeiten der frühkindlichen Betreuung, die sozioökonomische Situation sowie Bildungsressourcen in der Familie. Zu den *Prozessmerkmalen* zählen die Förderaktivitäten im Elternhaus sowie das Engagement der Eltern in der Schule. Zusätzlich spielen Einstellungen, Überzeugungen, Motive und Interessen der Eltern bezogen auf die Beteiligung an den Bildungsprozessen ihres Kindes, die Zusammenarbeit mit der Schule und die Naturwissenschaften eine Rolle.

© Waxmann Verlag GmbH

Die Erfassung von Bildungserfahrungen in der frühen Kindheit in PISA ist besonders herausfordernd, da diese Lebensphase zum einen schon zehn bis fünfzehn Jahre zurückliegt und zum anderen die Schülerinnen und Schüler selbst kaum Erinnerungen an diese frühe Zeit haben (Anders et al., im Druck). Daher erhebt der Elternfragebogen Indikatoren, die mit einer hohen Wahrscheinlichkeit zum Zeitpunkt der PISA-Befragung noch verlässlich erinnert werden können. Dazu zählen der *Umfang der Fremdbetreuung* der Kinder (Stunden pro Woche) und der *Grund für die Entscheidung*, eine außerfamiliale Betreuung zu wählen. Wichtige Informationen sind zudem der Fokus der besuchten Einrichtung(en), wie beispielsweise spezifische Schulvorbereitung oder allgemeine Förderung sowie in welchem Rahmen und durch wen die Betreuung stattfand (bei einer Tagesmutter oder in einer größeren Einrichtung).

Stichprobe

Die Teilnahme an der PISA-Elternbefragung ist in Deutschland freiwillig. Die Befragung richtet sich an alle Eltern und Erziehungsberechtigten der für die Teilnahme an der PISA-Erhebung ausgewählten Schülerinnen und Schüler. Diesen wird der Fragebogen in Papierform vor dem Testtag mit nach Hause gegeben. Die Fragebögen werden am Testtag wieder eingesammelt. Darüber hinaus haben die Eltern und Erziehungsberechtigten die Möglichkeit, die ausgefüllten Fragebögen in einem vorfrankierten Rückumschlag direkt an die Testleitung zu schicken.

Die Rücklaufquote der Fragebögen betrug in PISA 2015 für Deutschland 52 Prozent. Es ist davon auszugehen, dass die Teilnahmebereitschaft auch durch Hintergrundmerkmale der Familie beeinflusst wird, in PISA ist dies insbesondere der Zuwanderungshintergrund (siehe Hertel et al., 2010). Die Repräsentativität der Befunde auf der Grundlage des Elternfragebogens ist daher deutlich eingeschränkt und dies muss bei der Interpretation der Ergebnisse berücksichtigt werden.

In PISA 2015 wurde der Elternfragebogen wie auch in früheren Erhebungsrunden ausschließlich in deutscher Sprache ausgegeben, daher ist anzunehmen, dass die Beantwortung des Fragebogens für Eltern mit Zuwanderungshintergrund eine besondere Herausforderung darstellt. Differenzierte Analysen der Rücklaufquote zeigen, dass die Antwortquote der Eltern bei Schülerinnen und Schülern ohne Zuwanderungshintergrund mit 63 Prozent deutlich über der von Eltern der Schülerinnen und Schüler mit Zuwanderungshintergrund (1. Generation, 39 Prozent) liegt. Bei drei Viertel der Schülerinnen und Schüler wurde der Elternfragbogen von den Müttern bzw. einer weiblichen Erziehungsberechtigten ausgefüllt (76 Prozent).

Methodische und technische Aspekte

Bei der Auswertung und der Interpretation der Ergebnisse des Elternfragebogens sind drei Aspekte besonders zu berücksichtigen: Erstens ist ein umfassender internationaler Vergleich nur eingeschränkt möglich, da nicht alle Teilnehmerstaaten diesen Fra-

© Waxmann Verlag GmbH

gebogen einsetzen. Zweitens ist aufgrund der Freiwilligkeit der Teilnahme mit einem eingeschränkten Rücklauf zu rechnen. Analysen aus dieser sowie aus vorangegangen PISA-Erhebungsrunden weisen zudem darauf hin, dass der Ausfall systematisch mit familiären Hintergrundmerkmalen (sozioökonomischer Status sowie Migrationshintergrund) zusammenhängt. Drittens ist zu beachten, dass die PISA-Studie als Querschnittserhebung angelegt ist. Die untersuchte Kohorte wird nur einmalig am Ende der Pflichtschulzeit befragt. Dieses Design erlaubt allenfalls Aussagen über (korrelative) Zusammenhänge unterschiedlicher Indikatoren zum Zeitpunkt der Datenerhebung. Aussagen über (kausale) Ursache-Wirkungszusammenhänge sind auf der Ebene von Schülerinnen und Schülern sowie der Eltern nicht möglich. Mit entsprechender Vorsicht sind allerdings Schlussfolgerungen über die Veränderung von Indikatoren von einem Zeitpunkt zum nächsten auf Ebene des Bildungssystems möglich (vgl. Kapitel 11).

Da die Auswahl der Eltern zur Teilnahme am Fragebogen auf der Selektion der Schülerinnen und Schüler basiert (vgl. Kapitel 12), werden für die nachfolgenden Analysen die Daten des Elternfragebogens mit dem Schülergewicht gewichtet. Prozentuale Angaben beziehen sich demnach immer auf den Anteil der fünfzehnjährigen Schülerinnen und Schüler, die in Familien aufwachsen, in denen entsprechende Lernumgebungen vorliegen bzw. die Eltern auf eine bestimmte Art und Weise unterstützen oder sich an der Schule des Kindes engagieren.

Im Folgenden werden häusliche Lernumgebungen der Fünfzehnjährigen in Deutschland und die elterliche Beteiligung an (schulischen) Bildungsprozessen ihrer Kinder deskriptiv dargestellt. Deskriptive Vergleiche mit der Erhebung PISA 2006, bei der ebenfalls Naturwissenschaften im Fokus waren, werden für diejenigen Indikatoren vorgenommen, die zu beiden Zeitpunkten erhoben wurden. Darüber hinaus werden Zusammenhänge zwischen Faktoren der häuslichen Lernumgebung und der elterlichen Unterstützung mit der Kompetenz der Schülerinnen und Schüler in den Naturwissenschaften analysiert.

10.3 Die häusliche Lernumgebung von Fünfzehnjährigen aus Sicht der Eltern und das elterliche Engagement an der Schule ihres Kindes

Häusliche Lernumgebungen und Rahmenbedingungen in den Familien

In diesem Abschnitt wird der Fokus auf die Beschreibung der Prozessmerkmale in der Unterstützung der Eltern sowie auf die frühen Bildungserfahrungen der Schülerinnen und Schüler gelegt. Zunächst wurden die gemeinsamen Aktivitäten im Elternhaus erfragt, dabei wurden sowohl allgemeine soziale Aktivitäten einbezogen als auch solche, die sich spezifisch auf die Schule oder die Naturwissenschaften im Allgemeinen beziehen (siehe Abbildung 10.1). Die befragten Eltern konnten dabei angeben, wie häufig in ihrer Familie die jeweiligen Aktivitäten stattfinden.

© Waxmann Verlag GmbH

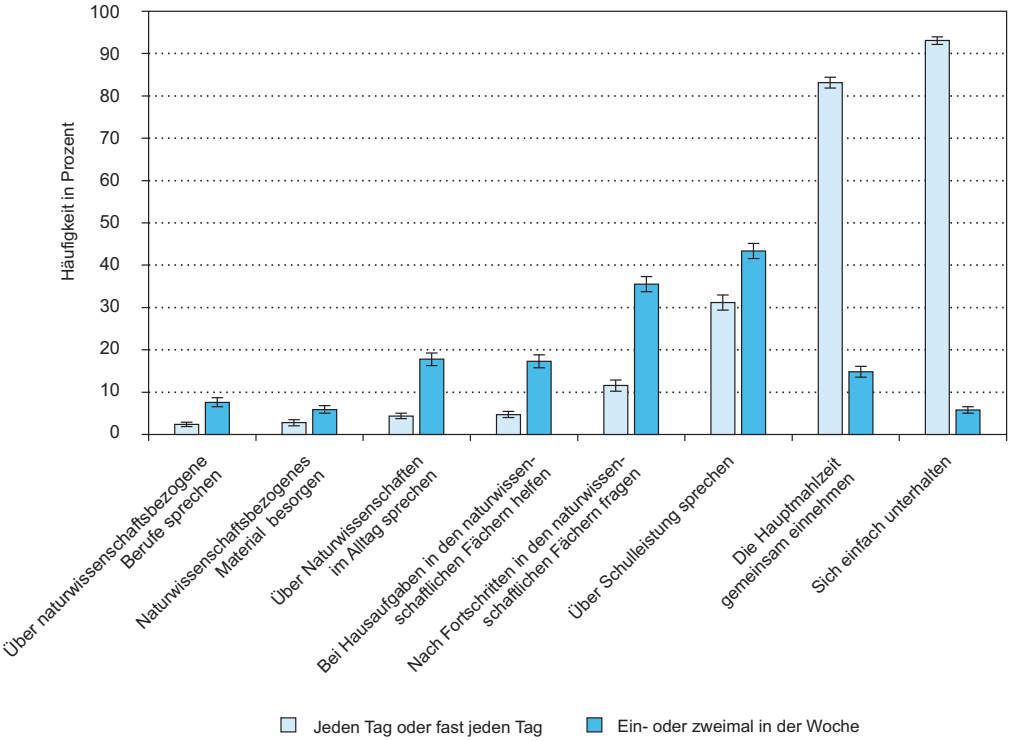


Abbildung 10.1: Allgemeine und schulbezogene Aktivitäten und Interaktionen in der Familie

Die von den Eltern berichteten Interaktionen zwischen ihnen und ihrem Kind haben meist keinen direkten Bezug zur Schule oder zum Lernen. Ein großer Teil der Fünfzehnjährigen nimmt laut Auskunft der Eltern täglich oder fast täglich die Hauptmahlzeit mit den Eltern ein und unterhält sich mit ihnen. Auch Gespräche mit einem Bezug zur Schule beziehungsweise zu schulischen Themen finden statt, jedoch bereits seltener. Knapp die Hälfte der Eltern fünfzehnjähriger Schülerinnen und Schüler führt einmal oder zweimal pro Woche mit ihnen Gespräche über die schulischen Leistungen. Spezifische Unterstützung oder Förderung im naturwissenschaftlichen Bereich kommt noch seltener vor. Nur 10 Prozent der Fünfzehnjährigen wachsen in Familien auf, in denen regelmäßig über naturwissenschaftliche Berufswege gesprochen wird, allerdings geben 22 Prozent der teilnehmenden Eltern an, häufig mit ihrem Kind darüber zu sprechen, wie Naturwissenschaften im Alltag angewendet werden.

Zusätzlich zu auf schulische Erfolge und Leistung ausgerichteten Unterstützungsaktivitäten der Eltern wurde in PISA 2015 erstmalig auch die emotionale Unterstützung der Schülerinnen und Schüler durch ihre Familie erfasst. Die Eltern wurden gefragt, wie sehr sie ihr Kind emotional unterstützen und Interesse an den schulischen Aktivitäten ihres Kindes zeigen (Abbildung 10.2).

© Waxmann Verlag GmbH

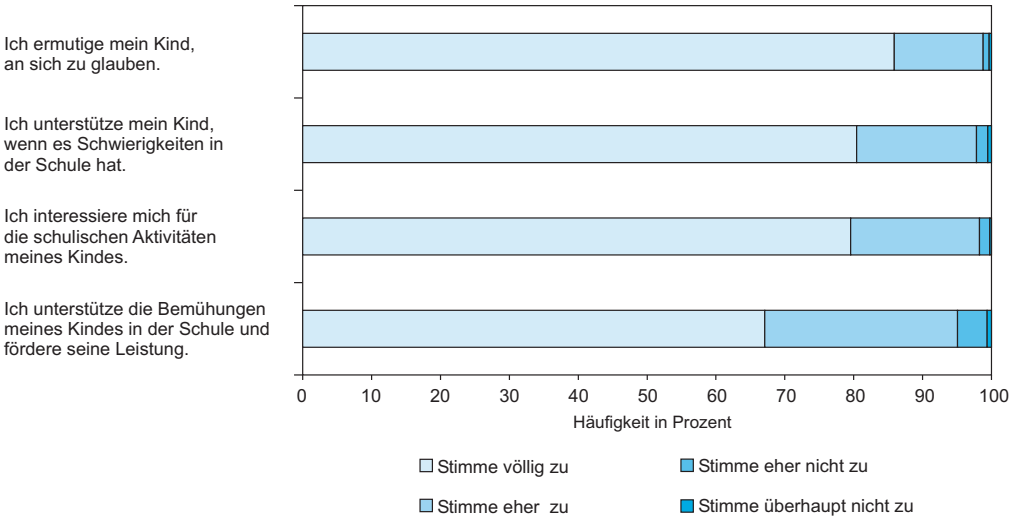


Abbildung 10.2: Emotionale Unterstützung der Schülerinnen und Schüler durch ihre Eltern

Ein Großteil der fünfzehnjährigen Schülerinnen und Schüler wird von den Eltern emotional unterstützt. Konkret geben diese Eltern an, ihr Kind zu ermutigen, an sich zu glauben, bei Schwierigkeiten für es da zu sein und sich für die schulischen Aktivitäten zu interessieren und es zu fördern. Aufgrund der hohen Zustimmungsraten sind diese Ergebnisse allerdings mit Vorsicht zu betrachten. Deckeneffekte oder sozial erwünschtes Antwortverhalten können nicht ausgeschlossen werden.

Einen wichtigen Stellenwert nehmen – wie einleitend dargestellt – die frühen Bildungserfahrungen der Schülerinnen und Schüler ein. Die Eltern wurden gebeten, Angaben über ihre Aktivitäten bei der Förderung und Unterstützung ihres Kindes während der Grundschulzeit sowie im Kleinkindalter zu machen. Hinsichtlich der naturwissenschaftlichen Förderung in der Familie während der Grundschulzeit legen die Ergebnisse nahe, dass eine Unterstützung im naturwissenschaftlichen Bereich während Grundschulzeit eher selten erfolgte (Abbildung 10.3).

Am häufigsten fand eine entsprechende Förderung durch gemeinsame Bauspiele statt. Etwa 60 Prozent der Fünfzehnjährigen haben laut Aussagen ihrer Eltern solche Spiele sehr oft beziehungsweise regelmäßig gemacht. Ein weiterer, von den Eltern etwa eines Drittel der fünfzehnjährigen Schülerinnen und Schüler gewählter Zugang, ist das gemeinsame Anschauen von Fernsehsendungen mit naturwissenschaftlichem Inhalt (sehr oft bzw. regelmäßig: 28 Prozent). Der gemeinsame Besuch einer naturwissenschaftlichen Arbeitsgruppe war hingegen vergleichsweise selten (sehr oft bzw. regelmäßig: 3 Prozent).

© Waxmann Verlag GmbH

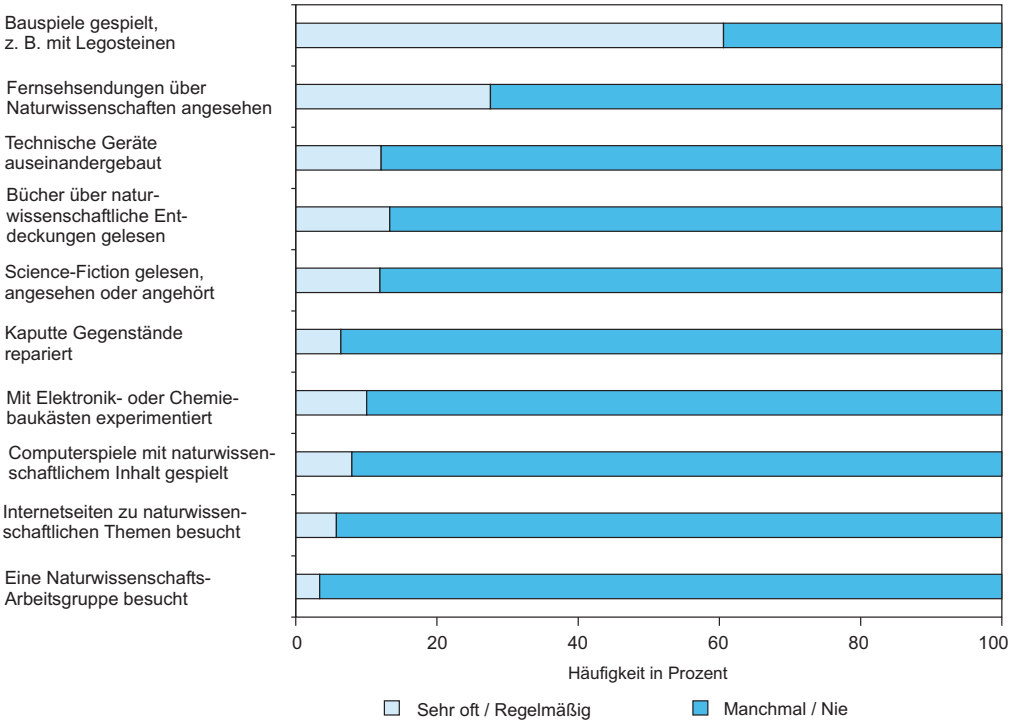


Abbildung 10.3: Naturwissenschaftliche Förderung in der Familie während der Grundschulzeit

Institutionelle Bildungserfahrungen in der frühen Kindheit

Die überwiegende Mehrheit der an PISA 2015 teilnehmenden Schülerinnen und Schüler – die im Jahr 1999 geboren wurden –, besuchten eine außerfamiliale Betreuungseinrichtung. Es lassen sich leichte Unterschiede in den Besuchsquoten von Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund in Einrichtungen der Kleinkindbetreuung (z. B. Tagesmutter, Kindertagesstätte, Kinderhort und Kinderkrippe) sowie in Einrichtungen der frühkindlichen Lernentwicklung (bspw. Frühförderung, Früherziehung) ausmachen. Für die Besuchsquoten von Einrichtungen für Vorschulerziehung (z. B. Schulkindergarten und Vorschule) ergeben sich keine Unterschiede für die Fünfzehnjährigen mit und ohne Zuwanderungshintergrund (vgl. Abbildung 10.4).

Nur 8 Prozent der fünfzehnjährigen Schülerinnen und Schüler haben nach Angabe ihrer Eltern keine dieser drei Einrichtungsarten besucht, demgegenüber besuchten 7 Prozent Einrichtungen aller drei Arten. Für die Mehrheit der fünfzehnjährigen Schülerinnen und Schüler (60 Prozent) wurde der Besuch einer Einrichtungsart bestätigt und für 23 Prozent der Besuch zweier Einrichtungsarten. Der Fragebogen enthielt keine Frage dazu, wie viele einzelne Einrichtungen eine Schülerin beziehungsweise ein Schüler in der Kindheit insgesamt besucht hat.

© Waxmann Verlag GmbH

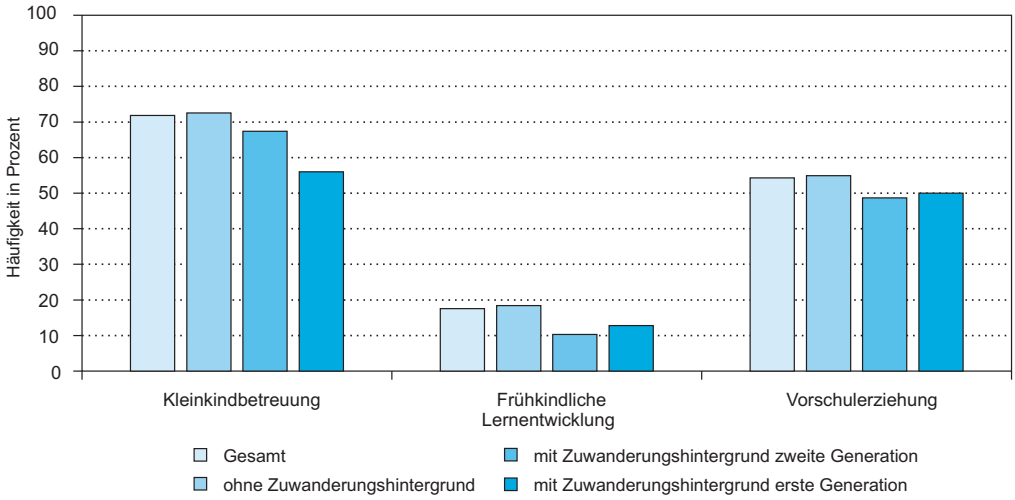


Abbildung 10.4 Einrichtungen zur frühkindlichen Betreuung und Erziehung

Die Fünfzehnjährigen selbst gaben Auskunft dazu, von welchem Alter an sie einen Kindergarten besucht haben und wann sie in die Grundschule eingeschult wurden (vgl. Abbildung 10.5).

Jungen waren im Schnitt beim Eintritt in einen Kindergarten (3;4 Jahre) und bei ihrer Einschulung (6;2 Jahre) etwas älter als Mädchen (3;2 Jahre bzw. 6;1 Jahre), der Altersunterschied ist aber statistisch nicht bedeutsam. Etwas anders sieht es bei Schülerinnen und Schülern mit Zuwanderungshintergrund aus. Zum einen haben sie eine insgesamt etwas niedrigere Besuchsquote des Kindergartens. Diejenigen, die einen Kindergarten besuchten, waren nach eigenen Angaben bei ihrem Eintritt zudem etwas älter als Schülerinnen und Schüler ohne Zuwanderungshintergrund. Schülerinnen und Schüler ohne Zuwanderungshintergrund waren beim Eintritt in den Kindergarten (bzw. Schuleintritt) im Durchschnitt 3;2 Jahre (6;1 Jahre) alt, Schülerinnen und Schüler mit einem Zuwanderungshintergrund zweiter Generation waren mit 3;6 Jahren schon durchschnittlich 4 Monate älter (6;3 Jahre, folglich zwei Monate älter bei Schuleintritt). Schülerinnen und Schüler mit Zuwanderungshintergrund erster Generation waren beim Eintritt in den Kindergarten 3;10 Jahre alt (6;4 Jahre) und damit beim Eintritt in den Kindergarten deutlich älter.

Die Gründe für den Besuch einer außerfamiliären Betreuung wurden mit dem Elternfragebogen erfasst. Die Eltern der überwiegenden Mehrheit der Schülerinnen und Schüler geben an, der Wunsch nach zusätzlichen Lernanreizen für das Kind wäre unabhängig von der besuchten Einrichtungsart der Hauptgrund dafür gewesen, ihr Kind außerfamilial betreuen zu lassen. Bei einem Drittel der fünfzehnjährigen Schülerinnen und Schüler (30 Prozent) geben die Eltern an, eine Einrichtung der Kleinkindbetreuung (z.B. Tagesmutter, Kinderkrippe) genutzt zu haben, weil sie sich aufgrund von Arbeit oder Krankheit nicht ausreichend selbst um das Kind kümmern konnten. Für knapp

© Waxmann Verlag GmbH

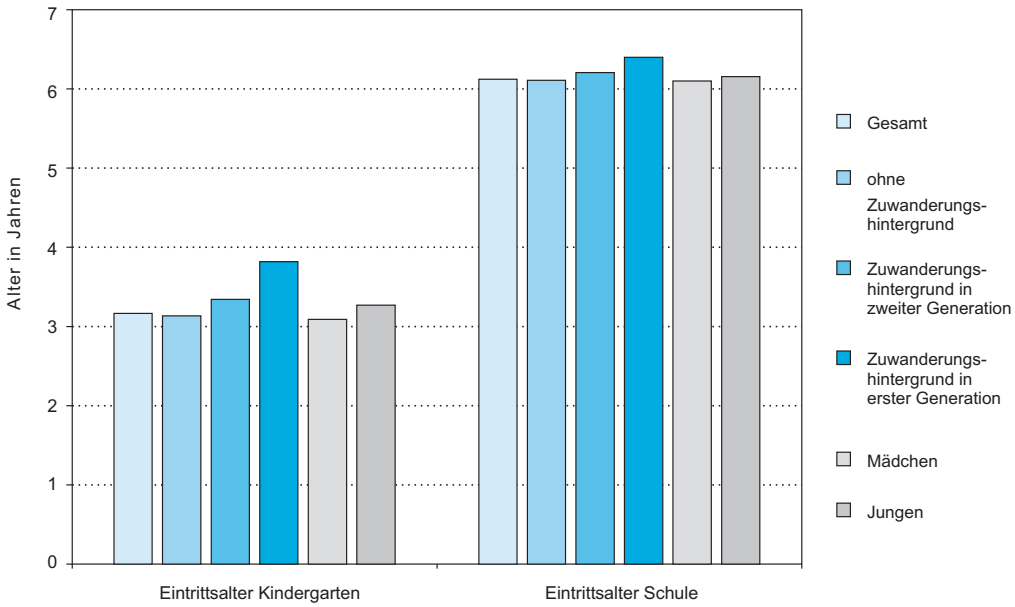


Abbildung 10.5: Eintrittsalter in Kindergarten und Schule

10 Prozent der Fünfzehnjährigen war ein Besuch einer Einrichtung der Vorschulerziehung nach Angaben der Eltern verpflichtend (vgl. Abbildung 10.6.).

An verschiedenen Stellen zeigen sich Zusammenhänge der Nutzung von außerfamilialer Bildung, Erziehung und Betreuung mit dem Bildungshintergrund der Eltern und dem sozioökonomischen Status der Familie. Die überwiegende Mehrheit der fünfzehnjährigen Schülerinnen und Schüler gab an, den Kindergarten ab einem Alter von drei Jahren besucht zu haben (58 Prozent). Dabei zeigt sich, dass die Eltern derjenigen Fünfzehnjährigen, die schon im Alter von zwei Jahren in einen Kindergarten gingen (13 Prozent der Schülerinnen und Schüler), über einen höheren Bildungs- und sozioökonomischen Hintergrund verfügen. Umgekehrt weisen Familien von Kindern, die erst im letzten Jahr vor der Einschulung in den Kindergarten gingen, deutlich ungünstigere Hintergrundmerkmale auf.

Das Engagement der Eltern an der Schule ihres Kindes

Die Mehrheit der fünfzehnjährigen Schülerinnen und Schüler lebt in einer Familie, in der die Eltern die Interaktion mit der Schule als relevant für die Entwicklung ihres Kindes beurteilen (85 Prozent). Für nur sehr wenige Fünfzehnjährige gilt, dass sich die Eltern nicht an der Schule ihres Kindes engagieren. Der Grund hierfür liegt bei einem Drittel der Schülerinnen und Schüler in der mangelnden zeitlichen Flexibilität der berufstätigen Eltern.

© Waxmann Verlag GmbH

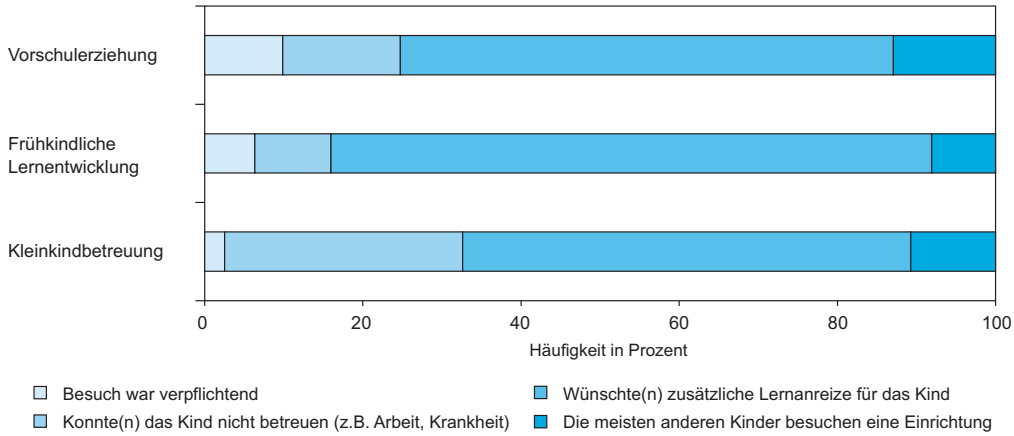


Abbildung 10.6: Gründe für die Wahl von frühkindlichen Betreuungseinrichtungen

Für den Bereich der allgemeinen Aktivitäten zeigt sich, dass die Eltern der fünfzehnjährigen Schülerinnen und Schüler vielfältig in der Schule engagiert sind. Dies geschieht beispielsweise durch die Teilnahme an schulischen Arbeitskreisen wie dem Elternbeirat (18 Prozent), Unterstützung durch handwerkliche Tätigkeiten oder Begleitung bei Ausflügen (17 Prozent) sowie durch Vorträge oder Unterstützung in der Schulbibliothek (17 Prozent). Zusätzlich zu diesen allgemeinen Formen des Engagements werden in PISA 2015 die spezifischen, auf das Kind bezogenen Unterstützungsaktivitäten differenziert betrachtet. Der Schwerpunkt liegt dabei auf Gesprächen von Eltern und Lehrkräften.

Die häufigste Form der Beteiligung der Eltern ist die Teilnahme an einem Elternabend. Bei neun von zehn Schülerinnen und Schülern geben die Eltern an, im letzten Schuljahr an einem Elternabend teilgenommen zu haben. Bei fast zwei Dritteln der Fünfzehnjährigen haben die Eltern im letzten Schuljahr auf eigenen Wunsch hin mit der Lehrerin oder dem Lehrer über das Verhalten ihres Kindes gesprochen. Bei etwa der Hälfte der fünfzehnjährigen Schülerinnen und Schüler fanden auf Initiative der Eltern hin Gespräche über die Entwicklung des Kindes statt. Aus Sicht der Eltern werden entsprechende Gespräche etwas seltener von den Lehrkräften ihres Kindes angeregt: Jeweils etwa ein Drittel der Schülerinnen und Schüler besucht Schulen, an denen Gespräche über das Verhalten des Kindes oder die Entwicklung des Kindes mit den Eltern durch Lehrkräfte angestoßen wurden (vgl. Abbildung 10.7.).

In diesen Gesprächen zwischen Eltern und Lehrkräften werden häufig auch Unterstützungsmöglichkeiten durch die Eltern thematisiert. Bei etwa der Hälfte der fünfzehnjährigen Schülerinnen und Schüler haben die Eltern mit den Lehrkräften darüber gesprochen, wie sie ihr Kind zu Hause bei den Hausaufgaben unterstützen können. Bei etwa einem Drittel der Fünfzehnjährigen fand ein Austausch zwischen Eltern und Lehrkräften über Erziehung, familiäre Unterstützung oder die Entwicklung des Kindes statt.

Ein weiterer Indikator für das Engagement und die Eingebundenheit der Eltern an der Schule ihres Kindes sind die sozialen Kontakte und Netzwerke im schulischen

© Waxmann Verlag GmbH

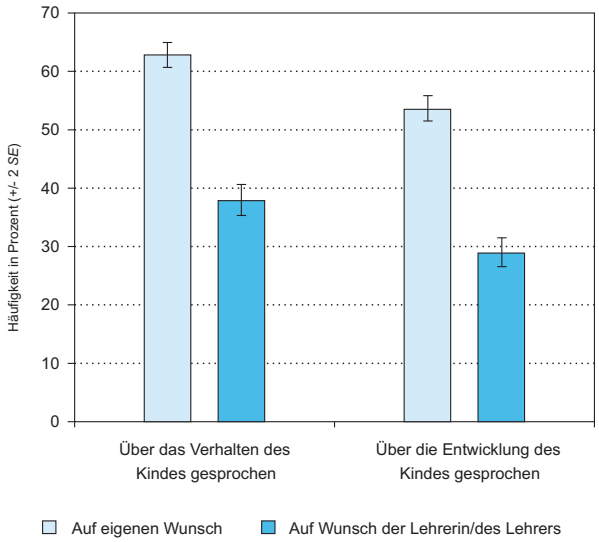


Abbildung 10.7: Häufigkeit der Gespräche zwischen Eltern und Lehrkräften

Umfeld (McNamara, Horvat, Weininger & Lareau, 2003). Diese können als soziales Kapital angesehen werden und eine wichtige Ressource darstellen. Die Eltern eines Großteils der fünfzehnjährigen Schülerinnen und Schüler kennen drei oder mehr Eltern von Freunden ihres Kindes an der Schule (67 Prozent) und können auch mindestens drei Freunde ihres Kindes namentlich benennen (78 Prozent). Hinsichtlich der Kontakte zu den Mitarbeiterinnen und Mitarbeitern an der Schule zeigt sich, dass die Eltern fast aller Fünfzehnjährigen mindestens eine, oft sogar drei oder mehr (54 Prozent) Personen an der Schule kennen, mit denen sie vertrauensvoll über ihr Kind sprechen können.

10.4 Elterliche Einstellungen, Überzeugungen, Motive und Interessen

Kriterien der Eltern bei der Schulwahl, Einschätzung der Schulqualität und Ausgaben für Bildung

Die Schulwahl ist eine wichtige Entscheidung in der Bildungsbiografie von Schülerinnen und Schülern. Dabei können verschiedene Motive und Argumente in Betracht gezogen sowie unterschiedlich gewichtet werden. Die Eltern wurden unter anderem danach gefragt, welche Möglichkeiten sie bei der Schulwahl für ihr Kind hatten und welche Kriterien ihnen dabei wichtig waren. Bei etwa einem Drittel der Schülerinnen und Schüler gab es keine alternative Schule in der Umgebung. Bei wiederum einem Drittel stand nach Auskunft der Eltern mindestens eine weitere Schule zur Auswahl, für 37 Prozent der Schülerinnen und Schüler bestand nach Einschätzung der Eltern die Möglichkeit, zwischen zwei oder mehr Schulen auszuwählen.

© Waxmann Verlag GmbH

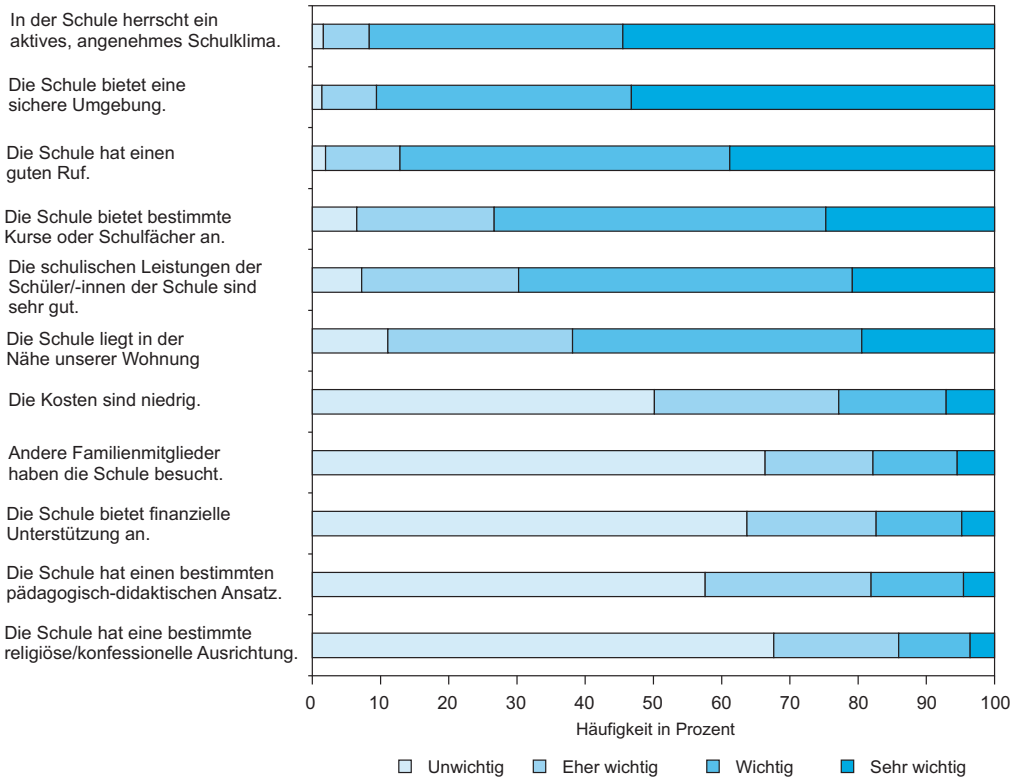


Abbildung 10.8: Kriterien der Eltern für die Schulwahl

Um zu erfahren, welche Aspekte und Kriterien Eltern bei der Schulwahl heranziehen, wurden sie gebeten einzuschätzen, wie wichtig unterschiedliche Merkmale von Schulen bei ihrer Entscheidung für die Wahl der weiterführenden Schule, die ihr Kind nach der Grundschule besuchen sollte, waren (vgl. Abbildung 10.8).

Die Auswertungen zeigen, dass die Eltern insbesondere das Schulklima (92 Prozent), eine sichere Umgebung (91 Prozent) sowie einen guten Ruf der Schule (87 Prozent) als wichtig beziehungsweise sehr wichtig für die Schulwahl angeben. Die fachlichen Schwerpunktsetzungen (73 Prozent), sehr gute Leistungen der Schule (70 Prozent) und die Nähe zum Wohnort (62 Prozent) waren ebenfalls eine wichtige Entscheidungsgrundlage für die Wahl der weiterführenden Schule. Demgegenüber sind Kostenaspekte (23 Prozent), der Besuch der Schule durch weitere Familienmitglieder (18 Prozent), vorhandene finanzielle Unterstützung durch die Schule (17 Prozent), spezifische pädagogisch-didaktische Ansätze (18 Prozent) oder eine konfessionelle Ausrichtung der Schule (14 Prozent) in den Familien der meisten fünfzehnjährigen Schülerinnen und Schüler weniger von Bedeutung.

Ein wichtiger Aspekt für das Engagement an der Schule des Kindes ist die Wahrnehmung der Schule durch die Eltern. Neben einer allgemeinen Zufriedenheit mit der Schule sind Einladungen der Eltern an die Schule sowie das Angebot von Informationen,

© Waxmann Verlag GmbH

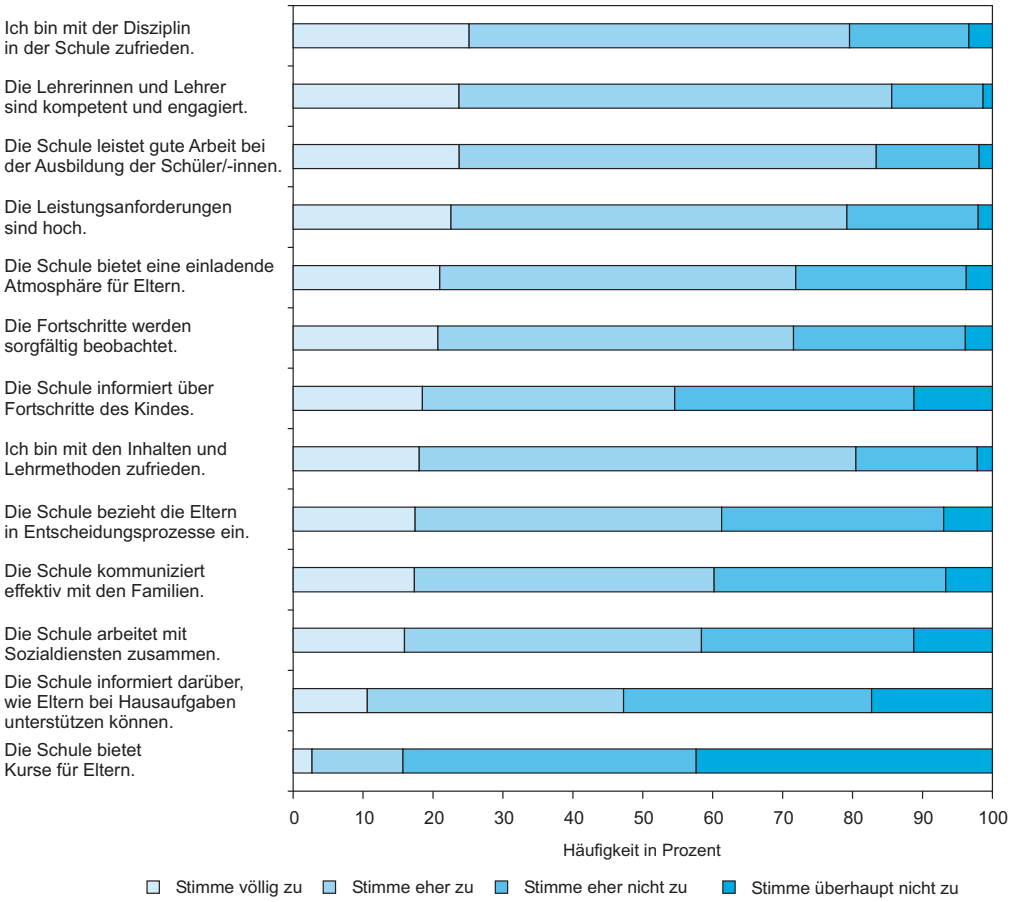


Abbildung 10.9: Merkmale der Schulqualität aus der Perspektive der Eltern

Fortbildungen oder Trainingsprogrammen für Eltern durch die Schule von Bedeutung. Diese Aspekte werden in PISA als Merkmale der Schulqualität erhoben (vgl. Abbildung 10.9.).

Die meisten Eltern nehmen die Lehrerinnen und Lehrer ihrer Kinder als kompetent wahr (86 Prozent), sie sind mit der Disziplin an der Schule ihres Kindes und mit den Lerninhalten (jeweils 80 Prozent) zufrieden und meinen, die Schule leiste gute Arbeit bei der Ausbildung der Schülerinnen und Schüler (83 Prozent) und lege hohe Leistungsanforderungen an (80 Prozent). Für die Zusammenarbeit von Schule und Familie sind eine einladende Atmosphäre an der Schule, eine regelmäßige Kommunikation von Lehrerinnen und Lehrern mit Eltern sowie eine Zusammenarbeit in professionellen Netzwerken sehr wichtig. Es zeigt sich, dass ein großer Anteil der Jugendlichen an Schulen unterrichtet wird, an denen aus Sicht der Eltern eine einladende und freundliche Atmosphäre herrscht, an denen die Fortschritte des Kindes sorgfältig beobachtet werden (jeweils 72 Prozent) und die nach Auskunft der Eltern in professionelle Netzwerke ein-

© Waxmann Verlag GmbH

gebunden sind (58 Prozent), das heißt mit Sozialdiensten und therapeutischen Einrichtungen zusammenarbeiten.

Auch die konkreten Kommunikations- und Unterstützungsangebote werden von den Eltern tendenziell positiv wahrgenommen. Etwa zwei Drittel der Fünfzehnjährigen besuchen Schulen, an denen eine effektive Kommunikation erfolgt und die Eltern regelmäßig in Entscheidungsprozesse einbezogen werden (jeweils 60 Prozent). Informationen über die Fortschritte ihres Kindes (55 Prozent), konkrete Unterstützungsangebote (47 Prozent) oder Elternfortbildungen (16 Prozent) erhalten allerdings vergleichsweise weniger der Eltern der fünfzehnjährigen Schülerinnen und Schüler.

Einige dieser Qualitätsmerkmale der Schulen wurden bereits in früheren PISA-Zyklen erfasst. Bezüglich der Zustimmung der Eltern lassen sich bei einer deskriptiven Betrachtung keine merklichen Unterschiede ausmachen. Über die Zeit lässt sich feststellen, dass die Schulen inzwischen häufiger die Fortschritte der Kinder dokumentieren und entsprechende Informationen zur Entwicklung des Kindes zur Verfügung stellen. Insgesamt sind die Eltern in Deutschland nach wie vor von der hohen Qualität der Schule ihres Kindes überzeugt.

In Deutschland besuchen 90 Prozent der Fünfzehnjährigen eine öffentliche Schule (destatis, 2016). Diese sind in aller Regel kostenlos und folgen den curricularen Vorgaben, wobei sie stets innerhalb bestimmter organisatorischer und rechtlicher Rahmenbedingungen agieren (vgl. Kapitel 5). Darüber hinaus besteht jedoch auch eine Nachfrage nach ergänzenden Dienstleistungen im Bildungsbereich, wie etwa nach Nachhilfe oder zusätzlichen kostenpflichtigen Kursen. Solche Dienstleistungen müssen privat finanziert werden. Studien in Deutschland weisen darauf hin, dass Eltern im Durchschnitt 93 Euro pro Monat für die Bildung ihrer Kinder ausgeben. Dabei steigen die Ausgaben mit dem Sozialstatus der Familie an (DIW Berlin, 2015).

Auch die Eltern der an PISA 2015 teilnehmenden Schülerinnen und Schüler wurden gefragt, wie viel Geld sie in den letzten zwölf Monaten ungefähr für Leistungen von Bildungseinrichtungen ausgegeben hatten. Dazu zählen allgemeine Schulgebühren, Honorare für Lehrpersonen und Gebühren für Nachhilfestunden. Ausstattung wie Bekleidung und Bücher war nicht eingeschlossen. Es gaben 80 Prozent der Familien an, zusätzliche Ausgaben für die Bildung getätigt zu haben. Bei knapp der Hälfte der fünfzehnjährigen Schülerinnen und Schüler lagen die Ausgaben zwischen 50 Euro und 250 Euro (45 Prozent), in den Familien von etwa einem Drittel der Fünfzehnjährigen wurden maximal 1200 Euro für die schulische Bildung des Kindes ausgegeben (27 Prozent), bei 9 Prozent der Familien waren es mehr als 1200 Euro.

Das Interesse der Eltern an den Naturwissenschaften

Eltern sind mit ihren Einstellungen und Interessen wichtige Vorbilder für ihre Kinder. Gleichzeitig können Erwartungen und Aspirationen der Eltern die Berufswahl ihrer Kinder beeinflussen. Vielfach kann beobachtet werden, dass innerhalb einer Familie gleiche Berufe beziehungsweise Berufe aus einem ähnlichen Berufsfeld gewählt werden. Ein

© Waxmann Verlag GmbH

Drittel der Fünfzehnjährigen hat Eltern, die einem naturwissenschaftlichen Beruf nachgehen (33 Prozent). Hierzu zählen alle Berufe, die eine weiterführende Ausbildung oder ein Studium in einem naturwissenschaftlichen Bereich (z.B. Ingenieure, Optiker, Ärzte, technische Assistenten) beinhaltet. Die Eltern jeder zehnten Schülerin beziehungsweise jedes zehnten Schülers erwarten, dass ihr Kind einen naturwissenschaftlichen Beruf ergreift oder eine entsprechende Ausbildung anstreben wird (11 Prozent).

Ein weiterer wichtiger Aspekt ist das Interesse der Eltern an den Naturwissenschaften selbst. In PISA 2015 wurden die Eltern nach der Bedeutung von Naturwissenschaften für die Gesellschaft und das eigene Leben gefragt. Es zeigt sich, dass praktisch alle Eltern naturwissenschaftlichen Themen eine hohe Bedeutung für die Gesellschaft (95 Prozent) und das Verständnis der Alltagswelt beimessen (97 Prozent). Fortschritte in den Naturwissenschaften werden dabei in den meisten Elternhäusern der Fünfzehnjährigen mit sozialen Fortschritten in Verbindung gebracht (Zustimmung: 83 Prozent).

10.5 Zusammenhänge der elterlichen Unterstützung mit den Kompetenzen der Schülerinnen und Schüler in den Naturwissenschaften

Die positive Wirkung von elterlichem Engagement und einer kontinuierlichen Zusammenarbeit der Eltern mit der Schule des Kindes konnte wie bereits dargelegt in vielen wissenschaftlichen Untersuchungen gezeigt werden. Im Rahmen von PISA ist es möglich, die Zusammenhänge der Strukturmerkmale der Familie sowie der Prozessmerkmale der Lernumgebung in der Familie und der Förderung durch die Eltern mit den Kompetenzen der Schülerinnen und Schüler zu untersuchen. In PISA 2009 wurden entsprechende Analysen für die Lesekompetenz, die Lesemotivation und das Interesse am Lesen durchgeführt (vgl. Hertel et al., 2010). In PISA 2015 wird der Schwerpunkt auf die Naturwissenschaften gelegt und es wird auf der Grundlage von Regressionsanalysen untersucht, welche Zusammenhänge zwischen Struktur- und Prozessmerkmalen der Förderung und der Kompetenz der Schülerinnen und Schüler in den Naturwissenschaften (PISA-Testwerte) bestehen (vgl. Abbildung 10.10).

Als individuelle Herkunftsmerkmale und als Strukturmerkmale der Familie wurden bei den Analysen – wie in PISA üblich – das Geschlecht des Kindes, der sozioökonomische Status, der Zuwanderungshintergrund und die besuchte Schulart berücksichtigt. Als Prozessmerkmale wurden das Alter bei Eintritt in eine Kinderbetreuungseinrichtung sowie dessen quadrierter Term, die naturwissenschaftsbezogene Förderung des Kindes durch die Eltern in der Grundschulzeit und aktuell sowie das Interesse der Eltern an Naturwissenschaften einbezogen. Die einzelnen Indikatoren wurden gleichzeitig in das Modell aufgenommen, sodass die berichteten Ergebnisse jeweils um den Einfluss der anderen Indikatoren in dem Modell kontrolliert werden.

© Waxmann Verlag GmbH

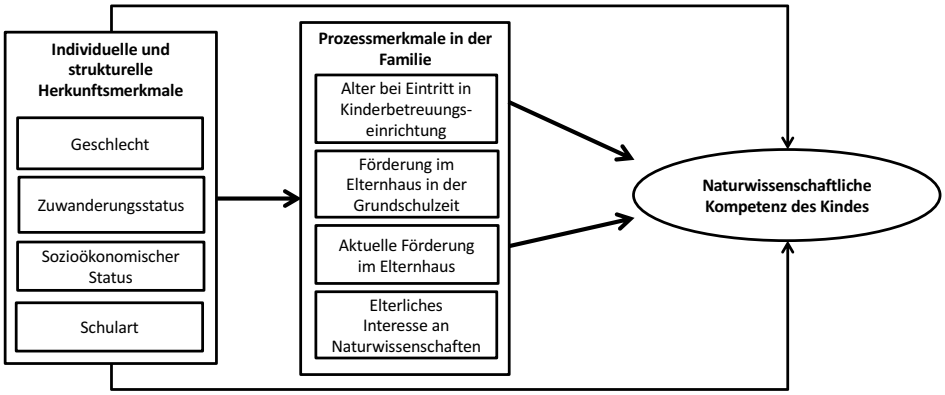


Abbildung 10.10: Theoretisches Pfadmodell der Zusammenhänge zwischen Struktur- und Prozessmerkmalen der Unterstützung in der Familie und der naturwissenschaftlichen Kompetenz der Schülerinnen und Schüler (in Anlehnung an Hertel et al., 2010, S. 268)

Es wurden zwei aufeinander aufbauende Regressionsmodelle spezifiziert (vgl. Tabelle 10.1). Im ersten Modell wurden nur die individuellen Herkunftsmerkmale und Strukturmerkmale einbezogen, im zweiten Modell wurden zusätzlich die Prozessmerkmale der Lernumgebung und Förderung im Elternhaus aufgenommen. Beide Modelle bestätigen die vielfach beschriebenen Zusammenhänge der Strukturmerkmale mit fachlichen Kompetenzen. Schülerinnen und Schüler aus bildungsnäheren sozialen Schichten, Schülerinnen und Schüler ohne Zuwanderungshintergrund sowie Schülerinnen und Schüler an Gymnasien verfügen über höhere Kompetenzen in den Naturwissenschaften. Zudem erreichen Jungen höhere Ergebnisse in den naturwissenschaftlichen Kompetenztests als Mädchen (vgl. auch Kapitel 2). Durch diese individuellen Herkunftsmerkmale und Strukturmerkmale lassen sich bereits etwa ein Drittel der Varianz in der naturwissenschaftlichen Kompetenz der Schülerinnen und Schüler erklären (R^2 = 37 Prozent).

Unter Hinzunahme der Prozessmerkmale in einem zweiten Modell lässt sich die Varianzaufklärung nur gering erhöhen (R^2 = 38 Prozent). Allerdings reduziert sich der Effekt einzelner Strukturmerkmale. Dabei kann festgehalten werden, dass alle untersuchten Prozessmerkmale signifikante Zusammenhänge zu den naturwissenschaftlichen Kompetenzen der Schülerinnen und Schüler aufweisen. Schülerinnen und Schüler, die während der Grundschulzeit mehr Förderung im naturwissenschaftlichen Bereich durch ihre Eltern erfahren haben und deren Eltern sich für naturwissenschaftliche Themen interessieren, erzielen höhere Werte in den naturwissenschaftlichen Kompetenztests in PISA 2015.

Für den Besuch einer institutionellen Frühförderung lässt sich feststellen, dass ein niedrigeres Alter beim Kindergarteneintritt tendenziell mit besseren Leistungen im Altern von 15 Jahren einhergeht. Allerdings zeigt sich bei genauerer Analyse ein qua-

© Waxmann Verlag GmbH

Tabelle 10.1 Regressionsmodelle zur Vorhersage der naturwissenschaftlichen Kompetenz unter Kontrolle der häuslichen Lernumgebung

	Modell I Strukturmerkmale Regressionskoeffizienten			**Modell II** Struktur- und Prozessmerkmale Regressionskoeffizienten		
Prädiktoren	*b*	(*SE*)	β	*b*	(*SE*)	β
Naturwissenschaftliche Kompetenz auf						
Zuwanderungshintergrund	**-40.59**	4.10	**-0.19**	**-35.11**	5.13	**-0.15**
ESCS	**18.87**	1.80	**0.18**	**15.70**	2.27	**0.15**
Geschlecht der Schülerin/des Schülers	**16.15**	2.52	**0.08**	**12.88**	3.79	**0.07**
Schulart Gymnasium versus nicht gymnasiale Schularten	**-89.87**	4.05	**-0.45**	**-79.00**	4.95	**-0.41**
R^2/korrigiertes R^2	.37/.37					
Eintrittsalter Kindergarten				**41.19**	7.69	**0.40**
Eintrittsalter quadrierter Term				**-48.79**	8.13	**-0.46**
Förderung im Grundschulalter				**13.35**	2.37	**0.12**
Aktuelle Förderung im Elternhaus				**-15.19**	2.74	**-0.12**
Einstellung der Eltern zu Naturwissenschaften				**6.70**	1.51	**0.08**
R^2/korrigiertes R^2	.43/.38					

signifikante Koeffizienten sind fett gedruckt

dratischer Zusammenhang. Im Durchschnitt sind die Kompetenzen in den Naturwissenschaften etwas niedriger bei Kindern, die bei Beginn der institutionellen Betreuung nach den selbst berichteten Angaben sehr jung (jünger als ein Jahr) beziehungsweise deutlich älter (älter als vier Jahre) waren.

Eine aktive Förderung und Beteiligung an den fachspezifischen Bildungsprozessen des Kindes in der Grundschulzeit durch die Eltern geht Jahre später bei den fünfzehnjährigen Schülerinnen und Schülern mit höheren Werten in den naturwissenschaftlichen Kompetenztests einher. Eine Förderung im Elternhaus zum Zeitpunkt der PISA-2015-Erhebung geht im Querschnitt betrachtet mit geringeren Kompetenzen in den Naturwissenschaften einher. Wie bereits in früheren PISA-Erhebungen zeigen sich hier sogenannte remediale Effekte. Schülerinnen und Schüler im Alter von 15 Jahren werden durch die Eltern am ehesten dann gefördert, wenn schwache Leistungen vorliegen. Vergleichbare Befunde wurden im Rahmen von PISA 2009 für den Zusammenhang einer frühen Leseförderung in der Familie und der Lesekompetenz der Fünfzehnjährigen gefunden (vgl. Hertel et al., 2010).

© Waxmann Verlag GmbH

10.6 Diskussion und Ausblick

Der Elternfragebogen der PISA-Studie ist ein wertvolles Instrument, um Merkmale der häuslichen Lernumgebung, der frühen Förderung sowie der Zusammenarbeit mit der Schule zu erfassen und über die Zeit als Merkmale des Bildungssystems abzubilden. In Ergänzung zum Fragebogen für Schülerinnen und Schüler ermöglicht es diese Befragung, die Sicht der Eltern und Erziehungsberechtigten auf die Schule ihres Kindes sowie auf die institutionellen Bildungserfahrungen in der frühen Kindheit darzustellen. Die in diesem Kapitel berichteten Ergebnisse geben zudem einen Einblick in die familienbezogenen Lernumgebungen und das häusliche Umfeld der Fünfzehnjährigen.

Allerdings gehen der vergleichsweise geringe Rücklauf ausgefüllter Elternfragebögen sowie die höhere Ausfallquote in Familien mit Zuwanderungshintergrund, wie auch in den vorangegangenen PISA-Erhebungsrunden, mit deutlichen Einschränkungen der Interpretierbarkeit und der Generalisierbarkeit der Befunde einher. Insbesondere ist die Repräsentativität für die Kohorte der fünfzehnjährigen Schülerinnen und Schüler nicht gegeben, weshalb Verallgemeinerungen auf der Grundlage der Daten aus dem Elternfragebogen nicht zulässig sind. Gleichwohl lassen sich aus den Ergebnissen der Befragung der Eltern und Erziehungsberechtigten vorsichtig Hinweise darauf ableiten, welche Ansatzpunkte es in den Familien und in ihrer Zusammenarbeit mit den Schulen geben könnte, um auch außerschulisches, elternunterstütztes Lernen im Bereich Naturwissenschaften in Deutschland zu stärken.

Insgesamt betrachtet sind die Voraussetzungen für die Förderung im Elternhaus und die Zusammenarbeit von Eltern und Schule in Deutschland – zumindest in Bezug auf die Gruppe der an der Befragung teilnehmenden Eltern – durchaus positiv zu beurteilen. Die meisten Eltern beschreiben sich selbst als motiviert, sich an der Schule ihres Kindes aktiv zu beteiligen, das Kind beim Lernen zu unterstützen und es zu fördern. An vielen Schulen nehmen Eltern eine hohe Schulqualität und eine zur Zusammenarbeit einladende Atmosphäre wahr. Dem gegenüber stehen die vergleichsweise selten genannten konkreten Förder- und Unterstützungsangebote durch die Eltern im Familienalltag – sowohl in der Grundschulzeit als auch bei den Fünfzehnjährigen. Dies kann ein Hinweis darauf sein, dass das Potenzial für die Unterstützung naturwissenschaftlicher Interessen und Kompetenzentwicklung im häuslichen und familiären Bereich in Deutschland nicht voll ausgeschöpft wird.

Für einen Großteil der Fünfzehnjährigen in Deutschland ist der Kompetenzerwerb im Bereich der Naturwissenschaften fast ausschließlich auf die Schule beschränkt. Dies entspricht dem wahrgenommenen Vermittlungsmonopol, das den Schulen im Bereich naturwissenschaftlicher Bildung ähnlich wie auch bei der Mathematik zugeschrieben wird (vgl. Baumert & Stanat, 2010). Eine mögliche Perspektive könnte darin bestehen, Eltern durch spezifische Informationsangebote und zielgerichtete Beratungsgespräche umfassender über konkrete Unterstützungsmöglichkeiten zu informieren. Das kann beispielsweise bedeuten, Ansprechpartner für Eltern zu benennen oder Programme für Freizeitgestaltung mit naturwissenschaftlichen Themen zur Verfügung zu stellen. Eine

© Waxmann Verlag GmbH

weitere, deutlich aufwendigere Möglichkeit ist es, Elternkurse oder Elterntrainings anzubieten, in denen konkrete Förder- und Unterstützungsstrategien vermittelt werden (Bruder, Perels & Schmitz, 2004).

Die Bedeutung von familiärer Lernumgebung und Förderung in der Familie wurde in den hier vorgestellten Analysen einmal mehr deutlich. Selbst im Alter von 15 Jahren konnten bei den Schülerinnen und Schülern positive Effekte einer Förderung durch die Eltern in der Grundschulzeit gefunden werden. Für die Förderung zum Zeitpunkt des PISA-Tests zeigte sich – wie auch in vorangegangenen PISA-Erhebungen – ein negativer Zusammenhang mit der Kompetenz der Schülerinnen und Schüler. Es ist anzunehmen, dass dies im Wesentlichen auf den beschriebenen Remedialeffekt zurückzuführen ist, da vor allem dann gefördert und unterstützt wird, wenn Probleme auftreten und Defizite bestehen. Ein solcher Effekt spricht dafür, dass Förderung tendenziell als eine Unterstützung im Bedarfsfall angesehen wird. Einschlägige Studien sprechen dafür, dass eine proaktive und kontinuierliche Förderung und Unterstützung durch die Eltern effektiver und nachhaltiger ist als eine bedarfsgesteuerte, die erst dann greift, wenn Probleme auftreten (z. B. in Bezug auf Hausaufgabenhilfe Dumont, Trautwein, Nagy & Nagengast, 2014).

Die PISA-Daten selbst erlauben aufgrund des querschnittlichen Designs jedoch keine entsprechenden Aussagen. Darüber hinaus wäre hinsichtlich der Erfassung des elterlichen Engagements eine Unterscheidung elterlicher Aktivitäten *mit* Schulbezug (z. B. dem Kind bei den Hausaufgaben helfen) und Aktivitäten *ohne* Schulbezug (z. B. gemeinsam die Hauptmahlzeit einnehmen) wünschenswert. Eine entsprechende Differenzierung würde spezifischere Analysen der Zusammenhänge mit den Kompetenzen und Interessen der Schülerinnen und Schüler erlauben.

Insgesamt ist zu überlegen, wie die frühe, proaktive Förderung bezogen auf die Naturwissenschaften in den Familien auch außerhalb von Bildungsinstitutionen stärker unterstützt werden kann. In Domänen wie etwa dem Lesen wird dies durch Leseprogramme für unterschiedliche Altersgruppen bereits umgesetzt. Trotz einer Fülle von Programmen wie Jugend forscht, MINTEC-Schulen oder naturwissenschaftlichen Wettbewerben, die vorwiegend an Gymnasien weit verbreitet sind, besteht für die Fünfzehnjährigen in Deutschland Bedarf in Bezug auf Unterstützung bei der fachlichen Entwicklung der naturwissenschaftlichen Kompetenz und Interesse. Ferienprogramme mit naturwissenschaftlichem Schwerpunkt, familienfreundliche Angebote naturwissenschaftlicher Museen oder auch elterliches Engagement in Form von Workshops an der Schule ihres Kindes, die naturwissenschaftliche Themen und den Alltag miteinander verknüpfen, können Ansatzpunkte hierfür bieten.

Zusammenfassend kann festgehalten werden, dass die Förderung der Entwicklung naturwissenschaftlicher Kompetenz im familiären Rahmen meist weniger explizit stattfindet als in der Schule. Sie könnte aber durchaus bedeutsame Effekte in Bezug auf Begegnungen mit naturwissenschaftlichen Themen, das Interesse für naturwissenschaftliche Phänomene und Fragestellungen oder auch den Kontakt mit Berufen im naturwissenschaftlichen oder technischen Bereich haben. Dem in der Schule institutionalisierten Rahmen für die Entwicklung naturwissenschaftlicher Kompetenz steht die Familie der

© Waxmann Verlag GmbH

Jugendlichen als nicht institutionalisierter Kontext zur Seite, der in Deutschland für die meisten Fünfzehnjährigen kaum auf die fachspezifische Förderung naturwissenschaftlicher Bildung ausgerichtet ist. Durch eine kontinuierliche und effektive Zusammenarbeit von Familie und Schule können sich die spezifischen Schwerpunktsetzungen ergänzen und zu optimierten Lernumgebungen für die Schülerinnen und Schülern beitragen.

Literatur

Ainsworth, M. D. S., Bell, S. M. & Stayton, D. J. (1974). Infant-mother attachment and social development: „Socialisation" as a product of reciprocal responsiveness to signals. In M. P. Richards (Hrsg.), *The integration of a child into a social world* (S. 99–135). London: Cambridge University Press.

Anders, Y., Roßbach, H.-G. & Kuger, S. (im Druck). Early childhood learning experiences. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing contexts of learning: An international perspective*. Berlin: Springer.

Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016*. Bielefeld: wbv.

Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften – Ergebnisse des Forschungsprogramms COACTIV* (S. 29–53). Münster: Waxmann.

Baumert, J. & Stanat, P. (2010). Internationale Schulleistungsvergleiche. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Aufl., S. 324–335). Weinheim: Beltz.

Berngruber, A., Alt, C. & Huber, S. (2014). *Das Recht auf den Platz – Tagesbetreuung für Kinder unter drei Jahren auf dem Prüfstand*. Zugriff am 17.10.2016. Verfügbar unter: http://www.dji.de/index.php?id=43589

Boudon, R. (1974). *Education, opportunity, and social inequality; changing prospects in Western society*. New York: Wiley.

Bruder, S., Hertel, S., Gerich, M. & Schmitz, B. (2014). Lehrer als Berater. In E. Terhart, I. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Aufl.) (S. 905–919). Münster: Waxmann.

Bruder, S., Perels, F. & Schmitz, B. (2004). Selbstregulation und elterliche Hausaufgabenunterstützung. Die Evaluation eines Elterntrainings für Kinder der Sekundarstufe I. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 36* (3), 139–146.

Catsambis, S. (1998). *Expanding knowledge of parental involvement in secondary education – Effects on high school academic success.* Baltimore, MD: CRESPAR.

Corno, L. (2000). Looking at homework differently. *Elementary School Journal, 100*, 529–548.

Cox, D. D. (2005). Evidence based interventions using home school collaboration. *School Psychology Quarterly, 20* (4), 473–497.

Department of Health. (2000). *Framework for the assessment of children in need and their families*. London: The Stationery Office.

Destatis. (2016). *Privatschulen nach Schularten*. Zugriff am 17.10.2016. Verfügbar unter: https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/BildungForschungKultur/Schulen/Tabellen/AllgemeinBildendeBeruflichePrivateSchulenSchularten.html

© Waxmann Verlag GmbH

DIW Berlin. (2015) *Private Bildungsausgaben für Kinder: Einkommensschwache Familien sind relativ stärker belastet*. Zugriff am 01.09.2016. Verfügbar unter: https://www.diw.de/six-cms/detail.php?id=diw_01.c.497259.de

Dumont, H., Trautwein, U., Nagy, G. & Nagengast, B. (2014). Quality of parental homework involvement: Predictors and reciprocal relations with academic functioning in the reading domain. *Journal of Educational Psychology, 106*, 144–161.

Freyaldenhofen, I. (2005). *Schule in der Krise? Psychologische Beratung als Antwort! Theoretische und praktische Hinweise für eine gelingende Schulberatung – lösungsorientiert und individuell.* Stuttgart: ibidem-Verlag.

Green, C. L., Walker, J. M. T., Hoover-Dempsey, K. V. & Sandler, H. M. (2007). Parents' motivations for involvement in children's education: An empirical test of a theoretical model of parental involvement. *Journal of Educational Psychology, 99* (3), 532–544.

Grossmann, K. & Grossmann, K. E. (2012). *Bindungen – das Gefüge psychischer Sicherheit*. Stuttgart: Klett-Cotta.

Hertel, S. (2016). Elternberatung im schulischen Kontext. In S. Frank & A. Sliwka (Hrsg.), *Eltern und Schule* (S. 116–126). Weinheim: Beltz Juventa.

Hertel, S., Bruder, S., Jude, N. & Steinert, B. (2013). Elternberatung an Schulen im Sekundarbereich: Schulische Rahmenbedingungen, Beratungsangebote der Lehrkräfte und Nutzung von Beratung durch die Eltern. In N. Jude & E. Klieme (Hrsg.), *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung. Zeitschrift für Pädagogik*, *Beiheft, 59,* 40–63.

Hertel, S., Jude, N. & Naumann, J. (2010). Leseförderung im Elternhaus. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 255–276). Münster: Waxmann.

Hill, N. E., Castellino, D. R., Lansford, J. E., Nowlin, P., Dodge, K. A., Bates, J. E. & Pettit, G. S. (2004). Parental academic involvement as related to school behavior, achievement, and aspirations: Demographic variations across adolescence. *Child Development, 75,* 1491–1509.

Hoover-Dempsey, K. V., Bassler, O. C. & Burow, R. (1995). Parents' report involvement in students' homework: Parameters of reported strategy and practice. *Elementary School Journal*, *95,* 435–450.

Hoover-Dempsey, K. V. & Sandler, H. M. (1997). Why do parents become involved in their children's education? *Review of Educational Research, 67*, 3–42.

Hoover-Dempsey, K. V., Walker, J. M. T., Sandler, H. M., Whetsel, D., Green, C. L., Wilkins, A. S. & Closson, K. E. (2005). Why do parents become involved? Research findings and implications. *Elementary School Journal*, *106* (2), 105–130.

Jeynes, W. H. (2007). The relationship between parental involvement and urban secondary school student academic achievement: a meta-analysis. *Urban Education, 42* (1), 82–110.

Jude, N. & Hertel, S. (2014). *PISA 2012. Die Perspektive der Eltern auf die Zusammenarbeit mit der Schule*. Vortrag auf dem 49. Kongress der Deutschen Gesellschaft für Psychologie. Bochum.

Kluczniok, K. & Roßbach, H.-G. (2014). Conceptions of educational quality for kindergartens. *Zeitschrift für Erziehungswissenschaft, 17* (6), 145–158.

Kohl, G. O., Lengua, L. J. & McMahon, R. J. (2000). Parent involvement in school conceptualizing multiple dimensions and their relations with family and demographic risk factors. *Journal of School Psychology, 38* (6), 501–523.

© Waxmann Verlag GmbH

Lehrl, S., Kluczniok, K. & Roßbach, H.-G. (2016). Longer-term associations of preschool education: The predictive role of preschool quality for the development of mathematical skills through elementary school. *Early Childhood Research Quarterly, 36,* 475–488.

Maaz, K., Neumann, M. & Baumert, J. (Hrsg.). (2014). *Herkunft und Bildungserfolg von der frühen Kindheit bis ins Erwachsenenalter: Forschungsstand und Interventionsmöglichkeiten aus interdisziplinärer Perspektive* (Sonderheft 24-2014 der Zeitschrift für Erziehungswissenschaft). Wiesbaden: Springer VS.

McElvany, N. & Becker, M. (2009). Häusliche Lesesozialisation: Wirkt das Vorbild der Eltern? *Unterrichtswissenschaft*, *37*, 246–261.

McNamara Horvat, E., Weininger, E. B. & Lareau, A. (2003). From social ties to social capital: Class differences in the relations between schools and parent networks. *American Educational Research Journal, 40* (2), 319–351.

Schober, P. S. & Spieß, C. G. (2014) Die Kita-Qualität ist für das Erwerbsverhalten von Müttern mit Kleinkindern relevant: Zusammenhang eindeutiger in Ostdeutschland. *DIW Wochenbericht*, *21*, 463–471.

Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., Taggart, B., Toth, K., Smees, R., Draghici, D., Mayo, A. & Welcomme, W. (2012). *Effective pre-school, primary and secondary education project (EPPSE 3-14): Final report from the key stage 3 phase: Influences on students' development from age 11–14 full report.* London: Department for Education.

Szczesny, M. & Watermann, R. (2011). Differenzielle Einflüsse von Familie und Schulform auf Leseleistung und soziale Kompetenzen. *Journal for Educational Research Online/Journal für Bildungsforschung Online*, *3* (1), 168–193. Zugriff am 24.10.2016. Verfügbar unter: http://www.j-e-r-o.de/index.php/jero/view/130/111

Tillmann, K.-J. (2010). *Sozialisationstheorien. Eine Einführung in den Zusammenhang von Gesellschaft, Institution und Subjektwerdung*. Reinbek: Rowohlt.

van de Pol, J., Volman, M. & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review, 22* (3), 271–296. http://dx.doi.org/doi: 10.1007/s10648-010-9127-6

© Waxmann Verlag GmbH

© Waxmann Verlag GmbH

11 Fünfzehn Jahre PISA: Bilanz und Ausblick

Kristina Reiss & Christine Sälzer

Nach mittlerweile sechs abgeschlossenen Erhebungsrunden blickt PISA auf eine fünfzehnjährige Geschichte zurück. Die Studie ist damit genauso alt wie die Schülerinnen und Schüler, die 2015 weltweit auf ihre Grundbildung getestet und zu ihren Entwicklungsbedingungen befragt wurden. Um die Ergebnisse wird es – wie eigentlich immer – Diskussionen geben. Welche Aussagekraft haben Studien wie PISA? Was bedeuten ein paar Prozentpunkte mehr oder weniger? Welchen Nutzen hat die Gesellschaft von PISA? Lohnt sich der Aufwand für die wiederholten Messungen? Sind es die richtigen Fragen, auf die PISA eine Antwort gibt?

Auch wenn 15 Jahre ein eher kurzer Zeitraum sind, so hat sich doch seit dem Jahr 2000 (nicht nur) in Deutschland vieles geändert. Es sind insbesondere technische Neuerungen, die auch das Leben von Jugendlichen in den letzten Jahren maßgeblich beeinflusst haben. Konkret sind Smartphone und Tablet, Podcast und Facebook allesamt Produkte des 21. Jahrhunderts. Bildung braucht Kontinuität, muss aber auch Innovationen aufnehmen, damit Kinder und Jugendliche bestmöglich auf die Welt vorbereitet werden. In diesem Spannungsfeld ist PISA 2015 zu sehen. Die in diesem Band vorgestellten Ergebnisse reihen sich ein in ein inzwischen breit etabliertes Bildungsmonitoring, das im Sinne der Trendberichterstattung den Anschluss an frühere Erhebungszyklen ermöglicht, zugleich aber den Anspruch hat, lebensweltnahe Bildungsergebnisse zu erfassen. Das bedeutet, dass die aktuelle Lebenswelt der untersuchten Fünfzehnjährigen mit Aufgaben, Testformaten und behandelten Themen erreicht werden soll. PISA kann sich daher nicht erlauben, lediglich einmal definierte Standards und Prozeduren in jeder Runde unreflektiert aufs Neue umzusetzen, sondern muss Veränderungen berücksichtigen und Anpassungen vornehmen.

Im Folgenden soll einerseits Bilanz aus 15 Jahren PISA gezogen und andererseits ein Ausblick auf künftige Herausforderungen an die Studie gewagt werden. International vergleichende Bildungsstudien befinden sich in mehreren Spannungsfeldern: Zwischen überstaatlicher Vergleichbarkeit und nationaler Adäquatheit, zwischen Konsens und Diversität, und – wie oben bereits begrifflich gefasst – zwischen Kontinuität und Innovation.

© Waxmann Verlag GmbH

11.1 Zwischen Kontinuität und Innovation

Der Blick zurück auf sechs PISA-Erhebungsrunden lässt sich gedanklich zunächst in „2 · 3“ gliedern, also in zwei abgeschlossene Durchgänge, in denen jeweils eine der drei Kompetenzdomänen als Schwerpunkt gemessen wurde (PISA 2000 und 2009: Lesen, PISA 2003 und 2012: Mathematik, PISA 2006 und 2015: Naturwissenschaften). Diese Sichtweise stellt die Kontinuität in den Vordergrund, die Trendaussagen über die Jahre hinweg im Blick hat. Trends können allerdings nur dann seriös berichtet werden, wenn ein ausreichend breiter Teil der Aufgaben und Fragen an die Schülerinnen und Schüler identisch in mehreren Erhebungsrunden zum Einsatz kommt. Daher sind die Leistungen in Mathematik nur bis 2003, in Naturwissenschaften – über alle Staaten hinweg – bis 2006 rückwärts vergleichbar. Davon losgelöst wurden in jeder neuen PISA-Runde auch neue Aufgaben eingesetzt, um dem Anspruch der Anbindung an die Lebenswelt und den Alltag der untersuchten Jugendlichen gerecht zu werden.

Mit Blick auf PISA 2015 ist eine Gliederung der sechs Runden in „5 + 1“ allerdings genauso plausibel, nämlich dann, wenn man nach den Erhebungsbedingungen, insbesondere dem Darbietungsmodus der Aufgaben differenziert. Dieser bestand von PISA 2000 bis PISA 2012 in Papier-und-Bleistift-Tests, mit PISA 2015 wurde das computerbasierte Testen eingeführt. Damit und mit dem gegenüber früheren PISA-Erhebungsrunden neuen Skalierungsmodell für die Daten wurde im Grunde eine Art zweiter Anfangspunkt für die Trendberichterstattung in PISA gesetzt, auf den sich künftige Erhebungsrunden beziehen werden. Diese Neuerungen müssen berücksichtigt werden, wenn man den Bezug der Kompetenzmittelwerte zu den entsprechenden Mittelwerten in früheren Runden herstellen möchte. Zwar lassen sich auch diesmal Differenzwerte abbilden und auf ihre Signifikanz testen, ihre Interpretation bedarf jedoch einer kontextrelevanten Beschreibung und Abwägung möglicher Bedingungsfaktoren. Die bloße Angabe der Differenz sagt in PISA 2015 mit Blick auf vergangene Runden weniger aus als bisher, da das Zustandekommen der Mittelwerte, Streuungen und Verteilungen auf Kompetenzstufen diesmal von mehreren Veränderungen gleichzeitig beeinflusst ist. Dennoch kann die Bedeutsamkeit von Veränderungen gegenüber früheren PISA-Erhebungsrunden sehr wohl stichhaltig interpretiert werden, wenn man zwei Dinge berücksichtigt: Zum einen gab es Modus-Effekte, die mit dem Wechsel von Papier zu Computer zusammenhängen. Vertiefende Analysen am Zentrum für Internationale Bildungsvergleichsstudien deuten darauf hin, dass die Aufgaben in PISA 2015 am Computer für die Schülerinnen und Schüler in Deutschland eher schwieriger waren als auf Papier. Dies entspricht der einschlägigen Forschungsliteratur und knüpft an die Befunde der ersten ICILS-Studie 2013 an, nach der Computer an Schulen in Deutschland nur vergleichsweise selten eingesetzt werden und die infrastrukturellen Rahmenbedingungen ausbaufähig sind – auch etwa jede dritte an PISA teilnehmende Schule in Deutschland verfügte nicht über die benötigte Computerausstattung für die Durchführung von PISA 2015 und musste mit Laptops versorgt werden. Unter Beachtung möglicher Modus-Effekte und weiterer Faktoren ist in PISA 2015 damit die Verbindung der für die Trendberichterstattung und das konti-

© Waxmann Verlag GmbH

nuierliche Bildungsmonitoring notwendigen Kontinuität mit innovativen, lebensweltnahen Elementen eine größere Herausforderung als bis PISA 2012, jedoch keine, die nicht zu meistern wäre. In diesem nationalen Bericht zu PISA 2015 für Deutschland wird entsprechend an die Trendberichterstattung vergangener Runden angeknüpft, wobei jede Domäne auf die jeweils letzte zurückliegende Hauptdomänenrunde bezogen wird (also z. B. Naturwissenschaften auf PISA 2006). Entsprechend vorsichtig wurden Befunde zu Trends als Entwicklungen über die Zeit interpretiert. Die OECD bietet – wenn auch nur in Form eines Anhangs – alternativ fiktive Mittelwerte für vergangene PISA-Runden an, die sich dabei nur auf das neue Skalierungsmodell beziehen und naturgemäß den Wechsel zum computerbasierten Testen nicht berücksichtigen können.

PISA 2015 nimmt also in gewisser Hinsicht eine Sonderrolle ein, da mehrere Variablen zur gleichen Zeit verändert wurden und deren Auswirkungen nicht separat quantifiziert werden können. Dennoch liefern die Ergebnisse aus PISA 2015 Indikatoren, die auch und gerade für die bildungspolitischen Fragen in Deutschland wichtige Anhaltspunkte geben.

11.2 Disparitäten: Thema des PISA-Schocks und dauerhafte Herausforderung

Eines der wichtigsten Themen, die den öffentlichen, politischen und wissenschaftlichen Diskurs seit PISA 2000 durchziehen, sind Disparitäten. Es gibt systematische Ungleichheiten zwischen Gruppen von Schülerinnen und Schülern, die mit bestimmten Merkmalen einhergehen. Kompetenzunterschiede zwischen Mädchen und Jungen, Kompetenzrückstände bei Jugendlichen mit Zuwanderungshintergrund oder Kompetenzunterschiede in Abhängigkeit von der sozialen Herkunft sind Aspekte, deren Ausmaß eigentlich erst durch ein langfristiges und systematisches Bildungsmonitoring und Benchmarking offenbar wurde. Auch PISA 2015 gibt Hinweise auf Disparitäten innerhalb Deutschlands, die in anderen – auch besonders leistungsstarken – Staaten wesentlich niedriger ausgeprägt sind.

Eine der Hauptbotschaften ist daher auch diesmal, dass weiterhin aktiv darauf hingearbeitet werden muss, einer systematischen Benachteiligung aufgrund welcher Merkmale auch immer möglichst entgegenzuwirken. Dabei kann es nicht das – allzu schlichte – Ziel sein, möglichst viele Kinder auf das Gymnasium zu schicken. Vielmehr muss gewährleistet werden, dass alle Schülerinnen und Schüler, gleich welcher Herkunft und welchen Geschlechts, bestmöglich gefördert werden und eine hohe Durchlässigkeit des Bildungssystems erreicht wird.

Blickt man zurück, so waren es bei PISA 2000 neben den unterdurchschnittlichen Kompetenzen der Jugendlichen in Deutschland gerade die Disparitäten, die zum sogenannten PISA-Schock führten. In Bezug auf die mittleren Leistungen zeigten sich durchgängig Diskrepanzen zwischen den Bundesländern (die bis einschließlich 2006 im Rahmen einer nationalen Erweiterung von PISA dargestellt wurden und seit 2009 mit dem

© Waxmann Verlag GmbH

IQB-Ländervergleich in einer eigenen, von PISA unabhängigen Studie, betrachtet werden), die weiterhin im Ländervergleich zur Überprüfung der Bildungsstandards sichtbar werden (zuletzt 2016). Auch für Jugendliche mit Zuwanderungshintergrund finden sich in allen in PISA untersuchten Domänen in jeder Erhebungsrunde signifikante Nachteile. Allerdings gab es in den letzten beiden Runden 2009 und 2012 positive Nachrichten, da sich die Disparitäten deutlich verringerten. Im Lesen (Hauptdomäne 2009) holten die Schülerinnen und Schüler mit Zuwanderungshintergrund auf, während für Jugendliche ohne Zuwanderungshintergrund kaum Kompetenzsteigerungen zu verzeichnen waren. Auch im Bereich Mathematik (Schwerpunkt 2012) zeichnete sich eine deutliche Annäherung der Leistungen von Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund ab. Während die Jugendlichen mit Zuwanderungshintergrund in PISA 2012 wesentlich bessere Leistungen als in PISA 2003 zeigten, blieb die durchschnittliche mathematische Kompetenz Jugendlicher ohne Zuwanderungshintergrund etwa auf dem gleichen Niveau. In PISA 2015 wird diese Entwicklung leider nicht bestätigt: Im Vergleich der Jahre 2006 und 2015 mit Schwerpunkt Naturwissenschaften zeigt sich weder für die Jugendlichen ohne Zuwanderungshintergrund noch für die Jugendlichen mit Zuwanderungshintergrund insgesamt eine signifikante Verbesserung der naturwissenschaftlichen Kompetenz. Die Disparitäten bleiben damit im Bereich Naturwissenschaften bestehen, was man ganz unabhängig von den genannten Änderungen in PISA 2015 feststellen kann.

Kontinuierlich haben sich in PISA auch Geschlechterdifferenzen gezeigt. Mädchen erzielten in allen sechs Erhebungsrunden signifikant höhere Kompetenzwerte im Lesen und signifikant niedrigere Kompetenzwerte in der Mathematik als Jungen. In den Naturwissenschaften ist das Bild inkonsistent; Geschlechterdifferenzen bestanden entweder nicht oder nur in Bezug auf Teilkompetenzen. In PISA 2015 fällt im Vergleich zu früheren PISA-Erhebungsrunden auf, dass erstmals sowohl im OECD-Durchschnitt als auch in Deutschland die Jungen eine signifikant bessere naturwissenschaftliche Kompetenz erreichen als die Mädchen. Dies geht mit entsprechend unterschiedlich ausgeprägten motivationalen Variablen einher. Die von den Schülerinnen und Schülern selbst berichtete Freude und ihr Interesse an Naturwissenschaften, ihre Selbstwirksamkeit und die instrumentelle Motivation sind Einstellungen, die typischerweise eng mit den erreichten Kompetenzen zusammenhängen. Wenn Jugendliche, wie sich dies in der Stichprobe Deutschlands abzeichnet, wenig begeistert und selbstbewusst an naturwissenschaftlichen Themen arbeiten, so sollten geeignete bildungspolitische Maßnahmen diskutiert werden. In allen diesen Bereichen existieren außerdem signifikante Geschlechterunterschiede zugunsten der Jungen. Im Vergleich zu PISA 2006 zeigt sich dieser Trend besonders unter den Mädchen, während sich die Jungen sich im Vergleich zu 2006 weder bei Freude und Interesse noch bei der Selbstwirksamkeit signifikant verändert haben; für die instrumentelle Motivation gibt es sowohl bei den Jungen als auch bei den Mädchen einen Rückgang. Die naturwissenschaftliche Grundbildung, wie sie in PISA erfasst wird, ist auf einem erfreulich konstanten und überdurchschnittlich hohen mittleren Kompetenzniveau. Dennoch zeigt sich Handlungsbedarf, nämlich einerseits in Bezug auf eine

© Waxmann Verlag GmbH

gezielte Förderung und Talententwicklung und andererseits – hinsichtlich mehrdimensionaler Bildungsziele – in Bezug auf eine Stärkung von Interesse, Begeisterung und Forscherdrang bei Schülerinnen und Schülern der Sekundarstufe I. Die in PISA 2015 in den Naturwissenschaften erstmals auffälligen Geschlechterdifferenzen sollten Anlass sein, intensiv an der Aufgabe zu arbeiten, systematische Kompetenz- und Interessensunterschiede zwischen Mädchen und Jungen in den Blick zu nehmen und Maßnahmen für eine Förderung, die das Ausschöpfen von Potenzialen und die Entwicklung von Talenten im Blick hat, forschungsbasiert zu entwickeln und zu implementieren.

Schließlich stehen bei international vergleichenden Bildungsstudien soziale Disparitäten, also Ungleichheiten aufgrund eines unterschiedlichen sozialen Hintergrunds, immer im Zentrum des Interesses der Berichterstattung. Ein gewisses Maß an Kopplung zwischen sozialer Herkunft und dem eingeschlagenen Bildungsweg mag zwar – durch die Funktion der Familie als primäre Sozialisationsinstanz – als schwierig zu überwinden angesehen werden, der internationale Vergleich zeigt aber, dass es durchaus gelingen kann, Bildungswege und -institutionen durchlässiger zu machen. Es gibt Staaten, in denen die Entscheidung etwa für oder gegen weiterführende Schularten, für eine duale Berufsausbildung oder für ein Studium nicht systematisch damit zusammenhängt, wie hoch oder niedrig der sozioökonomische Status einer Familie ist. Bedeutsamer als die oft geäußerte Empfehlung, die Bildungsbeteiligung am Gymnasium in Deutschland zu erhöhen, scheint die Sicherstellung einer möglichst guten Passung zwischen Fähigkeiten und Interessen der Jugendlichen einerseits und einem gut strukturierten, transparenten Angebot für (Aus-)Bildungswege andererseits. Wenn möglichst jede Schülerin und jeder Schüler einen Bildungs- und Berufsweg einschlagen kann, der ihren oder seinen Kompetenzen und Neigungen entspricht, so bedeutet dies auch, dass nicht jedes Akademikerkind studieren wird. Zwar lässt sich bei PISA 2015 – wie schon in früheren Erhebungsrunden – nachweisen, dass der sozioökonomische Status des Elternhauses und die naturwissenschaftliche Kompetenz der Fünfzehnjährigen in keinem OECD-Staat völlig entkoppelt sind, allerdings gibt es in allen sechs erfolgten PISA-Erhebungsrunden eine hohe Variabilität zwischen den Staaten. Der wünschenswerte Befund, eine geringe Kopplung bei gleichzeitig hohem Durchschnittsniveau in der Kompetenz vorzufinden, gelingt in einer Reihe von Bildungssystemen deutlich besser als in Deutschland. Mit Blick auf die bisher vorliegenden PISA-Runden scheint sich Deutschland diesbezüglich am Anfang eines guten Weges zu befinden. So haben sich die Abstände in der Lesekompetenz zwischen den sozialen Lagen über die Zeit verringert, wobei insbesondere Jugendliche aus Arbeiterfamilien profitiert haben. Die Bildungsbeteiligung im Sinne des Gymnasialbesuchs ist jedoch nach wie vor stark an die berufliche Stellung der Eltern gebunden, wenn auch ein leichter Rückgang dieser Disparitäten festzustellen ist. Mit Blick auf die Zukunft könnten parallele Förderangebote im unteren und oberen Leistungsbereich über die Zeit dazu beitragen, dass sich das Ausmaß der Kopplung zwischen sozialem Hintergrund und mittlerem Kompetenzniveau kontinuierlich verringert. Solange dabei die Gruppe der leistungsschwachen Schülerinnen und Schüler stetig abnimmt, wäre diese Entwicklung aus bildungspolitischer Perspektive positiv zu bewerten.

© Waxmann Verlag GmbH

11.3 Spitzenförderung

Bereits nach der fünften Erhebungsrunde, PISA 2012, waren Hinweise auf erhöhten Bedarf für Spitzenförderung und Talententwicklung formuliert worden. Jugendliche an Gymnasien erzielten zwar nach wie vor sehr hohe Kompetenzmittelwerte, konnten allerdings im Gegensatz zu Schülerinnen und Schülern an nicht gymnasialen Schularten keine Leistungssteigerung erreichen. Die positive Entwicklung, die im Rahmen von PISA 2012 sichtbar wurde, war vorwiegend Jugendlichen an nicht gymnasialen Schularten, Jugendlichen mit Zuwanderungshintergrund und Jugendlichen aus wenig privilegierten Elternhäusern zu verdanken. Entsprechend klar war die Empfehlung, insbesondere auf Kompetenzentwicklungen an Gymnasien Wert zu legen, auch weil die Bildungsbeteiligung an dieser Schulart mit mehr als einem Drittel der Fünfzehnjährigen beträchtlich ist.

In PISA 2015 bleibt Deutschlands Rangposition innerhalb der OECD-Staaten für die Naturwissenschaften als Schwerpunkt im Wesentlichen unverändert. Folglich ist es in Deutschland gelungen, die naturwissenschaftliche Kompetenz auf einem hohen Niveau zu konsolidieren. Gleichwohl sind keine besonderen Fortschritte zu erkennen, vielmehr sprechen einige Befunde für relative Schwächen und weitere Möglichkeiten der Leistungssteigerung. So ist zum Beispiel an den Gymnasien im Vergleich zu PISA 2006 ein bedeutsamer Rückgang der naturwissenschaftlichen Kompetenz zu beobachten. Für die nicht gymnasialen Schularten zeigt sich diese Entwicklung nicht. Daraus kann man schließen, dass Deutschland das Potenzial des differenzierten Schulsystems zur besseren Förderung von Talenten in der naturwissenschaftlichen Bildung weiterhin nur unzureichend ausschöpft.

Auch wenn belastbare Daten in PISA 2015 vor allem zur naturwissenschaftlichen Grundbildung erhoben wurden, weisen die Ergebnisse für Lesen und Mathematik in eine ähnliche Richtung. In diesen Domänen ist es in den letzten Jahren ebenfalls gelungen, den Anteil sehr schwacher Schülerinnen und Schüler zu verringern. Das ist ein großer Erfolg, auf den ganz besonders die Lehrerinnen und Lehrer in den nicht gymnasialen Schularten stolz sein dürfen. Eine Förderung nach Neigung, Interesse und Begabung sollte sich allerdings genauso an leistungsstarke Jugendliche wenden und sie bestmöglich unterstützen.

Mit ihrer Strategie zur Spitzenförderung hat die KMK vor Kurzem bereits einen wichtigen Impuls gesetzt, dessen Wirkung im Frühjahr 2015 noch nicht nachweisbar sein konnte. Es ist allerdings plausibel anzunehmen, dass die folgenden PISA-Erhebungsrunden 2018, 2021 und 2024 entsprechende Erfolge erkennen lassen werden.

© Waxmann Verlag GmbH

11.4 Fazit

PISA 2015 hat zunächst eine übergreifend gute Nachricht zu verkünden. Deutschland ist einer von nur 15 OECD-Staaten, die in allen drei Domänen signifikant über dem OECD-Durchschnitt liegen. Darüber hinaus zeigt die Gesamtbilanz von PISA 2000 bis PISA 2015 in Deutschland eine einzigartige positive Entwicklung in allen drei Kompetenzbereichen. Im Umkehrschluss heißt das, dass PISA bei den Jugendlichen während der letzten 15 Jahre in sehr vielen OECD-Staaten signifikante Schwankungen des Kompetenzniveaus festgestellt hat. Diese Trendaussagen sind für alle Staaten von Bedeutung, geben sie doch insbesondere über längere Zeiträume hinweg Anhaltspunkte dafür, ob und wie staatliche Maßnahmen in Bezug auf Bildung die gewünschte Wirkung zeigen. In einer globalisierten Welt kann die Basis dafür nicht ausschließlich auf nationaler Ebene gesehen werden; vielmehr liefert erst der internationale Vergleich die Grundlage für die Beschreibung und Einschätzung von Veränderungen. Nur so bekommen die Daten Aussagekraft.

Mit Blick auf die Zukunft ist davon auszugehen, dass beginnend mit PISA 2018 die Trendschätzungen über mehrere Erhebungsrunden wieder stabiler sein werden. Sicherlich sind die Daten aber auch unabhängig davon mit der notwendigen Ruhe zu interpretieren. Es geht nicht um ein paar Punkte nach oben oder unten und genauso wenig um kleine Änderungen der Rangfolge, beides gehört in den Bereich des Messfehlers. Aber über signifikante Änderungen gibt PISA Informationen, die eine sinnvolle Grundlage für politische Entscheidungen sind. Insofern hat die Studie einen konkreten und direkten gesellschaftlichen Nutzen. Verlässliche Informationen erfordern einen gewissen Aufwand, denn nur repräsentative Daten können als Basis für systembezogene Entscheidungen dienen. Dass PISA entsprechend Aussagen über den Bildungsstand Jugendlicher gegen Ende ihrer Pflichtschulzeit erlaubt und Einblicke gibt, inwieweit die bis zum PISA-Testtag kumulativ erworbene Bildung im Sinne eines anwendbaren und anschlussfähigen Kompetenzprofils eine gute Prognose über lebenslanges Lernen erlaubt, ist einer der zentralen Gewinne der Studie. Sind es also die richtigen Fragen, auf die PISA eine Antwort gibt? Wenn man Bildung als wichtigste Ressource einer Gesellschaft begreift, dann kann man tatsächlich nur mit einem „Ja“ antworten. Wir brauchen belastbares Wissen zum Stand unseres Bildungssystems, und genau das liefert PISA regelmäßig und mit internationaler Perspektive.

© Waxmann Verlag GmbH

© Waxmann Verlag GmbH

12 Kompetenzmessung in PISA 2015

Jörg-Henrik Heine, Julia Mang, Lars Borchert, Jens Gomolka,
Ulf Kröhne, Frank Goldhammer & Christine Sälzer

PISA 2015 setzt bei der Auswahl der Stichprobe von Jugendlichen sowohl im nationalen als auch im internationalen Kontext auf die bewährten Verfahren vorheriger Runden. Auch die Prinzipien der Erstellung von Aufgaben sind im Wesentlichen gleich geblieben. Insbesondere in methodischer Hinsicht gibt es im Rahmen der Studie allerdings verschiedene Neuerungen, die bei der Interpretation der Ergebnisse zu beachten sind. Konkret gibt es Änderungen in Bezug auf den Modus der Erhebung der Daten, die Ziehung der Stichprobe, die Aufbereitung des Testmaterials zur Kompetenzmessung und die statistische Auswertung der Daten anhand eines psychometrischen Modells. Bei der quasi-längsschnittlichen Interpretation der Ergebnisse vor dem Hintergrund vergangener PISA-Runden müssen diese Aspekte entsprechend bedacht werden. Zwei neue Aspekte, die hervorgehoben werden sollten, sind der Einsatz des Computers bei der Bearbeitung von Testaufgaben und die Anwendung des 2-Parameter-Logistischen-Modells (2-PL-Modell) bei der Skalierung und Auswertung der Daten. Mit der Einführung der computerbasierten Datenerhebung ist eine veränderte Zusammenstellung des Testmaterials verbunden. Zur Bestimmung der Schwierigkeit der PISA-Aufgaben (Itemkalibrierung) wurde bei der aktuellen Runde mit der sogenannten *Concurrent Item Calibration* ein gegenüber früheren Runden anderes Verfahren eingesetzt.

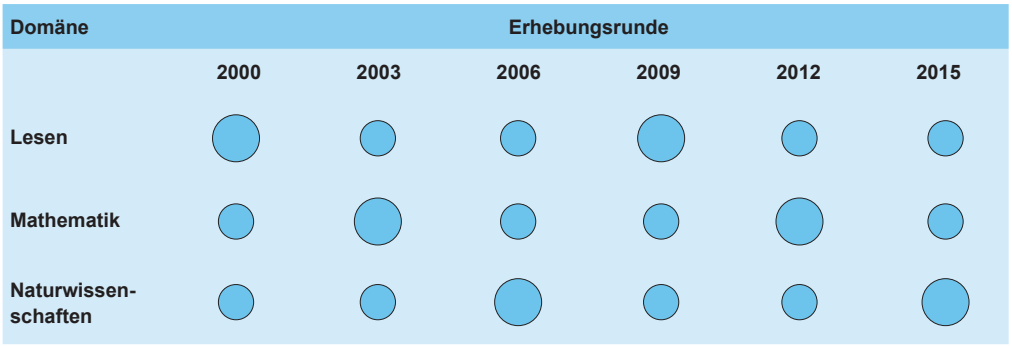
12.1 Einleitung und Überblick

Das *Programme for International Student Assessment* (PISA) blickt mittlerweile auf fünf *querschnittlich* angelegte Erhebungsrunden mit internationalen Vergleichen zurück. Mit der in diesem Band vorgestellten sechsten Erhebungsrunde liegen nun zwei vollständige PISA-Zyklen vor, wobei jeder Zyklus aus drei Erhebungsrunden besteht. Bei jeder Runde wurde der Schwerpunkt jeweils auf einen der drei PISA-Kompetenzbereiche (Domänen, vgl. Abbildung 12.1) Lesen, Mathematik und Naturwissenschaften gesetzt.

Die technischen Standards und die grundlegende Konzeption zur praktischen Durchführung der PISA-Studien wurden in ihren wesentlichen Aspekten bei der ersten Erhebungsrunde (PISA 2000) festgelegt. Für die in den Jahren 2003, 2006, 2009 und 2012 folgenden PISA-Erhebungsrunden wurde diese grundlegende Konzeption, nicht

© Waxmann Verlag GmbH

zuletzt wegen praktischer Überlegungen vor dem Hintergrund der einfachen Durchführbarkeit von Trendbetrachtungen, jeweils nur geringfügigen Modifikationen unterzogen. Diese Konzeption und die damit verbundenen technischen Standards integrierten so den jeweils aktuellen Stand der psychometrischen Forschung zur Kompetenzmessung zum Zeitpunkt der jeweiligen PISA-Runde und nahmen gleichzeitig Rücksicht auf die zu Beginn gesetzte Konzeption der PISA-Studien.



Anmerkung: Große Kreise = Schwerpunkt-Domäne; kleine Kreise = Neben-Domäne

Abbildung 12.1 Übersicht der inhaltlichen Schwerpunktsetzung in den PISA-Runden seit 2000

Nachdem die erste technische Konzeption der PISA-Studien inzwischen mehr als eineinhalb Jahrzehnte zurückliegt, war es für die sechste PISA-Erhebung 2015 an der Zeit, Änderungen an der Basiskonzeption in diesem Bereich vorzunehmen. Dies ist eine Folge des wissenschaftlichen Erkenntnisgewinns im Bereich internationaler Bildungsvergleichsstudien sowie des technischen Fortschritts bei der Nutzung von Computertechnologien für die Datenerhebung. Die Anpassungen in der technischen Konzeption haben allerdings Konsequenzen, mit denen bei der praktischen Durchführung der Erhebung sowie bei der anschließenden Datenauswertung sorgfältig umgegangen werden muss. So ergeben sich etwa Änderungen

- bei den organisatorischen Abläufen,
- in Bezug auf die Datenstrukturen,
- bei den zugrunde liegenden *psychometrischen Modellen* zur international vergleichenden und übergreifenden Kalibrierung der Testaufgaben sowie zur Modellierung der *Kompetenzniveaus* der Schülerinnen und Schüler.

Konkret beziehen sich die technischen Weiterentwicklungen zur Durchführung und Auswertung der PISA-Studien zunächst auf Aspekte der Ökonomie und Effizienz der *Datenerhebung* (z. B. computerbasierte Testung) sowie deren Weiterverarbeitung. Hinzu kommen bessere Möglichkeiten zur Datenauswertung, indem neben Ergebnisdaten auch Prozessdaten gesammelt werden, die während der Testbearbeitung am Computer anfallen (z. B. Goldhammer, Naumann, Rölke, Stelter & Tóth, im Druck; Lee & Chen, 2011;

© Waxmann Verlag GmbH

vgl. auch Abschnitt 12.3). Bei der Auswertung der Daten aus PISA 2015 – im Vergleich zu den Erhebungen vergangener Jahre – werden *erweiterte psychometrische Skalierungsmodelle* eingesetzt, deren Anwendung auf Daten aus *Large-Scale-Assessments* (LSA) in der psychometrischen Literatur (auch kontrovers) diskutiert wird (z. B. Adams, 2011; Kreiner, 2011; Kreiner & Christensen, 2014; Macaskill, 2008; Oliveri, Olson, Ercikan & Zumbo, 2012). Hier wurden speziell die Daten aus vorangegangenen PISA-Erhebungsrunden im Rahmen der Bestimmung der Charakteristika der Testaufgaben vom internationalen Konsortium (ETS, vgl. Kapitel 1) einer Reanalyse unterzogen. Sie zeigte die bessere Passung dieser erweiterten Skalierungsmodelle bei der Auswertung der Daten (ETS, 2015a; OECD, 2013; Oliveri & von Davier, 2014).

Die Veränderungen aufgrund der technischen Entwicklungen gehen organisatorisch mit dem Wechsel beziehungsweise der Vergabe der Durchführung der internationalen Datenauswertung an ein neu zusammengesetztes internationales Konsortium einher. Einzelne Akteure der ersten Stunde erhielten dabei erneut den Zuschlag für die Mitwirkung an PISA, andere sind in veränderter Rolle tätig oder neu im Projekt (vgl. Kapitel 1).

12.2 Population und Stichprobe der PISA-Teilnehmer in Deutschland

In PISA wird – wie in ähnlichen Bildungsstudien auch – eine systematische Teilerhebung realisiert. Dies geschieht durch Ziehung einer geeigneten *Stichprobe* von fünfzehnjährigen Schülerinnen und Schülern. Diese erfolgt nach exakten statistischen Regeln, sodass anhand der Stichprobe Verallgemeinerungen über die Grundgesamtheit, das heißt über alle fünfzehnjährigen Schülerinnen und Schüler in Deutschland, möglich sind (vgl. Bortz & Schuster, 2010; Brown, 2010; Häder, 2015; Kish, 1995; Levy & Lemeshaw, 2008; Thompson, 2012).

12.2.1 Populationsdefinitionen, Stichprobenplan und Ziehung der Stichproben

Bildungsvergleichsstudien – insbesondere *Large-Scale-Assessments* – stellen besondere Anforderungen an ihre Stichproben. So kommt im Allgemeinen keine *einstufige Zufallsauswahl* zum Einsatz (z. B. direkte Ziehung einzelner Schülerinnen und Schüler aus der Grundgesamtheit aller Schüler), sondern es werden zwei- oder mehrstufige Auswahlverfahren verwendet. In den meisten Teilnehmerstaaten wurde ein zweistufiges Ziehungsverfahren für die Stichprobe eingesetzt,[1] bei dem die Schulen die erste Stichprobenein-

1 Eine genaue Beschreibung der in den bisherigen PISA-Erhebungsrunden verwendeten Methodologie kann den sogenannten *Technical Reports* entnommen werden. Diese finden sich auf der Website der OECD – unter: http://www.oecd.org/pisa/pisaproducts/.

© Waxmann Verlag GmbH

heit und Schülerinnen und Schüler die zweite, also die eigentliche Untersuchungseinheit darstellen. Für Deutschland trifft diese generelle PISA-Vorgehensweise der Stichprobengewinnung nur bedingt zu, da aufgrund eines erweiterten Studiendesigns auch ein komplexeres Stichprobenziehungsverfahren erforderlich wurde. Dies wird im Folgenden dargestellt.

Wie in allen bisherigen PISA-Runden besteht auch bei PISA 2015 die international vorgegebene Zielpopulation aus allen Schülerinnen und Schülern einer *Alterskohorte*. Es handelt sich um alle Fünfzehnjährigen, die sich in der 7. Klassenstufe oder einer höheren Klassenstufe befinden. Die genaue Altersdefinition wurde hierbei mit dem internationalen PISA-Konsortium abgestimmt und kann bedingt durch den Erhebungszeitraum in den einzelnen Staaten, die an PISA teilnehmen, leicht variieren. Für Deutschland wurden alle Schülerinnen und Schüler der Alterskohorte der Fünfzehnjährigen zugerechnet, die zwischen dem 1. Januar und dem 31. Dezember 1999 geboren wurden.

Formale Definition der Schüler-Zielpopulation:
Alle Schülerinnen und Schüler („Fünfzehnjährige“), die im Zeitraum 1. Januar 1999 bis 31. Dezember 1999 (einschließlich) geboren sind und mindestens die 7. Klassenstufe besuchen.

Die aus dieser Alterskohorte stammenden Schülerinnen und Schüler bilden die *Grundgesamtheit* für die spätere Stichprobenziehung und sind zugleich Teil der PISA-Grundstichprobe, die in jedem an der Studie teilnehmenden Staat gezogen wurde. Für Deutschland wurde diese Grundstichprobe erweitert – insbesondere, um vertiefende Analysen durchführen und weitere Forschungsfragen bearbeiten zu können (z. B. Fragestellungen zum Klassenkontext, Auswertungen zu den Perspektiven der Lehrkräfte und der Schülerinnen und Schüler). Dabei wurden zusätzliche Schülerinnen und Schüler der 9. Jahrgangsstufe in die Stichprobe einbezogen: So wurde in allgemeinbildenden Schulen zunächst eine 9. Klasse per Zufall gezogen, aus der wiederum 15 Schülerinnen und Schüler zufällig für die Teilnahme an PISA ermittelt wurden.[2] In Förderschulen, in denen generell weniger Neuntklässler anzutreffen sind und in denen eine übliche Klassenstruktur nicht immer gegeben ist, wurden zunächst alle Neuntklässler berücksichtigt, aus denen dann die 15 zusätzlichen Schülerinnen und Schüler zufällig ausgewählt werden konnten. Diese Zusatzschüler des 9. Jahrgangs können entsprechend aus verschiedenen Altersgruppen (Fünfzehnjährige und Nicht-Fünfzehnjährige) stammen. Das Ziehungskriterium ist entweder, dass sie zu einer gezogenen 9. Klasse gehören (allgemeinbildende Schulen) oder dass sie der 9. Jahrgangsstufe zugeordnet werden können

2 Die Größenordnung ergab sich aus der Notwendigkeit, mit den vorhanden Ressourcen (z. B. Computerressourcen in Klassenzimmern bzw. Laptop-Sets) sparsam umzugehen. Aus statistischer Sicht sind 15 zufällig gezogene Schülerinnen und Schüler für die Zusatzerhebung „einer“ 9. Klasse ausreichend. Es handelt sich insbesondere um eine Größenordnung, welche in späteren Analysen die sinnvolle Berücksichtigung bzw. Untersuchung der Datenstruktur im Rahmen eines Multilevel-Ansatzes noch zulässt (z. B. Luke, 2004).

© Waxmann Verlag GmbH

(Förderschulen). An den beruflichen Schulen blieb – im Gegensatz zu den allgemeinbildenden Schulen und Förderschulen – das Studiendesign unverändert, die Stichprobe wurde hier also nicht erweitert. Es sei betont, dass der vorliegende PISA-Bericht ausschließlich die Ergebnisse zur PISA-Grundstichprobe der Fünfzehnjährigen ausweist, die Daten zu den zusätzlich gezogenen Neuntklässlern werden in weiteren Forschungsarbeiten analysiert.

Neu im Vergleich zu den vorherigen Erhebungsrunden ist bei PISA 2015, dass als nationale Zusatzoption auch eine Stichprobe von Lehrerinnen und Lehrern gezogen werden konnte.[3] Deutschland nahm als einer von 18 Staaten an dieser zusätzlichen Lehrerbefragung teil. Im Mittelpunkt des Forschungsinteresses standen dabei Lehrkräfte, die in Klassenstufen mit einem hohen Anteil von fünfzehnjährigen Schülerinnen und Schülern unterrichten. Dies sind in Deutschland die 9. und 10. Klassenstufe (Statistisches Bundesamt & DESTATIS, 2014). Da zudem der Forschungsschwerpunkt bei PISA 2015 auf dem Kompetenzbereich Naturwissenschaften lag (OECD, 2016a), nahmen insbesondere Lehrkräfte, aus dem Bereich der Naturwissenschaften teil.[4] Generell wurde die Zielpopulation der Lehrkräfte wie folgt definiert:

Formale Definition der Lehrer-Zielpopulation:
Die Lehrerbefragung ist an alle Lehrkräfte (inkl. Referendarinnen und Referendare) gerichtet, die eine Lehrbefähigung besitzen, um Schülerinnen und Schüler der 9. Jahrgangsstufe und/oder der 10. Jahrgangsstufe zu unterrichten – ungeachtet dessen, ob eine Lehrkraft dies aktuell tut, jemals getan hat oder künftig tun wird oder könnte. Vollzeit- sowie Teilzeitlehrkräfte, angestellte und verbeamtete Lehrerinnen und Lehrer sind dabei gleichermaßen zu berücksichtigen. Auch Lehrerinnen und Lehrer, die ihre Lehrtätigkeit an mehreren verschiedenen Schulen ausüben, sind für die Listung vorgesehen.

Als Lehrkraft gilt dabei eine Person, deren vorrangige oder hauptsächliche Aktivität in der Schule die Ausbildung von Schülerinnen und Schülern ist und die den Schülerinnen und Schülern Unterrichtsstunden erteilt. Lehrerinnen und Lehrer können mit den Schülerinnen und Schülern im ganzen Klassenverband in Klassenräumen arbeiten, in Kleingruppen oder im Einzelunterricht inner- oder außerhalb der regulären Klassenräume.

3 In Deutschland wurden bei allen vorhergehenden PISA-Zyklen auch Lehrkräfte befragt, jedoch war dies eine rein nationale Zusatzerhebung. In PISA 2012 wurde darüber hinaus keine spezifische Zielpopulation definiert. Vielmehr wurden alle Mathematiklehrkräfte (d. h. auch fachfremde und teilzeitbeschäftigte) ab der Jahrgangsstufe 5 sowie das gesamte sonderpädagogische Personal für Schülerinnen und Schüler mit besonderem Förderbedarf (ebenfalls ab der 5. Jahrgangsstufe) berücksichtigt.

4 Laut internationaler PISA-Definition wurden der Domäne „Naturwissenschaften“ die folgenden Fächer zugeordnet: Physik, Chemie, Biologie, Geologie, Astronomie, angewandte Naturwissenschaften und Technik oder Fächer, die innerhalb eines einzelnen allgemeinen, integrierten oder fächerübergreifenden Unterrichts unterrichtet werden (z. B. Natur und Technik).

© Waxmann Verlag GmbH

Von allen Lehrkräften einer Schule, die unter die oben beschriebene Definition fielen, wurden zwei Gruppen per Zufall ausgewählt: Eine Gruppe von 15 Lehrkräften, die gegenwärtig ein naturwissenschaftliches Fach unterrichten, sowie eine Gruppe von 20 sonstigen Lehrkräften. Das erweiterte Forschungsdesign für PISA in Deutschland ist entsprechend mit einem relativ komplexen mehrstufigen Ziehungsverfahren für die Stichproben verbunden, das in Abbildung 12.2 dargestellt ist.

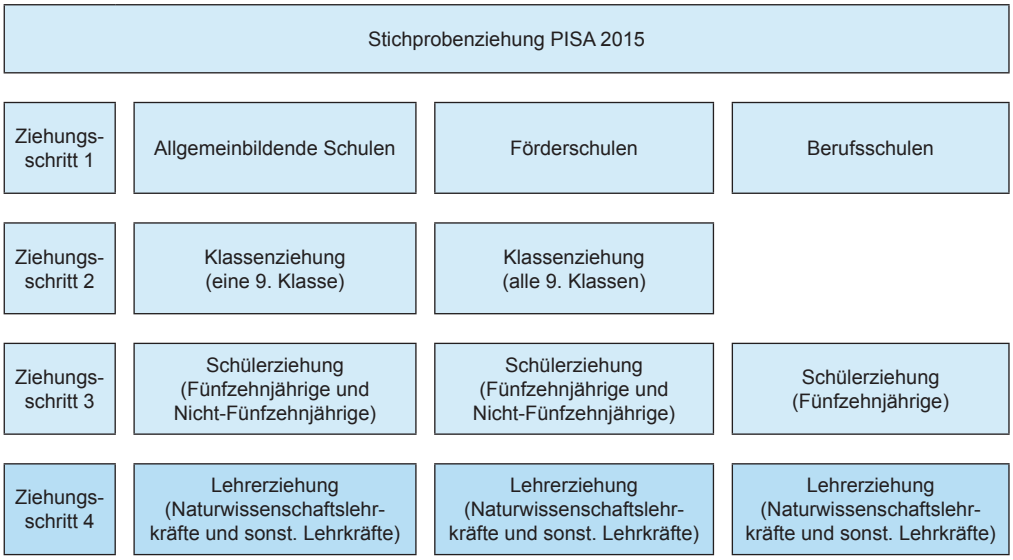


Abbildung 12.2: Vier Schritte der PISA-Stichprobenziehung

Schritt 1: Vorbereitung der Schulstichprobenziehung – vollständige Schulliste

Als Grundlage für die Stichprobenziehung der Schulen dient der sogenannte *Sampling Frame* (Stichprobenrahmen). Hier sind alle Schulen in Deutschland aufgelistet, an denen sich potenziell fünfzehnjährige Schülerinnen und Schüler befinden. Als Schulen werden dabei in der Regel ganze Schulen verstanden, weniger dagegen einzelne Schulprogramme oder Bildungsgänge. Für Deutschland orientiert sich die Auflistung sämtlicher Schulen im *Sampling Frame* an der Definition der Kultusministerkonferenz (KMK) von Schule als „schulartspezifische Einrichtung". Demnach müssen Schulen „keine verwaltungsrechtlich eigenständigen Organisationseinheiten" bilden (KMK, 2015, S. 8). „Vielmehr werden schulartspezifische Schulteile, die im Hinblick auf Lehrpläne, Bildungsziel beziehungsweise Qualifikationsniveau einen eigenständigen Charakter haben, als Schulen gezählt" (KMK, 2015, S. 8). Gemäß dieser Definition ist es beispielsweise möglich, dass eine Schule doppelt im *Sampling Frame* gelistet wurde, wenn an der betreffenden Schule zwei Schularten vorliegen: Zum Beispiel dann, wenn es sich bei einer Schule um eine nominelle Haupt- und Realschule oder um eine kooperative Gesamtschule mit gymnasialer Oberstufe handelt; nicht aber, wenn es um eine Schule mit mehreren Bil-

© Waxmann Verlag GmbH

dungsgängen geht, denn diese zählen als eine einzige Schule mit Klassen, in denen Schülerinnen und Schüler unterschiedlicher Bildungsgänge gemeinsam unterrichtet werden.

Um den *Sampling Frame* nach den Vorgaben des internationalen PISA-Konsortiums erstellen zu können, wurden sämtliche hierfür benötigten Schulinformationen bei den 14 Statistischen Landesämtern eingeholt.[5] Bei diesen Informationen handelt es sich um Listen der Schulen (allgemeinbildende Schulen, Förderschulen, Berufsschulen) im jeweiligen Land, die folgende Angaben enthalten:

- die offizielle Schulnummer (die später im *Sampling Frame* pseudonymisiert wurde),
- die Schulart,
- die Anzahl der Schülerinnen und Schülern in den Geburtsjahrgängen 1998, 1999, 2000,
- die Anzahl der Schülerinnen und Schüler in den 7. bis 10. Klassenstufen,
- die Anzahl der 7. bis 10. Klassen,
- die Trägerschaft (öffentlich oder privat),
- Informationen über Veränderungen der Schulart, Schulzusammenlegungen und Schulschließungen sowie
- für Förderschulen die Informationen über die Förderschwerpunkte gemäß den Vorgaben der Kultusministerkonferenz (KMK), wobei in Anlehnung an alle vorhergehenden PISA-Erhebungsrunden die Förderschwerpunkte Lernen, Sprache sowie emotionale und soziale Entwicklung berücksichtigt wurden.

Datengrundlage für sämtliche der oben genannten Schulinformationen war die Schulstatistik des jeweiligen Landes für das Schuljahr 2013/2014. Falls diese für eine oder mehrere Informationen nicht verfügbar war, wurde auf die jeweils zuletzt veröffentlichten Daten zurückgegriffen.[6] In den *Sampling Frame* selbst wurden später allerdings nicht alle der bei den Statistischen Landesämtern erfragten Informationen mit aufgenommen. So dienten einige Informationen lediglich zur Sicherung der Datenqualität.

PISA wendet zur Zufallsziehung der Schulen das sogenannte PPS-Verfahren (*Probabilities Proportional to Size*) an. Hierbei wird die Ziehungswahrscheinlichkeit umgekehrt proportional zur Schulgröße festgesetzt; große Schulen haben somit eine erhöhte Wahrscheinlichkeit, gezogen zu werden. Umgekehrt haben Schülerinnen und Schüler innerhalb großer Schulen eine kleine Wahrscheinlichkeit, für die Studie ausgewählt zu werden. Diese Methode führt zu geringen Varianzen in den Stichprobengewichten und trägt somit zu niedrigen Standardfehlern bei. Um dieses Ziehungsverfahren anwenden zu können, muss der *Sampling Frame* ein sogenanntes *Measure of Size* (MOS) aufweisen.

5 Es gibt in Deutschland insgesamt 14 Statistische Landesämter, da für Berlin und Brandenburg sowie für Hamburg und Schleswig-Holstein jeweils ein Landesamt zuständig ist.

6 Ausnahmen bilden die allgemeinbildenden Schulen in Mecklenburg-Vorpommern sowie die Berufsschulen in Baden-Württemberg, für die auf die Schulstatistik von 2011/12 zurückgegriffen werden musste. Weitere Ausnahmen bilden die allgemeinbildenden Schulen und Berufsschulen im Saarland sowie die Berufsschulen in Niedersachen. Für diese wurden die Schulstatistiken von 2012/13 verwendet.

© Waxmann Verlag GmbH

Für die PISA-Alterskohorte (fünfzehnjährige Schülerinnen und Schüler) war dies die jeweilige Schülerzahl pro Schule des Geburtsjahrganges 1998.[7] Dieser Wert wurde dann mit den Schüleranzahlen der Geburtsjahrgänge 1999 und 2000 abgeglichen, um Schwankungen in den Geburtenzahlen zu identifizieren. Darüber hinaus war es aufgrund des erweiterten Forschungsdesigns in Deutschland zusätzlich erforderlich, die Anzahl der 2015 erwarteten Neuntklässler im *Sampling Frame* zu listen. Hierfür wurde die Anzahl der jeweiligen Neuntklässlerinnen und Neuntklässler im Schuljahr 2013/2014 verwendet, wobei die Schülerzahlen der 8. bis 10. Klassenstufen sowie die Anzahl der dazugehörigen Klassen überprüft wurden, um mögliche Inkonsistenzen aufzudecken und entsprechend auszugleichen.

Um die Qualität der notwendigen Informationen für den *Sampling Frame* und damit für die Ziehung der Schulstichprobe zu optimieren, wurden sämtliche Schuldaten der Statistischen Landesämter noch einmal mit den Daten des Statistischen Bundesamts, den Daten der Fachserie 11 (Statistisches Bundesamt & DESTATIS, 2014), abgeglichen. Sämtliche Auffälligkeiten wurden dann in Rücksprache mit dem entsprechenden Statistischen Landesamt geklärt.

Schritt 2: Ziehung der Schulstichprobe

Bevor die Stichprobenziehung durchgeführt werden konnte, mussten die Schulen im *Sampling Frame* nach bestimmten Kriterien, den sogenannten Stratifizierungsvariablen, geschichtet werden (vgl. Daniel, 2012; Kalton, 1983; Kish, 1995; Levy & Lemeshaw, 2008). Dies bedeutet, dass mittels der *Stratifizierung* alle gelisteten Schulen (gemäß ihrer Ähnlichkeit) in Gruppen aufgeteilt wurden. Ziel dieser Gruppierung ist es, mit einer möglichst geringen Stichprobengröße möglichst genaue Populationsschätzungen vornehmen zu können. Des Weiteren wird mittels Stratifizierung sichergestellt, dass die Verteilung der Schulen innerhalb der Stichprobe die Verteilung in der Grundgesamtheit widerspiegelt und somit zugleich spezifische Teilpopulationen (z. B. bestimmte Schularten) in der Stichprobe in adäquater Weise repräsentiert werden. Dadurch wird die Effizienz des Sampling Designs verbessert, wodurch die Schätzer für die unbekannten Parameter (z. B. Kompetenzmittelwerte) in der Grundgesamtheit (fünfzehnjährige Schülerinnen und Schüler in Deutschland) – im Vergleich zu einer ungeschichteten Stichprobe derselben Größe – verlässlicher werden.

Die Stratifizierung kann explizit und implizit vorgenommen werden. Bei der expliziten Stratifizierung werden die Schulen in einzelne Gruppen beziehungsweise Strata aufgeteilt, die unabhängig voneinander behandelt werden und aus denen jeweils eine sepa-

7 Der Geburtsjahrgang 1998 wurde als Schätzer für die Anzahl der 2015 erwarteten Fünfzehnjährigen gewählt, da man die Anzahl der Schülerinnen und Schüler an den allgemeinbildenden Schulen überschätzen würde, wenn man auf den Geburtsjahrgang 1999 zurückgreift: Da in einigen Bundesländern beispielsweise viele Fünfzehnjährige die Hauptschule verlassen, um berufsbildende Schulen zu besuchen, eignen sich die Vierzehnjährigen (das wäre hier der Jahrgang 2000) nicht als Schätzer für die Fünfzehnjährigen des nächsten Schuljahres.

© Waxmann Verlag GmbH

rate Zufallsauswahl getroffen wird. Die implizite Stratifizierung dient dazu, die Schulen noch einmal innerhalb der expliziten Strata zu sortieren, um eine näherungsweise proportionale Verteilung der Schulen über sämtliche Strata innerhalb der Stichprobe zu gewährleisten.

Für Deutschland kamen bei PISA 2015 zwei *explizite* und zwei *implizite* Stratifizierungsvariablen zum Einsatz:

- Einteilung der Schulen in allgemeinbildende Schulen, Förderschulen und Berufsschulen (erste explizite Stratifizierungsvariable)
- Gruppierung der *allgemeinbildenden Schulen* nach Ländern (zweite explizite Stratifizierungsvariable)
- Gruppierung der *Förder- und Berufsschulen* nach Ländern (erste implizite Stratifizierungsvariable)
- Einteilung der *allgemeinbildenden Schulen* in die Schularten Hauptschule, Integrierte Gesamtschule, Realschule, Schule mit mehreren Bildungsgängen und Gymnasium (zweite implizite Stratifizierungsvariable)

Damit lässt sich das Stichprobendesign für Deutschland folgendermaßen beschreiben: Zunächst werden die zu ziehenden Schulen nach erster und zweiter Stratifizierungsvariable eingeteilt (= explizite Stratifizierung). Dann erfolgt die Stichprobenziehung in der Art und Weise, dass sich die gezogenen allgemeinbildenden Schulen innerhalb der 16 einzelnen Länder näherungsweise proportional zur Gesamtzahl der Schülerinnen und Schüler je Schulart auf die verschiedenen Schularten verteilen (= implizite Stratifizierung). Förder- und Berufsschulen bilden zwei weitere explizite Strata. Das heißt, dass diese beiden Schularten stichprobentechnisch gesondert behandelt werden und daher auch unabhängig von den allgemeinbildenden Schulen gezogen werden. Der Grund hierfür ist, dass die relative Häufigkeit dieser beiden Schularten innerhalb der Bundesländer unterschiedlich ist. Des Weiteren befindet sich in der Stichprobe eine so geringe Anzahl von Schülerinnen und Schülern dieser beiden Schularten, dass diese bereits in der expliziten Stratifizierung beachtet werden müssen. Damit wird die Population der Schulen in Deutschland, die potenziell von fünfzehnjährigen Schülerinnen und Schülern besucht werden, in insgesamt 18 explizite Strata unterteilt.

Durch die implizite Stratifizierung nach Ländern wird dann gesichert, dass sich die zu ziehenden Förder- und Berufsschulen ebenfalls näherungsweise proportional zur Gesamtzahl der Schülerinnen und Schüler dieser beiden Schularten je Land verteilen.

Im Anschluss an die Stratifizierung wird die Schulstichprobe für Deutschland gezogen. Hierbei muss die Verteilung der Gesamtzahl der gezogenen Schulen auf die einzelnen Strata so erfolgen, dass der erwartete Anteil der Schülerinnen und Schüler in jedem dieser Strata in der Stichprobe näherungsweise dem proportionalen Anteil der Schülerinnen und Schüler (PISA-Alterskohorte) in jedem Stratum der Grundgesamtheit entspricht. Insgesamt wurden 267 Schulen in Deutschland gezogen. Tabelle 12.1 zeigt die Anzahl gezogener Schulen in der Stichprobe in den einzelnen Strata.

© Waxmann Verlag GmbH

Tabelle 12.1: Bruttostichprobe der Schulen nach Bundesland und Schulart

	HS	IG	MBG	RS	GY	FS	BS	Σ
Baden-Württemberg	8	0	0	13	13	3	4	**41**
Bayern	10	1	0	14	12	1	4	**42**
Berlin	0	4	0	1	3	1	0	**9**
Brandenburg	0	0	3	0	3	0	0	**6**
Bremen	0	1	0	0	1	0	0	**2**
Hamburg	0	2	0	0	2	0	1	**5**
Hessen	2	4	0	5	6	2	1	**20**
Mecklenburg-Vorpommern	0	0	2	0	1	0	0	**3**
Niedersachsen	5	2	1	10	9	2	1	**30**
Nordrhein-Westfalen	10	10	1	15	21	4	1	**62**
Rheinland-Pfalz	0	2	6	0	4	0	1	**13**
Saarland	0	0	1	1	1	1	1	**5**
Sachsen	0	0	5	0	3	0	0	**8**
Sachsen-Anhalt	0	0	2	0	2	1	1	**6**
Schleswig-Holstein	0	4	1	1	3	0	0	**9**
Thüringen	0	0	3	0	2	0	1	**6**
Gesamt	35	30	25	60	86	**15**	**16**	267

Anmerkung: Hervorgehobene Schulzahlen entsprechen den expliziten Strata; HS: Haupschule, IG: Integrierte Gesamtschule, MBG: Schule mit mehreren Bildungsgängen, RS: Realschule, GY: Gymnasium, FS: Förderschule, BS: Berufsschule, Σ: Summen.

Es ist durchaus möglich, dass in einem Stratum – insbesondere in einem kleinen – nicht alle Schularten gezogen werden und somit in der Schulstichprobe vertreten sind. Sämtliche Schulen hatten aber eine von null verschiedene Ziehungswahrscheinlichkeit.[8] Aufgrund der geringen Repräsentativität einzelner Schularten innerhalb der Bundesländer bildet PISA auch nicht den richtigen Rahmen für Vergleiche zwischen den Bundesländern. Hierfür muss auf den Ländervergleich und fundierte Analysen auf Basis der Bundesländer verwiesen werden (Stanat, Böhme, Schipolowski & Haag, 2016).

Schritt 3 und 4: Ziehung der Schüler- und Lehrerstichprobe

Nach der Bestimmung der Schulen, an denen die Datenerhebung durchgeführt werden sollte, wurden mittels einer weiteren Stichprobenziehung die einzelnen Schülerinnen und Schüler innerhalb dieser Schulen ausgewählt. Laut Stichprobendesign war dabei vorgesehen, in jeder Schule 30 Fünfzehnjährige per Zufall zu ziehen. Darüber hinaus wurde – wie bereits oben ausgeführt – an allen allgemeinbildenden Schulen zunächst eine vollständige 9. Klasse zufällig ausgewählt, aus der dann wiederum zufällig 15 Fünf-

8 Ausführliche Beispiele zur Schulstichprobenziehung sowie die dazugehörigen Ziehungsalgorithmen können den PISA *Technical Reports* unter: http://www.oecd.org/pisa/pisaproducts/entnommen werden.

© Waxmann Verlag GmbH

zehnjährige für die PISA-Studie selektiert wurden. An den Förderschulen erfolgte zunächst eine Vollerhebung aller Neuntklässlerinnen und Neuntklässler, aus denen dann ebenfalls zufällig 15 Fünfzehnjährige gezogen wurden – falls möglich aus einer 9. Klasse. Gab es jedoch keine 9. Klasse mit mindestens 15 Fünfzehnjährigen an einer Schule, so wurden 15 zufällig ausgewählte Neuntklässlerinnen und Neuntklässler berücksichtigt.

In Deutschland wurden außerdem an jeder der Schulen, die an PISA teilnehmen, Lehrkräfte für eine Befragung ausgewählt. Das vollständige Stichprobendesign für Deutschland kann Abbildung 12.3 entnommen werden.

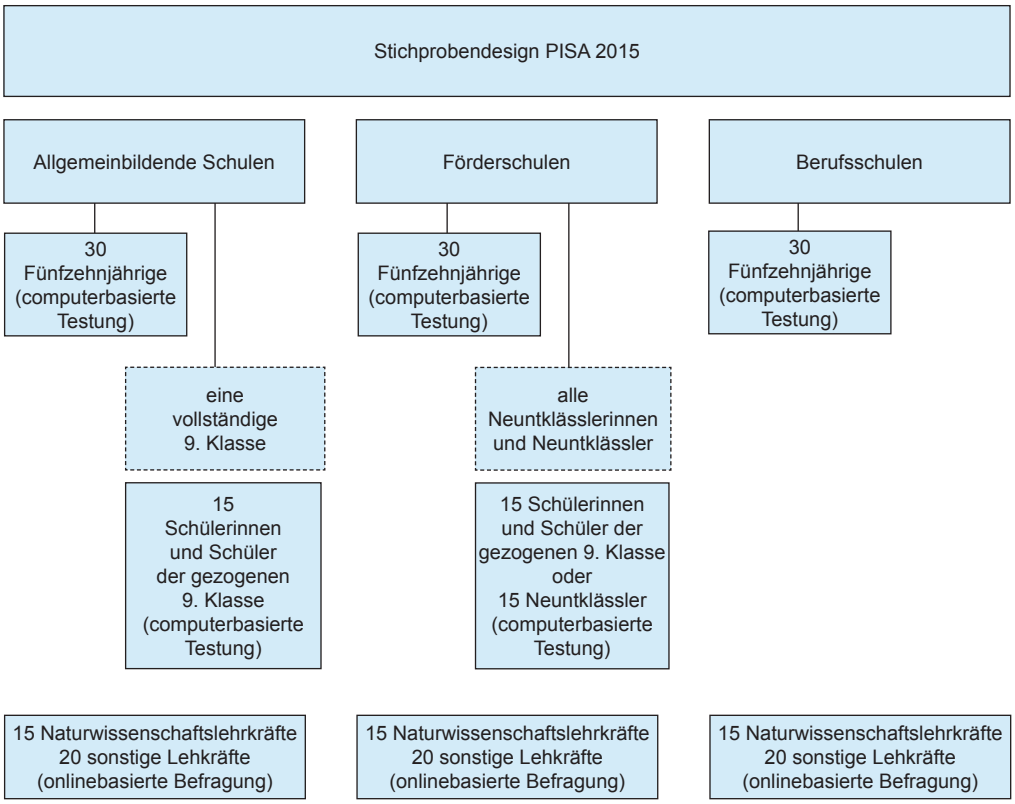


Abbildung 12.3: Stichprobendesign PISA 2015 Deutschland

An den allgemeinbildenden Schulen und Förderschulen ist es möglich, dass von den 30 gezogenen Fünfzehnjährigen (Alterskohorte: Geburtsjahrgang 1999) vereinzelt Schülerinnen und Schüler auch Teil der gezogenen 9. Klasse sind (aber der Testgruppe der Fünfzehnjährigen zugerechnet werden). Auf der anderen Seite kann es vorkommen, dass Schülerinnen und Schüler sowohl der gezogenen 9. Klasse als auch unterschiedlichen Geburtsjahrgängen angehören. Somit können sie zwar Bestandteil des Geburtsjahrgangs 1999 sein (der Gruppe der Fünfzehnjährigen), wurden aber der Testgruppe der Neunt-

© Waxmann Verlag GmbH

klässlerinnen und Neuntklässler zugeordnet.[9] Sobald nämlich 30 Fünfzehnjährige für die Testgruppe der Fünfzehnjährigen gezogen waren, wurden die restlichen Fünfzehnjährigen – die zu der gezogenen 9. Klasse gehören – nur der Testgruppe der entsprechenden 9. Klasse zugeordnet.

Um die beschriebene Schüler- und Lehrerstichprobe umsetzen zu können, mussten in den gezogenen Schulen Merkmale wie das Geburtsjahr, das Geschlecht, die Klassenstufe oder die Klassenbezeichnung in Listen zusammengestellt werden. Aus Datenschutzgründen verblieben personenbezogene Daten wie die Namen der Schülerinnen und Schüler in der Schule. Gelistet wurden alle Schülerinnen und Schüler, die entweder 15 Jahre alt waren (Geburtsjahrgang 1999) oder die zu einer 9. Klasse gehörten (ohne Geburtsjahrgang 1999). An Berufsschulen mussten lediglich die Fünfzehnjährigen gelistet werden. Hierfür wurde von den Schulen das vom *IEA DPC* entwickelte Online-System *IEA OnlineStudyExpert* (kurz: *IEA OSE*) verwendet. Es handelt sich dabei um eine Online-Plattform, über die Informationen zwischen den einzelnen Schulen und dem Datenerhebungsinstitut (*IEA DPC*) ausgetauscht werden können und die strikt alle datenschutzrechtlichen Belange berücksichtigt. Hervorzuheben sind hier die sichere und verschlüsselte Datenübertragung sowie die *Pseudonymisierung* der Namen der Schülerinnen und Schüler sowie der Lehrkräfte.

Die Schülerliste wie auch die Lehrerliste existierte in zwei verschiedenen Versionen. Es gab jeweils eine vollständige Version mit den Namen der Schülerinnen und Schülern bzw. Lehrerinnen und Lehrern, die aus datenschutzrechtlichen Gründen nur von den Schulen eingesehen werden konnte, und jeweils eine pseudonymisierte Liste (für das *IEA DPC*) für Schülerinnen und Schüler bzw. Lehrerinnen und Lehrer, die statt der Namen nur noch Ordnungsnummern enthielt.[10] Nur die beiden pseudonymisierten Listen wurden für die Durchführung der Schüler- und Lehrerstichprobenziehung verwendet.

Waren alle für die PISA-Studie relevanten Informationen an das *IEA DPC* übermittelt, wurde zunächst die Klassenziehung vorbereitet und durchgeführt. Dazu wurde die Klassenliste in das vom internationalen PISA-Konsortium bereitgestellte Stichprobenziehungsprogramm *KeyQuest* eingelesen und anschließend die Klassenstichprobe gezogen.

Die gezogenen Klassen wurden dann bei der Schülerstichprobenziehung aller allgemeinbildenden Schulen und aller Förderschulen berücksichtigt.[11] Hierzu wurde die gezogene Klasse bzw. wurden die gezogenen Klassen je Schule in den bereits vorliegenden Schülerlisten vermerkt, die Schülerlisten ebenfalls in *KeyQuest* eingelesen und dann die Schülerstichprobe für jede einzelne PISA-Schule gezogen. Insgesamt wurden

9 Wie eingangs schon erwähnt werden diese Schülerinnen und Schüler dann sowohl für die Analysen der Gruppe der Fünfzehnjährigen als auch für die Analysen der Gruppe der Neuntklässler herangezogen.

10 Die Kommunikation zwischen den Schulen und dem *IEA DPC* über die für PISA gezogenen Schülerinnen und Schüler sowie Lehrerinnen und Lehrer verlief später nur über diese Ordnungsnummern, welche in beiden Schülerlisten und beiden Lehrerlisten enthalten waren.

11 Für die Berufsschulen war die Klassenziehung nicht relevant, da an diesen Schulen nur die Testgruppe der fünfzehnjährigen Schülerinnen und Schüler gezogen wurde.

© Waxmann Verlag GmbH

so n = 7 189 Fünfzehnjährige für die Testgruppe der fünfzehnjährigen Schülerinnen und Schüler ermittelt sowie weitere n = 3 395 Schülerinnen und Schüler für die Testgruppe der 9. Klasse gezogen. Auf die letztgenannten wird – wie oben bereits erwähnt – im vorliegenden PISA-Bericht nicht eingegangen.

In ähnlicher Weise wurde die Lehrerstichprobe gezogen: Auch die Lehrerlisten wurden zunächst in das Stichprobenziehungsprogramm *KeyQuest* importiert, anschließend wurde mit diesem die Stichprobenziehung durchgeführt. Um sicherzustellen, dass alle Lehrkräfte befragt wurden, die in der gezogenen 9. Klasse einer Schule unterrichteten und um insbesondere Analysen zum Klassenkontext durchführen zu können, wurden diese Lehrkräfte zusätzlich in die Lehrerstichprobe aufgenommen, auch wenn sie bei der vorherigen Stichprobenziehung zufällig *nicht* gezogen wurden. Damit ergibt sich eine Gesamtzahl der naturwissenschaftlichen Lehrkräfte von n = 2 949 und der sonstigen Lehrkräfte von n = 5 498. Von den naturwissenschaftlichen Lehrkräften wurden n = 2 827 zufällig gezogen und n = 122 zusätzlich (als Lehrkräfte einer PISA-Klasse) aufgenommen. Bei den sonstigen Lehrkräften wurden n = 4 588 zufällig gezogen und n = 910 zusätzlich aufgenommen. Die Gesamtstichprobe der Lehrkräfte liegt damit bei n = 8 447.

Die Ergebnisse der Klassen-, Schüler- und Lehrerstichprobenziehung wurden an die Schulen zurückgemeldet, damit dort die Durchführung der PISA-Studie organisiert und vorbereitet werden konnte.

12.2.2 Realisierte Schul-, Schüler- und Lehrerstichproben

Die bisher beschriebenen Schul-, Schüler- und Lehrerstichproben sind sogenannte *Bruttostichproben*. Hiervon zu unterscheiden sind die *Nettostichproben*. Während es sich bei den Bruttostichproben ganz allgemein um diejenigen Schulen, Schülerinnen und Schüler sowie Lehrerinnen und Lehrer handelt, die zufällig für die Teilnahme an PISA ausgewählt wurden, umfassen die Nettostichproben diejenigen Schulen und Personen, die später tatsächlich an der Studie teilgenommen haben und für die zugleich auswertbare Daten vorlagen. Die Brutto- und Nettostichprobe in PISA 2015 unterscheiden sich wie folgt: Nicht alle der zunächst gezogenen 267 Schulen (vgl. Tabelle 12.1) nahmen tatsächlich an der PISA-Studie teil. Dies ist auf verschiedene Ursachen zurückzuführen: So gab es eine Schule, die zum Zeitpunkt der Stichprobenziehung nicht mehr existierte.[12] Darüber hinaus kam es zu zehn weiteren Schulausfällen, da es vor Ort keine fünfzehnjährigen Schülerinnen und Schüler gab. Bei diesen Schulen handelt es sich um fünf Berufsschulen, drei Förderschulen sowie zwei kleine Schulen mit nur sehr geringer Schüleranzahl. Bei diesen Fällen wurden gemäß der internationalen PISA-Stan-

12 Fälle wie dieser sind darauf zurückzuführen, dass die amtlichen Daten zur Schulstatistik zu einem bestimmten Zeitpunkt erhoben werden, die Stichprobenziehung jedoch erst später erfolgt und die deutsche Schullandschaft teilweise recht dynamisch ist.

© Waxmann Verlag GmbH

dards keine *Ersatzschulen* gezogen.[13] Im Gegensatz dazu wurden acht Originalschulen durch ihre jeweils erste Ersatzschule ersetzt. Zudem wiesen drei Schulen, die zwar an PISA teilgenommen hatten, eine zu geringe Beteiligung auf Schülerebene auf (weniger als 50 Prozent der gezogenen Schülerinnen und Schüler) und wurden aus der Datenanalyse ausgeschlossen. Zieht man von allen n = 267 Schulen die Ausfälle ab, so nahmen schließlich n = 253 Schulen an der PISA-Studie 2015 teil. Damit liegt die gewichtete Teilnahmequote auf Schulebene bei 98.9 Prozent. Diese Zahl bezieht sich auf die ursprünglich gezogenen Schulen und zeigt auf, wie viele dieser Schulen nach eventuellen Ersetzungen an der Erhebung teilgenommen haben.

Des Weiteren gab es auch auf der Schülerebene Ausfälle. Nicht alle für die PISA-Studie vorgesehenen Schülerinnen und Schüler nahmen am Test bzw. an der Befragung teil. Gründe hierfür können beispielsweise sein, dass eine Schülerin oder ein Schüler erkrankt war, aufgrund sonderpädagogischen Förderbedarfs offiziell von der Studie ausgeschlossen wurde oder die Schule bereits verlassen hatte.[14] Von den n = 7 189 gezogenen Schülerinnen und Schülern nahmen insgesamt n = 6 504 teil. Damit liegt die gewichtete Teilnahmequote auf Schülerebene bei 93.2 Prozent.

Da für die Lehrkräfte eine Teilnahme an der PISA-Studie nicht verpflichtend war, liegt ihre Teilnahmequote deutlich unterhalb der Schülerteilnahmequote: Von den n = 8 447 Lehrkräften nahmen n = 5 600 an der Online-Lehrerbefragung teil, womit die ungewichtete[15] Teilnahmequote bei 66.3 Prozent liegt. Differenziert man hierbei noch einmal zwischen den naturwissenschaftlichen Lehrkräften (Teilnahme: n = 2 032 von n = 2 949) sowie den sonstigen Lehrkräften (Teilnahme: n = 3 568 von n = 5 498), dann zeigen sich keine wesentlichen Unterschiede in den Teilnahmequoten zwischen den Gruppen. Die ungewichtete Quote beträgt 68.9 Prozent für die Erstgenannten und 64.9 Prozent für die Letztgenannten.

12.2.3 Gewichtung als Adjustierung unterschiedlicher Ziehungswahrscheinlichkeiten

Aufgrund des mehrstufigen Stichprobenziehungsverfahrens haben nicht alle Schulen sowie alle Schülerinnen und Schüler innerhalb dieser Schulen dieselbe Wahrscheinlichkeit, Teil der PISA-Stichprobe zu werden. Außerdem wird nicht jede und jeder in die Stichprobe gezogene Jugendliche an der Erhebung teilnehmen (z. B. wegen Krankheit). Das Vorliegen gleicher Ziehungswahrscheinlichkeiten für jede Untersuchungsein-

13 Gemäß der internationalen PISA-Standards sind für jede gezogene Originalschule zwei Ersatzschulen vorgesehen, die jedoch nur unter bestimmten Bedingungen zum Einsatz kommen können.

14 Letzteres kann vorkommen, da zwischen dem Zeitpunkt der Schülerlistung an den Schulen und der Durchführung der Stichprobenziehung mehrere Wochen vergehen können, in denen sich – wenn auch selten – die Zusammensetzung der Schülerinnen und Schüler in den Schulen verändern kann (z. B. aufgrund von Zuzug oder Wegzug einzelner Schüler).

15 Eine gewichtete Teilnahmequote macht in diesem Zusammenhang keinen Sinn, da nur Gewichte auf Schul- bzw. Schülerebene vorhanden waren.

© Waxmann Verlag GmbH

heit (Schülerinnen und Schüler) und die Teilnahme bei Auswahl sind aber eigentlich notwendige Voraussetzungen *für die Verallgemeinerbarkeit von Stichprobenbefunden für die Population (Bortz* & Döring, 2006). Ungleichen Ziehungswahrscheinlichkeiten und unterschiedliche Teilnahmeraten werden daher durch sogenannte Schulbasis- bzw. Schülerbasisgewichte ausgeglichen. Beide Gewichte resultieren in einem gesamten Schülergewicht.

In einem ersten Schritt werden die Schulbasisgewichte errechnet, welche sich umgekehrt proportional zu den Ziehungswahrscheinlichkeiten verhalten. So haben z. B. Gymnasien innerhalb eines Bundeslandes mit vielen Gymnasien (vgl. explizite und implizite Stratifizierung) eine geringere Wahrscheinlichkeit, in die Stichprobe zu gelangen, und somit dann einen höheren Gewichtungsanteil als andere Schularten in diesem Bundesland. Ferner geht die Nichtbeteiligung von einzelnen Schulen aufgrund von zu geringer Teilnahmebereitschaft in diesem Schulbasisgewicht als Korrekturfaktor mit ein.

In einem zweiten Schritt werden die Schülerbasisgewichte innerhalb der gezogenen Schulen ermittelt, ebenfalls umgekehrt proportional zu den Schülerziehungswahrscheinlichkeiten. Hierbei spielt die Größe der Schule eine zentrale Rolle. So haben Schülerinnen und Schüler einer größeren Schule eine geringere Wahrscheinlichkeit in die Stichprobe zu gelangen als Jugendliche innerhalb einer kleineren Schule und somit einen größeren Gewichtungsfaktor. Des Weiteren wird das Schülerbasisgewicht auch auf Ebene der Schülerinnen und Schüler um deren Nichtteilnahme aufgrund von Krankheit oder Ausschluss korrigiert.

Neben den beiden oben erwähnten Korrekturfaktoren gibt es drei weitere Korrekturwerte. Einer dieser Werte bezieht sich auf teilnehmende Staaten, in denen nur fünfzehnjährige Schülerinnen und Schüler der Klassenstufe befragt werden, in welcher die meisten Fünfzehnjährigen zu erwarten sind. Zwei weitere Korrekturwerte gleichen etwaige Differenzen der Schulbasis- bzw. Schülerbasisgewichte zwischen der Stichprobenziehung und der tatsächlichen Größe dieser Ziehung zum Zeitpunkt der Erhebung aus. Genauere Angaben zu diesen Kennwerten können dem *Technical Report* zu PISA 2012 (OECD, 2014) sowie dem – zum Zeitpunkt der Drucklegung dieses Bandes noch nicht vorliegenden – *Technical Report* der OECD zu PISA 2015 entnommen werden.

Aus dem Schulbasis- und Schülerbasisgewicht sowie den drei weiteren Korrekturfaktoren wird durch Multiplikation das Schülergesamtgewicht berechnet. Um für die Population repräsentative Aussagen zu treffen, muss das Schülergesamtgewicht für alle gängigen Analysen der PISA-Daten verwendet werden. Mit den gewichteten Daten können nun die Ergebnisse für die Stichprobe auf die gesamte Population aller fünfzehnjährigen Schülerinnen und Schüler in Deutschland bzw. in allen an PISA teilnehmenden Staaten verallgemeinert werden.

© Waxmann Verlag GmbH

12.3 Tests zur Erfassung der Kompetenzen

Das Ziel international vergleichender Schulleistungsstudien wie PISA besteht darin, Erkenntnisse zum mittleren Kompetenzniveau bestimmter Grundgesamtheiten (Populationen) zu gewinnen. Im Gegensatz zur individuellen Diagnostik, wie sie beispielsweise im schulischen Kontext mit dem Ziel einer objektiven Aussage über einzelne Schülerinnen und Schüler durchgeführt wird (z. B. Klassenarbeiten), ist der individuelle Beitrag einzelner Personen dabei von untergeordneter Bedeutung. Aufgrund dieser spezifischen Zielsetzung weichen die Methoden der Datengewinnung und Auswertung, die im Rahmen von international vergleichenden Schulleistungsstudien eingesetzt werden, substanziell von denjenigen mit individualdiagnostischen Zielsetzungen ab. Hauptaspekte sind dabei die Aufbereitung und Darbietungsform des Testmaterials sowie die psychometrischen Modelle zur Datenauswertung. Beide Aspekte werden in diesem Abschnitt eingehender dargestellt.

12.3.1 Das PISA-Testmaterial im Multi-Matrix-Design

Psychometrische Tests und Fragebögen werden dazu eingesetzt, bestimmte Eigenschaften wie Einstellungen (vgl. Kapitel 13) oder Kompetenzen aus unterschiedlichen Inhaltsbereichen zu erfassen. In beiden Fällen werden die Reaktionen auf die einzelnen Teilaufgaben (*Items*) dabei als manifeste (d. h. direkt beobachtbare) Indikatoren für eine latente (d. h. verborgene) Variable angesehen. Im Rahmen von international vergleichenden Schulleistungsstudien wie PISA werden Leistungstests eingesetzt, um durch die Analyse von repräsentativen Stichproben Rückschlüsse auf das Niveau der Kompetenzen in der Population ziehen zu können (vgl. Abschnitt 12.2). Das im Rahmen von PISA eingesetzte Testmaterial muss daher im Hinblick auf die angestrebten Aussagen über die Population der fünfzehnjährigen Schülerinnen und Schüler im Wesentlichen zwei zentralen Punkten gerecht werden. Dies sind (1) eine hinreichende Genauigkeit der Messung und (2) eine möglichst kurze Testzeit für jede einzelne Schülerin und jeden einzelnen Schüler. Inhaltlichen Überlegungen folgend sollen gleichzeitig mehrere Kompetenzbereiche erfasst werden. Um dabei eine hinreichende Genauigkeit (Reliabilität – vgl. Abschnitt 12.5) bei der Messung zu erzielen, werden möglichst viele Testaufgaben entwickelt und eingesetzt. Bei PISA besteht eine Testaufgabe in der Regel aus einem Aufgabenstamm (Stimulus), also einem zu lesenden Textabschnitt (z. B. für Lesekompetenz), oder einer Darstellung eines naturwissenschaftlichen Phänomens, in der Bild (Grafik) und Text kombiniert werden (z. B. für naturwissenschaftliche Kompetenz). Auf einen solchen Stimulus beziehen sich mehrere einzelne Fragen, welche in Teilaufgaben zusammengefasst sein können. Die Summe aller Teilaufgaben zu einem Stimulus wird im PISA-Kontext auch als *Unit* bezeichnet. Beim sogenannten *Scoring* (etwa: Bewertung, Auswertung) werden die Antworten der Schülerinnen und Schüler zu den einzelnen Teilaufgaben in ein zwei- oder mehrstufiges (quantitatives) Kategoriensystem überführt

© Waxmann Verlag GmbH

(z. B. 0 = falsch, 1 = teilweise richtig, 2 = richtig). Im Rahmen der psychometrischen Auswertung (vgl. Abschnitt 12.3.2) wird eine so ausgewertete Teilaufgabe als kleinste Analyseeinheit eines Tests auch in der deutschsprachigen Literatur als *Item* bezeichnet (vgl. Osterlind, 1990). Die über die Anzahl der Items definierte Testlänge (vgl. Lienert & Raatz, 1994, S. 33) ist dabei eine maßgebliche Größe zur Beeinflussung der Reliabilität (z. B. Moosbrugger & Kelava, 2012, S. 34). Aus der Perspektive der Psychometrie ist ein möglichst umfangreiches Testmaterial erwünscht. Demgegenüber ist die zur Verfügung stehende Testzeit, also die Summe der Bearbeitungsdauer für alle vorgelegten Items, begrenzt. Aus diesen Gründen wird in PISA und auch in anderen international vergleichenden Schulleistungsstudien für die Datenerhebung ein sogenanntes *Multi Matrix Sampling* (MMS) eingesetzt (Lord, 1965). Der Kerngedanke besteht darin, bezogen auf die gesamte Stichprobe zwar eine große Anzahl von Items einzusetzen, dabei aber nicht alle Items sämtlichen Personen der Stichprobe vorzulegen. Das Prinzip der Ziehung einer Stichprobe von Personen (vgl. Abschnitt 12.2) wird so auch auf die einzelnen Units bzw. deren Items übertragen und als Item-Sampling bezeichnet (vgl. Johnson & Lord, 1958; Lord, 1962; Pugh, 1971). Jeder einzelnen teilnehmenden Schülerin bzw. jedem Schüler wird so nur eine Teilmenge der Items vorgelegt. Dabei muss darauf geachtet werden, dass die einzelnen PISA-Domänen gleichmäßig in den verschiedenen Testformen vertreten sind. Für die PISA-Runde 2015 wurden einerseits die Items vergangener Runden übernommen – und dabei lediglich an die computerbasierte Testplattform angepasst (vgl. Abschnitt 12.4) – und andererseits neue Items entwickelt, die keine Entsprechung als Papier-und-Bleistift-Version haben. In PISA 2015 wurde, wie in früheren Erhebungsrunden auch, neben den Domänen Naturwissenschaften, Mathematik und Lesen eine fachübergreifende Domäne erfasst, das sogenannte Problemlösen im Team.[16] Zur Darbietung und Bearbeitung der Items für die in 2015 insgesamt vier PISA-Domänen wurden diese in übergeordneten Gruppen – sogenannten *Clustern* – zusammengestellt (vgl. Anhang B; OECD, 2016b). Die Summe der Bearbeitungszeit aller Items eines Clusters beträgt für alle Kompetenzbereiche 30 Minuten. Für die drei PISA-Kompetenzbereiche ergeben sich so zunächst jeweils sechs Cluster, welche ausschließlich geheim gehaltene Aufgaben enthalten, die bereits in vergangenen PISA-Runden eingesetzt wurden. Für den Bereich Naturwissenschaften wurden zusätzlich sechs Cluster mit neu entwickelten Aufgaben erstmalig in PISA 2015 eingesetzt. In der fachübergreifenden Testkomponente *Problemlösen im Team* wurden die einzelnen Aufgaben in drei Clustern zusammengestellt (vgl. Tabelle 12.2).

Die einzelnen Cluster wurden für die Erhebung in unterschiedlicher Reihenfolge zusammengestellt und den Schülerinnen und Schülern über ein Computersystem (vgl. Abschnitt 12.4) dargeboten. Diese unterschiedlichen Zusammenstellungen von Aufgaben werden als *Booklets*, *Itemsets* oder auch *Testformen* bezeichnet. Die sich aus den einzelnen Testformen ergebende Struktur des gesamten Instrumentariums zur Daten-

16 Die Ergebnisse für diese Testkomponente sind nicht Teil der Erstberichterstattung im vorliegenden Buch, sondern werden seitens der OECD für 2017 erwartet.

© Waxmann Verlag GmbH

Tabelle 12.2: Überblick über die Cluster und Kompetenzbereiche mit Trend-Aufgaben und neu entwickelten Aufgaben

Kompetenzbereich	neu entwickelte Items	Trend-Items
Lesen	*keine*	Sechs 30-Min.-Cluster (R1 – R6)
Mathematik	*keine*	Sechs 30-Min.-Cluster (M1 – M6)
Naturwissenschaften	Sechs 30-Min.-Cluster (S7 – S16)	Sechs 30-Min.-Clusters (S1 – S6)
Problemlösen im Team	Drei 30-Min.-Cluster (C1 – C3)	*keine*

Anmerkungen: Standardaufgaben (a); leichte Aufgaben (b).

erhebung, wird Booklet-Design genannt. Die einzelnen Cluster für die vier Domänen werden bei PISA 2015 zunächst in 66 verschiedenen Basis-Testformen organisiert, welche in unterschiedlichen Anteilen an die teilnehmenden Schülerinnen und Schüler ausgegeben werden (vgl. Abbildung 12.4).

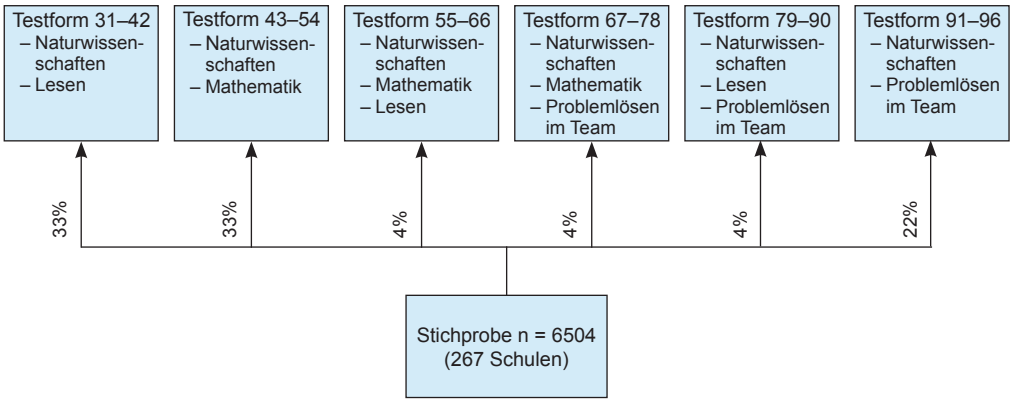


Abbildung 12.4: Zuordnung der PISA-Domänen zu Basis-Testformen und deren prozentuale Verteilung in der Stichprobe

Wie die beiden Kästen links und der Kasten rechts in Abbildung 12.4 zeigen, existieren 30 ($2 \cdot 12 + 6$) verschiedene Basis-Testformen, die jeweils zwei der vier Domänen kombinieren.

Diese Basis-Testformen werden an insgesamt 88 Prozent der Schülerinnen und Schüler bei der Erhebung ausgegeben. Eine detaillierte Erklärung und Darstellung des bei PISA 2015 eingesetzten Booklet-Designs auf der Ebene der einzelnen Cluster sowie der darin enthaltenen Aufgaben ist im Anhang B zu diesem Kapitel wiedergegeben.

© Waxmann Verlag GmbH

Wie die Darstellungen in der Tabelle 12.2 und Abbildung 12.4 zeigen, bedeutet die Anwendung eines Booklet-Designs, dass verschiedene Gruppen der Schülerinnen und Schüler unterschiedliche Aufgaben bearbeiten. Bei der Konstruktion dieser Designs sowie der anschließenden Zuordnung der Schülerinnen und Schüler zu den einzelnen Testformen muss daher darauf geachtet werden, dass sich kein systematischer Zusammenhang zwischen der Zuordnung einer Testform und den darin enthaltenen Fragen zu Schülern mit bestimmten Merkmalen (wie etwa dem Geschlecht oder der besuchten Schulart) ergibt. In diesem Sinne muss auch sichergestellt werden, dass bei der Zuordnung der Items zu den Clustern und der Cluster zu den Testformen keine systematischen und einseitigen Verzerrungen entstehen. So ist beispielsweise zu vermeiden, dass in einer Testform ausschließlich schwierige Mathematikaufgaben enthalten sind, welche dann möglicherweise systematisch an Schülerinnen oder Schüler mit niedriger Mathematikkompetenz ausgegeben werden. Das Prinzip der unsystematischen Zuweisung von Untersuchungseinheiten (hier: Schülerinnen und Schüler) zu experimentellen Bedingungen (hier: Items bzw. Testformen) ist aus dem Bereich der experimentellen Versuchsplanung bekannt. Solche Versuchspläne werden in wissenschaftlichen Untersuchungen immer dann eingesetzt, wenn unerwünschte Zusammenhänge zwischen dem zu messenden Merkmal und bestimmten experimentellen Bedingungen vermieden werden sollen (Fisher & Yates, 1963; Gelman, 2005). Durch die experimentelle *Versuchsplanung* werden die Untersuchungseinheiten hinsichtlich relevanter Merkmale balanciert den experimentellen Bedingungen zugewiesen. Weil dabei in der Regel nicht alle Merkmalskombinationen und experimentellen Bedingungen berücksichtigt werden können, spricht man in diesem Zusammenhang von *Balanced Incomplete Block Designs* [dt. balancierte unvollständige Block-Designs] (Frey & Bernhardt, 2012; Frey, Hartig & Rupp, 2009). Einen vertiefenden Überblick über solche „ausbalancierten" Testdesigns und deren Anwendung im Bereich international vergleichender Studien geben Rutkowski, Gonzalez, von Davier & Zhou (2013).

Das Booklet-Design für PISA 2015 ist so konzipiert, dass die Gesamttestzeit für die vier Domänen Lesen, Mathematik und Naturwissenschaften und Problemlösen im Team für jede Schülerin bzw. jeden Schüler etwa zwei Stunden beträgt. Im Vergleich zu früheren PISA-Runden weist das Booklet-Design für PISA 2015 erheblich mehr unterschiedliche Testformen auf, was wiederum einer besseren Balancierung der Cluster und der darin enthaltenen Items zur Folge hat. Die besondere Struktur des bei PISA 2015 eingesetzten Testmaterials hat zur Folge, dass die einfache summative Verrechnung der gelösten Aufgaben (z. B. Anzahl der erreichten Punkte in einer Klassenarbeit) nicht zu präzisen Schätzungen der Populationsparameter (mittlere Kompetenz in einzelnen Teilnehmerstaaten) herangezogen werden kann.

© Waxmann Verlag GmbH

12.3.2 Psychometrische Modelle zur Testauswertung

Den Ausgangspunkt der bei PISA 2015 eingesetzten *psychometrischen Modelle* zur Auswertung der Items (Teilaufgaben) bildet das von dem dänischen Mathematiker Georg Rasch formulierte logistische Testmodell, welches daher als *Rasch-Modell* (RM; Rasch, 1960) bezeichnet wird. Für Items mit einem mehrstufigen Kategoriensystem zur Bewertung (z. B: *falsch*, *teilweise richtig*, *richtig*) wurde das RM von Masters (1982) zum sogenannten *Partial-Credit-Modell* (PCM)[17] erweitert (vgl. auch Heine, Sälzer, Borchert, Siberns & Mang, 2013). Den Kerngedanken dieser beiden Modelle bildet die Überlegung, dass sich die Lösungswahrscheinlichkeit eines Items als Ergebnis der Interaktion von zwei Größen ergibt. Dies ist einerseits die Merkmalsausprägung bzw. *Kompetenz* (θ) der Schülerinnen und Schüler und andererseits die *Schwierigkeit* (σ) des jeweiligen Items. In der methodisch orientierten Literatur werden diese beiden Größen als *Parameter* eines psychometrischen Antwortmodells bezeichnet (vgl. Heine et al., 2013). Solche Modelle bieten eine Erklärung für das Zustandekommen von empirischen Testdaten an, weshalb sie unter dem Oberbegriff der *Item Response Theory* subsumiert werden (IRT; DeMars, 2010; Hambleton & Swaminathan, 1985). Im Laufe der psychometrischen Forschung wurde neben dem RM und dem PCM eine ganze Reihe weiterer (komplexerer) Modelle zur Erklärung von empirischen Antwortdaten entwickelt und vorgestellt (für einen Überblick vgl. z. B. Fischer & Mollenaar, 1995; van der Linden & Hambleton, 2010). Der Unterschied zwischen den verschiedenen (parametrischen) Modellen ist im Wesentlichen die Anzahl der *Modellparameter*, welche zur Beschreibung der empirischen Daten herangezogen werden. Während im RM und PCM, die in den vergangen PISA-Erhebungsrunden die Grundlage der Datenauswertung bildeten, lediglich *ein* Parameter zur Beschreibung der Aufgabencharakteristik verwendet wird (die Aufgabenschwierigkeit σ), werden die Aufgaben in PISA 2015 zusätzlich mit einem zweiten Modellparameter beschrieben. Die Ermittlung dieser psychometrischen Aufgabencharakteristiken (Itemparameter) wird mit dem Begriff *Itemkalibrierung* bezeichnet. Die Ergebnisse einer solchen Itemkalibrierung lassen sich graphisch anhand sogenannter *Item Characteristic Curves* (ICCs) darstellen (vgl. Abbildung 12.5). Abgebildet wird dabei der Verlauf der auf der y-Achse abgetragenen Lösungswahrscheinlichkeit p in Abhängigkeit der Differenz der beiden Parameter[18] σ und θ. Der Itemparameter σ bestimmt in solchen Darstellungen die horizontale Position der ICC an deren Wendepunkt. Weitere Eigenschaften dieser Kurve für die Antwortwahrscheinlichkeit p können durch die Aufnahme zusätzlicher Parameter in die Funktion von p modelliert werden. Zur Itemkalibrierung wird in PISA 2015 neben den Itemschwierigkeiten neu auch die *Itemtrennschärfe* α anhand der empirischen Daten ermittelt, welche die Steigung der Kurven für

17 Der englische Begriff *„partial credit"* leitet sich bei der Modellbezeichnung aus der Tatsache ab, dass für teilrichtige Antworten eben auch Teil- (*partial*) Punkte (*credit*) vergeben werden, was letztlich eine weniger strenge Bewertung der gegebenen Antwort impliziert.

18 Die Erklärungen zur Modellierung der Antwortwahrscheinlichkeit p über nur zwei Parameter beziehen sich auf das Rasch-Modell für Items mit dichotomem (zweistufigem) Antwortformat.

© Waxmann Verlag GmbH

die ICCs der Items bestimmt. Das Konzept der *Itemtrennschärfe* ist auch aus der Klassischen Test-Theorie (KTT) bekannt und besagt letztlich, dass einzelne Items unterschiedlich gut geeignet sein können, zwischen hoch und niedrig kompetenten Schülerinnen und Schülern zu unterscheiden. Items mit sehr geringer Trennschärfe werden in diesem Sinne sowohl von hoch kompetenten als auch von niedrig kompetenten Schülerinnen und Schülern mit annähernd gleicher Wahrscheinlichkeit gelöst. Es ist offensichtlich, dass eine solche Itemcharakteristik (sehr niedrige Trennschärfe) im Sinne einer differenzierenden Messung nicht anzustreben ist. In den vor PISA 2015 eingesetzten psychometrischen Modellen war der Trennschärfe-Parameter über alle Items implizit auf einen Wert von $\alpha = 1$ fixiert, was einer Steigung der ICC an deren Wendepunkt von 1 entspricht. Das in PISA 2015 erstmalig eingesetzte 2-Parameter-Logistische-Modell (2-PL-Modell) erlaubt nun unterschiedliche Itemtrennschärfen. Im 2-PL-Modell kann dieser zusätzliche Parameter für jedes Item frei geschätzt werden. Er kann damit – je nach empirischer Datenlage – vom Wert $\alpha = 1$ mehr oder weniger abweichen. In Erweiterung der Modellgleichung des Rasch-Modells für dichotome Antwortformate (vgl. Heine et al., 2013, S. 327) um den Trennschärfe-Parameter α ergibt sich die Modellgleichung des 2-parametrigen Modells formal wie in Gleichung 12.1 dargestellt.

$$p_i(X_{vi} = x|\theta_v, \sigma_i, \alpha_i) = \frac{exp(x \cdot (\alpha_i \cdot (\theta_v - \sigma_i)))}{1 + exp(\alpha_i \cdot (\theta_v - \sigma_i))}; x \in \{0, 1\} \tag{12.1}$$

Dieses Modell wurde als Spezialfall des allgemeinen logistischen Testmodells von Birnbaum (1968) vorgestellt. Modelliert wird hier die Antwortwahrscheinlichkeit p für die Wahl einer der beiden Kategorien ($x = 0$ oder $x = 1$) der *Zufallsvariable* X (die Antworten der Schülerinnen und Schüler), gegeben die Modellparameter θ, σ und α. Die Funktionskurven der einzelnen Items – die ICCs – werden dabei durch ihre horizontale Lage (σ) auf dem Kontinuum der latenten Merkmalsdimension und zusätzlich durch deren Steigung (α) definiert.

Analog zur Erweiterung des RM zum PCM für mehrstufige Antwortformate lässt sich auch das 2-PL-Modell für mehrstufige Antwortformate generalisieren. Diese Erweiterung geht auf Muraki (1992) zurück und wird als *Generalized Partial Credit Model* (GPCM) bezeichnet. Analog zu der Beziehung zwischen PCM und RM reduziert sich das GPCM zum 2-PL-Modell, wenn Items mit lediglich zwei Antwortkategorien vorliegen (sogenannte dichotome Items). Dies erlaubt die gleichzeitige Skalierung von zwei- und mehrstufigen Items wie sie bei PISA eingesetzt werden. Eine ausführliche Darstellung des GPCM wird von Muraki (1992) gegeben; das GPCM wird hier als Verallgemeinerung des PCM (Masters, 1982) und der *Ratingskalen*-Parametrisierung des PCM von Andrich (1978) hergeleitet. Für eine allgemeine Übersicht zu den einzelnen Modellen und deren unterschiedlicher *Parametrisierung* sei auf das Handbuch von van der Linden und Hambleton (2010) verwiesen. Der zusätzliche Itemparameter im zweiparametrigen Testmodell erweitert die Möglichkeiten zur Beschreibung der Charakteristik

© Waxmann Verlag GmbH

und für die Auswertung einzelner PISA-Teilaufgaben (Items). Abbildung 12.5 illustriert exemplarisch die Auswirkungen unterschiedlicher Werte für den Trennschärfeparameter α auf die ICC anhand der vergleichenden Darstellung von drei Items mit gleicher Schwierigkeit, welche allerdings unterschiedliche Trennschärfen aufweisen.

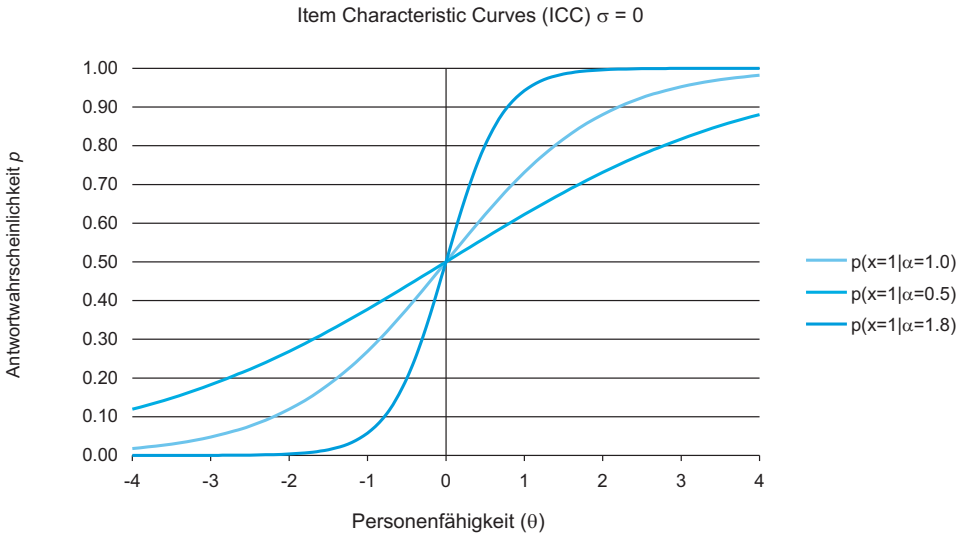


Abbildung 12.5: Item Characteristic Curves (ICC) für drei Items mit zweistufigen Antwortformat (richtig / falsch) mit unterschiedlichen Trennschärfen α

12.3.3 Das *1- und 2-Parameter-Logistische-Modell*: eine vergleichende Diskussion von Merkmalen

Die Erweiterung des RM und PCM um zusätzliche Modellparameter und die damit einhergehenden Konsequenzen werden in der psychometrischen Literatur durchaus kontrovers diskutiert (vgl. z. B. Masters, 1988). So gehen mit den von einigen Autoren als zu restriktiv empfundenen Annahmen des RM und PCM einige vorteilhafte Eigenschaften im Hinblick auf die Messung einher. Eine insbesondere für die *Individualdiagnostik* wesentliche Eigenschaft ist die *spezifische Objektivität* der Messung (vgl. Fischer, 1974). Der Begriff der spezifischen Objektivität definiert sich vereinfacht ausgedrückt dadurch, dass für alle Personen – unabhängig vom Kompetenzniveau – die nach ihrer Schwierigkeit geordnete Rangreihe der Items als gleich „empfunden" wird. In der psychometrischen Literatur wird dieser Umstand auch unter Begriffen wie *Invarianz der Messung* oder Invarianz der Item-Ordnung (IIO, *Invariant Item Ordering;* und invariante Ordnung der Personen) zusammengefasst (vgl. z. B. Sijtsma & Molenaar, 2002). Formal lässt sich diese Eigenschaft bezogen auf das RM in der folgenden Relation ausdrücken:

© Waxmann Verlag GmbH

$$p(X_j = 1|\theta_v) \leq p(X_i = 1|\theta_v) \text{ für alle Personen } v = 1 \ldots n.$$

Dieser Ausdruck besagt, dass die Lösungswahrscheinlichkeit des schwierigeren Items *j* stets geringer ausfallen muss als die des leichteren Items *i* – und zwar unabhängig vom individuellen Kompetenzniveau aller Personen *v* einer Stichprobe vom Umfang *n*.

Es lässt sich nun allerdings leicht zeigen, dass diese an und für sich plausibel und notwendig erscheinende Forderung und Messeigenschaft beim RM beziehungsweise PCM gegeben ist (Wright, 1977), nicht aber bei dessen Modellerweiterungen wie etwa dem 2-PL-Modell oder GPCM. Abbildung 12.6 zeigt beispielhaft die bezüglich der Invarianz der Personen- und Item-Ordnung paradoxe Situation, welche sich durch den „Verlust" gleicher Trennschärfen für alle Items (also durch die Einführung eines α-Parameters) und damit dem Verlust der spezifischen Objektivität ergibt. Betrachtet werden hier beispielhaft zwei unterschiedlich schwierige Items *i* (*hellblaue Kurve – leichteres Item*) und *j* (*dunkelblaue Kurve – schwierigeres Item*), für die unterschiedliche Trennschärfen im Rahmen der Bestimmung der Itemcharakteristiken ermittelt wurden.

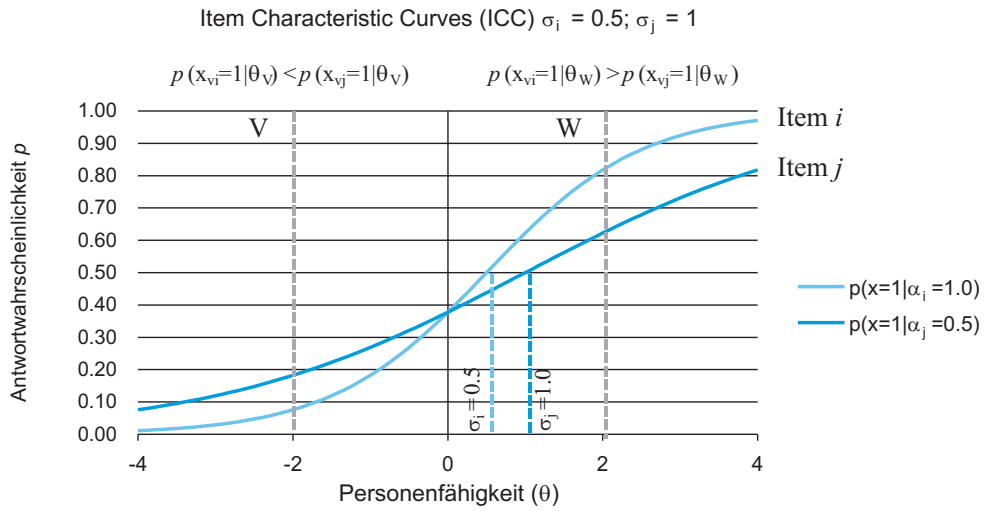


Abbildung 12.6: *Item Characteristic Curves* (ICC) für zwei Items *i* und *j* mit zweistufigen Antwortformat (richtig / falsch) mit unterschiedlichen Trennschärfen α und Schwierigkeiten σ

Des Weiteren werden in diesem Beispiel zwei unterschiedlich kompetente Personen V und W angenommen. Das jeweilige Kompetenzniveau (auf der x-Achse) der beiden Personen ist durch zwei vertikale Linien markiert. Die Position der Person V auf dem latenten Kontinuum der Kompetenz liegt bei der linken grau gestrichelten Linie *(niedrigere Kompetenz)* und die der Person W liegt bei der rechten grau gestrichelten Linie *(höhere Kompetenz)*. Die mit der unterschiedlichen Trennschärfe der beiden Items *i* und *j* einhergehende Überschneidung der ICCs impliziert nun, dass die Lösungswahrscheinlich-

© Waxmann Verlag GmbH

keit des schwierigeren Items *j* für die weniger kompetente Person V höher ist als die Lösungswahrscheinlichkeit derselben Person für das (einfachere) Item *i*. Für die kompetentere Person W dagegen entsprechen die Lösungswahrscheinlichkeiten der beiden Items den Erwartungen, die sich gemäß dem Verhältnis ihrer Schwierigkeiten ergeben. Dieser für den Bereich der Individualdiagnostik relevante Umstand könnte nun zu der Schlussfolgerung verleiten, dass derartige Modelle mit mehreren Parametern in allen Anwendungsbereichen eine objektive und faire vergleichende Messung von Kompetenzen nicht zulassen. Allerdings lässt sich im Hinblick auf die bei PISA intendierten inferenzstatistischen Aussagen auf *Populationsebene* ebenso leicht zeigen, dass der Verlust der spezifischen Objektivität der Messung auf Individualebene nicht notwendigerweise mit etwa „unfairen" Vergleichen der Populationen einzelner Teilnehmerstaaten einhergeht. So zeigen beispielsweise Analysen von Macaskill (2008) empirisch, dass die Interkorrelationen zwischen Populationsmittelwerten nach dem 1- und 2-PL-Modell für 31 untersuchte Teilnehmerstaaten in den PISA Runden 2003 und 2006 durchweg sehr hoch ausfallen. Eine wichtige Voraussetzung ist dabei, dass bezogen auf die internationale Gesamtstichprobe keine systematischen Zusammenhänge zwischen den Itemparametern σ sowie α und den einzelnen teilnehmenden Staaten – sogenannte *Item-by-Country-Interactions* – nachweisbar sind.

Bezogen auf den durch die Skalierung mit dem 2-PL-Modell erzielten Messwert lässt sich zeigen, dass dieser einer gewichteten Verrechnung der einzelnen Items zu einem Gesamtmesswert für das Kompetenzniveau entspricht (Sijtsma & Hemker, 2000). Den Gewichtungsfaktor bildet dabei der Trennschärfeparameter α, welcher jeweils das Ausmaß des Beitrages eines einzelnen Items zum Gesamtmesswert für die Kompetenz quantifiziert (Rost, 2004, S. 135). Prinzipiell ist durch die unterschiedliche Gewichtung der Items und ihrer Inhalte nicht auszuschließen, dass sich der Messinhalt der gesamten Skala (geringfügig) verändert. Gegen eine solche geringfügige Verschiebung der inhaltlichen Schwerpunktsetzung ist zunächst aus der Perspektive der objektiven psychometrischen Erfassung des Kompetenzniveaus nichts einzuwenden. Im Hinblick auf den internationalen Vergleich aller an PISA teilnehmenden Staaten, der ja den Kern der Studie bildet, muss allerdings sichergestellt werden, dass sich derartige Verschiebungen der Messinhalte für alle zu vergleichenden Populationen in den einzelnen Staaten in derselben Weise ergeben. In diesem Sinne konnte Macaskill (2008) anhand von Reanalysen der Daten aus PISA 2006 und 2003 zeigen, dass sich durch die Anwendung des 2-PL-Modells keine grundlegend unterschiedliche Inferenz auf der Ebene der Population bezüglich der *Rangfolge* der teilnehmenden Staaten ergibt. Als vorteilhafter Effekt der genaueren Modellierung der empirischen Daten über einen zusätzlichen Modellparameter ergibt sich demgegenüber allerdings eine Verbesserung der Modellpassung aufgrund des zusätzlichen Trennschärfe-Parameters (Oliveri & von Davier, 2011). Diese verbesserte Modellpassung[19] führt wiederum zu verringerten Standardfehlern

19 Grundsätzlich muss angemerkt werden, dass sich die Modellpassung durch zusätzliche Parameter weiter verbessert. So lässt sich z. B. die Möglichkeit zum Erraten der richtigen Antwort bei Mehrfachwahlaufgaben im sogenannten 3-PL-Modell berücksichtigen (z. B. Aitkin & Aitkin, 2011, S. 42).

© Waxmann Verlag GmbH

für die Populationsstatistiken und verbessert in diesem Sinne die Genauigkeit (Reliabilität) und damit die Güte der Messung auf der Ebene des Staatenvergleichs (Oliveri & von Davier, 2014).

Im Hinblick auf die psychometrischen Modelle, die in PISA zur Testauswertung herangezogen werden, lässt sich zusammenfassen: Spezifische Objektivität auf der Ebene einzelner Personen ist insbesondere dann erstrebenswert, wenn vergleichende *Individualdiagnosen* einzelner Personen oder auch kleinerer Personengruppen angestrebt werden, denn nur so kann ein fairer, spezifisch objektiver Vergleich gesichert werden (vgl. Fischer, 1974). Die Besonderheit der inferenzstatistischen Aussagen, welche mit den PISA-Studien gemacht werden, liegt demgegenüber darin, dass Aussagen auf *Populationsebene* angestrebt werden. Die dabei über das 2-PL-Modell erzielte bessere Modellpassung erhöht entsprechend die Genauigkeit der Messung auf Populationsebene (vgl. Oliveri & von Davier, 2014). Insofern muss argumentiert werden, dass der „Verlust" der spezifischen Objektivität zugunsten einer detaillierteren Modellierung in der Summe vorteilhafte Konsequenzen mit sich bringt. Vorteile bestehen insbesondere in einer verbesserten Modellpassung und geringeren Standardfehlern der Messung. Wichtig ist gleichzeitig, dass bezüglich des Diskriminationsparameters α keine substanziellen *Item-by-Country-Interaktionen* vorliegen.

12.4 Umstellung auf computerbasiertes Assessment (CBA)

Computerbasierte Erhebung (Computer-Based-Assessment; CBA) bedeutet, dass mithilfe des Computers Beobachtungsdaten (z. B. Antworten auf Testaufgaben) gesammelt werden, auf deren Grundlage Schlüsse über Individual- und Gruppenunterschiede in einem bestimmten Merkmal gezogen werden können (vgl. Scalise, 2012). Konkret wird der Computer dafür eingesetzt, die Aufgaben zu präsentieren (z. B. Bild, Text, Audio, Video, Simulation), deren Abfolge zu steuern, die Interaktionen der Testperson mit der Aufgabe zu registrieren (z. B. Maus-Klicks, Tastatureingaben, Touch-Eingaben) und eventuell auch automatisch auszuwerten.

Aus diagnostischer Sicht sprechen mehrere Gründe für einen Wechsel von papierbasierter Erhebung (Paper-Based-Assessment; PBA) zu CBA. So sind am Computer innovative Aufgabenformate realisierbar, die komplex, multimedial und interaktiv sein können (Parshall, Harmes, Davey & Pashley, 2010). Anhand dieser neuen Aufgaben lassen sich auch Merkmale erfassen, die mit einem PBA empirisch nicht zugänglich waren – wie etwa die Kompetenz, eine Simulation am Computer für die Lösung eines Problems zu verwenden. Computerbasierte Assessments können außerdem sowohl Ergebnisdaten als auch Prozessdaten sammeln, wodurch Einblicke in das Lösungsverhalten, also in Auswahl, Sequenz und Dauer von Bearbeitungsschritten möglich sind (z.B. Goldhammer et al., im Druck). Solche in Log-Dateien gespeicherten Interaktionen der Testperson mit dem Computer tragen dazu bei, den Lösungsprozess besser zu verstehen und die Datenqualität zu überprüfen.

© Waxmann Verlag GmbH

Mit der Umstellung von Papiertestheften auf CBA sind organisatorische und logistische Änderungen und Herausforderungen verbunden. Dazu zählen Neuerungen in den internationalen Entwicklungsprozessen, aber auch Übersetzungs-, Anpassungs- und Prüfprozessen in den einzelnen Teilnehmerstaaten einschließlich der dafür eingesetzten Werkzeuge (z. B. spezielle Diagnosesoftware für die Testung der Systemvoraussetzungen vorhandener Schulcomputer). Die Durchführung von CBAs an Schulen stellt spezifische Anforderungen an die Infrastruktur. Diese Anforderungen waren an etwas mehr als einem Viertel der Schulen in der PISA-Stichprobe in Deutschland nicht erfüllt, sodass die Testleiterinnen und Testleiter diese Schulen mit geliehenen Notebooks ausstatteten.

Die mit der CBA-Umstellung einhergehenden psychometrischen Herausforderungen betreffen insbesondere die Frage der Vergleichbarkeit von PBA-Aufgaben (Trend-Items) und ihren computerisierten Varianten (CBA-Aufgaben) und damit das Problem, ob der Wechsel des Testmodus die psychometrischen Eigenschaften der Messung beeinflusst (sogenannter Moduseffekt, vgl. Kröhne & Martens, 2011): Misst die CBA-Variante das gleiche Konstrukt wie die PBA-Aufgabe oder vielleicht eine Mischung von Merkmalen (z. B. weil Computerfertigkeiten eine Rolle spielen könnten)? Erfasst die CBA-Variante das Konstrukt mit der gleichen Genauigkeit? Sind alle oder einzelne Items in der CBA-Variante schwerer oder leichter als die ansonsten identischen PBA-Aufgaben? Fallen solche Schwierigkeitsunterschiede für Testpersonen oder Testgruppen (z. B. je nach Geschlecht) unterschiedlich aus? Variiert bei gleicher oder unterschiedlicher Schwierigkeit die Stärke des Zusammenhangs zwischen zu messendem Konstrukt und richtiger Antwort, das heißt, gibt es einen Effekt des Administrationsmodus auf die Itemdiskrimination (Trennschärfe), die in PISA 2015 erstmals mit modelliert wurde (vgl. Abschnitt 12.3)?

Um systematische Moduseffekte zu vermeiden, wurden die PBA-Aufgaben in PISA 2015 möglichst vergleichbar auf den Computer übertragen. Die in der folgenden Abbildung 12.7 dargestellte Leseverständnisaufgabe „Geizhals“ kann als Beispiel dafür angesehen werden. Beispielsweise passt hier der Lesetext vollständig auf den Bildschirm mit einer Auflösung von 1024x768 Bildpunkten, sodass sich trotz der Formatänderung (Querformat vs. Hochformat) keine Notwendigkeit zum Scrollen oder Blättern ergibt. Insbesondere Scrollen erwies sich in verschiedenen Studien gegenüber Blättern als tendenziell erschwerender Faktor (z. B. Pommerich, 2004). Änderungen ergeben sich dort, wo in der computerbasierten Präsentation spezifische Umsetzungen für einzelne Eigenschaften gebräuchlich sind. Konkret werden am Computer Fortschrittsinformationen in Form eines Balkens und der verbleibenden Testzeit angezeigt, da die Menge der noch zu bearbeitenden Aufgaben nicht aus der Dicke des Testhefts abgeschätzt werden kann. Zudem kann eine Hilfe mit generellen Instruktionen zur Testbearbeitung aufgerufen werden.

Eine ausschließlich auf die Vermeidung von Moduseffekten bedachte Computerisierung von PBA-Aufgaben unter Berücksichtigung bisheriger Forschung kann eine eigene empirische Überprüfung möglicher Effekte nicht ersetzen (vgl. Kröhne & Martens, 2011). Für die Untersuchung der Vergleichbarkeit stehen unterschiedliche

© Waxmann Verlag GmbH

Frage 7: GEIZHALS

Der GEIZHALS und sein Gold. Eine Fabel von Aesop

Ein Geizhals verkaufte alles, was er hatte, und kaufte einen Klumpen Gold, den er in einem Loch in der Erde neben einer alten Mauer vergrub. Jeden Tag ging er, um danach zu sehen. Einer seiner Arbeiter bemerkte die regelmäßigen Besuche des Geizhalses an dem Ort und beschloss, dessen Kommen und Gehen zu beobachten. Der Arbeiter entdeckte bald das Geheimnis des versteckten Schatzes, grub im Boden, fand den Klumpen Gold und stahl ihn. Der Geizhals fand das Loch bei seinem nächsten Besuch leer vor und begann, sich die Haare zu raufen und laut zu klagen. Ein Nachbar, der ihn in seiner Verzweiflung sah und den Grund dafür erfuhr, sagte zu ihm: „Bitte grämt Euch nicht so; nehmt Euch einen Stein, legt ihn ins Loch und stellt Euch vor, das Gold läge noch dort. Er wird Euch den gleichen Dienst erweisen, denn auch, als das Gold noch da war, besaßt Ihr es nicht, da Ihr nicht den geringsten Gebrauch davon gemacht habt."

Verwende die Fabel „Der Geizhals und sein Gold" auf der vorhergehenden Seite, um die folgenden Fragen zu beantworten.

Wie kam der Geizhals zu einem Klumpen Gold?

_____

_____

Abbildung 12.7: Beispielaufgabe „Geizhals" im bisherigen Testheft-Layout (oben) und im neuen CBA-Layout (unten)

© Waxmann Verlag GmbH

Untersuchungsdesigns und statistische Methoden zur Verfügung (für eine Übersicht vgl. Bürger, Kröhne & Goldhammer, 2016). Bei den experimentellen Designs kann grundsätzlich zwischen einem Zwischen- und Inner-Subjekt-Design unterschieden werden. Bei Ersterem werden Schülerinnen und Schüler in eine PBA- und eine CBA-Gruppe unterteilt und bearbeiten Aufgaben ausschließlich im Testheft oder am Computer. Unterschiede zwischen der PBA- und der CBA-Gruppe können bei einer zufälligen Zuweisung zu diesen Gruppen, die sich in keiner Eigenschaft vor der Administration der Aufgaben unterscheiden, eindeutig auf Unterschiede im Modus zurückgeführt werden. Beim Inner-Subjekt-Design bearbeiten die Schülerinnen und Schüler Testteile sowohl am Computer als auch im Testheft (mit Balancierung der Reihenfolge von Testteil und Modus zufällig über Personen). Dieses Design ist mit einem höheren administrativen Aufwand verbunden, bietet aber auch psychometrische Vorteile, insbesondere kann zur Prüfung der Konstruktäquivalenz die Korrelation zwischen dem per PBA und dem per CBA erfassten Merkmal berechnet werden. Eine annähernd perfekte Korrelation würde dafür sprechen, dass in beiden Modi dasselbe Konstrukt gemessen wird.

In PISA 2015 ist die Frage nach potenziellen Moduseffekten besonders relevant, da von ihr abhängt, ob Unterschiede auf Populationsebene über Modi hinweg im Längs- und Querschnitt national und international vergleichend beschreibbar sind. Äußern sich Moduseffekte nur in einer generellen oder itemspezifischen Änderung der Schwierigkeit des Tests, besteht unter Annahme von Konstruktäquivalenz die Möglichkeit, eine gemeinsame Metrik für das PBA und das CBA herzustellen. Bei erfolgreicher Umsetzung eines Zufallsgruppen-Designs und Nutzung dieses Designs zur Schätzung von Itemparametern könnten theoretisch beliebig viele Items in ihrer Schwierigkeit abweichen. Denn durch die – zwischen den Zufallsgruppen identische – Verteilung des Personenmerkmals ist im Rahmen einer gemeinsamen Kalibrierung eine modusübergeifende Metrik, das heißt eine gemeinsame Kompetenzskala, identifizierbar. Es ist aber damit zu rechnen, dass sich zwar einige Items in ihrer Schwierigkeit unterscheiden, andere dagegen nicht (partielle Invarianz, vgl. Bürger et al., 2016). In diesem Fall können die in der Schwierigkeit invarianten Items bei Bedarf als Anker für eine gemeinsame Metrik von CBA und PBA dienen (d. h. ihre Schwierigkeit ist gleichgesetzt), und variante Items erhalten je Modus einen spezifischen Schwierigkeitsparameter (d. h. ihre Schwierigkeit ist nicht gleichgesetzt). Für diese Verankerung müssen ausreichend viele invariante Items vorliegen, die repräsentativ für das zu erfassende Konstrukt sind. Hierbei ist es notwendig, dass die Parameter der Items, die als invariant identifiziert werden, tatsächlich vom Moduswechsel nicht beeinflusst sind und sie nicht nur wegen zu geringer Teststärke als invariant angenommen wurden (Fehler 2. Art). Außerdem stellt sich in Bezug auf die Itemparameter die Frage, ob die realisierten Stichproben beim Erhebungsdesign groß genug waren, um einen relevanten Moduseffekt überhaupt nachweisen zu können. Darüber hinaus ist es notwendig, dass die Items in allen zu vergleichenden Subgruppen (bspw. Geschlecht, Zuwanderungshintergrund, Staaten) als invariant angenommen werden können.

© Waxmann Verlag GmbH

12.5 Reliabilität, Stichprobenfehler und Beurteilerübereinstimmung

Eine der Grundfragen bei jeglicher Messung ist die Frage nach deren *Messgenauigkeit*. Die Messgenauigkeit wird speziell bei der psychometrisch fundierten Erfassung von menschlichen Eigenschaften, Einstellungen und Kompetenzen als *Zuverlässigkeit* oder *Reliabilität* der Messung bezeichnet. Dieses Kriterium stellt neben der *Gültigkeit* (Validität) der Messung ein zentrales Gütekriterium psychometrischer Tests dar (vgl. z. B. Bühner, 2011; Krebs & Menold, 2014). Die Gültigkeit bzw. Validität der Messung muss, zumindest im Hinblick auf die *Inhaltsvalidität* eines Testinstruments („*Was wird hier gemessen?*"), letztlich immer auf der Basis theoriegeleiteter Überlegungen zur Definition der zu erfassenden Eigenschaft und im Hinblick auf die Inhalte der einzelnen Aufgaben (Items) beurteilt werden (Murphy & Davidshofer, 2005). Demgegenüber lässt sich die Reliabilität der Messung anhand verschiedener statistischer Koeffizienten quantifizieren und damit die Größe des eventuell bestehenden Fehlers bei der Messung über unterschiedliche Methoden abschätzen (Balakrishnan & Rao, 2001; Schmidt & Hunter, 1996). Dabei ist es sinnvoll, zunächst die möglichen Fehlerquellen zu identifizieren, welche beim eigentlichen Messvorgang entstehen können (Faulbaum, 2014; Groves, 2005). In der entsprechenden Literatur wird in diesem Zusammenhang in der Regel von unterschiedlichen Varianzquellen (im Hinblick auf die Variabilität der Werte der gemessenen Größe) gesprochen, die zur Abschätzung der Messgenauigkeit bzw. der Reliabilität zueinander ins Verhältnis gesetzt werden (z. B. Miller, 2010).

Bezogen auf die PISA-Studien lassen sich hier vier elementare Varianzquellen benennen. Dies ist zunächst die *Variabilität* der tatsächlichen Kompetenz der Schülerinnen und Schüler in einer der PISA-Kompetenzbereiche, welche das eigentliche Ziel der Messung ist. Als Fehlervarianzquellen, welche im Sinne einer hohen Reliabilität als eigentlich störend anzusehen sind, lassen sich im Hinblick auf die angestrebten Populationsaussagen der *Stichprobenfehler* (Rust & Rao, 1996), der *Designeffekt* (Adams, 2005) und – bei PISA-Aufgaben mit offenem Antwortformat – die *Beurteilerübereinstimmung* (Wirtz & Caspar, 2002) nennen.

12.5.1 Kodierung offener Antworten und Beurteilerübereinstimmung

Ein Teil der bei PISA 2015 eingesetzten Aufgaben wird nicht über ein fest vorgegebenes Kategoriensystem als sogenannte Mehrfach-Wahlaufgabe (*Multiple Choice*) beantwortet, sondern erfordert zur Lösung die Formulierung eines eigenen freien Antwortsatzes bzw. eines kurzen Antworttextes. Diese Antworten der Schülerinnen und Schüler müssen vor der Skalierung in ein quantitatives Kategoriensystem überführt werden (vgl. Abschnitt 12.3). Dazu wird in jedem an PISA teilnehmenden Staat speziell trainiertes Personal zur Kodierung der vorliegenden *offenen Antworten* eingesetzt. Die von

© Waxmann Verlag GmbH

den Schülerinnen und Schülern als Kurztext gegebenen Antworten werden durch die Kodiererinnen und Kodierer in ein geschlossenes Bewertungssystem mit Kategorien wie „*richtig*“, „*teilweise richtig*“ oder „*falsch*“ überführt. Für den Bereich Naturwissenschaften müssen 59 PISA-Aufgaben von Kodiererinnen und Kodierern beurteilt werden. Davon sind 30 Aufgaben neu für PISA 2015 entwickelt worden. Für die Domäne Mathematik mussten 18 Aufgaben und für die Lesekompetenz 43 Aufgaben manuell kodiert werden. Für die deutsche Teilstichprobe in PISA 2015 wurden zur Kodierung der offenen Aufgaben aus den Naturwissenschaften acht, für die Mathematik vier und für die Lesekompetenz sechs Kodiererinnen und Kodierer eingesetzt.

In vorgeschalteten Übungssitzungen wurden diese Kodiererinnen und Kodierer anhand von Kodier-Richtlinien und beispielhaften offenen Antworten bereits während der Feldtest-Phase der PISA-Erhebung geschult. Um die Zuverlässigkeit der Kodierung offener Antworten zu bestimmen, sind zumindest für einige Antworten mehrfache Kodierungen erforderlich. Hierbei werden jeweils zwei Personen eingesetzt, die unabhängig voneinander die Bewertung vornehmen. Dadurch soll einer gegenseitigen Beeinflussung bei der Beurteilung der Schülerantworten vorgebeugt werden (vgl. Bechger, Maris & Hsiao, 2010). Zur Umsetzung dieser Kodierungen wurde in der PISA-Runde 2015 erstmals eine spezielle Software eingesetzt, das sogenannte *Open-Ended Coding System* (OECS). Dieses System wird staatenübergreifend zur Erfassung der Kodierungen der computerbasiert erfassten Antworten verwendet und implementiert das *Kodierdesign* (ETS, 2015b). Gemäß den internationalen Vorgaben werden, je nach landesspezifischem Stichprobenumfang und PISA-Kompetenzdomäne, verschiedene *Kodierdesigns* angewendet. Die Aufgaben aus den Naturwissenschaften wurden für das Kodierdesign in vier Gruppen und für den Kompetenzbereich Lesen in drei Gruppen zusammengefasst. Für den Kompetenzbereich Mathematik wurden keine Aufgabengruppen für das Kodierdesign gebildet. Die international vorgegebenen Kodierdesigns verfolgen zwei Strategien. Es soll die Übereinstimmung der Kodiererinnen und Kodierer bei der Bewertung der Aufgaben innerhalb eines Staates und zwischen den verschiedenen PISA-Teilnehmerstaaten quantifiziert werden. Dabei sollte jede der offenen Fragen mindestens einmal über die gesamte Stichprobe hinweg einfach und für einen Teil der Stichprobe doppelt bewertet werden. Dies bildet dann die Datengrundlage zur Bestimmung der *Beurteilerübereinstimmung*. Tabelle 12.3 zeigt exemplarisch das in Deutschland eingesetzte Kodierdesign für naturwissenschaftliche Kompetenz und die mittlere prozentuale Beurteilerübereinstimmung für Trend- und neu entwickelte Aufgaben nach Aufgabengruppe und bewertender Person.

Die prozentuale Beurteilerübereinstimmung für den Kompetenzbereich Naturwissenschaften liegt für die Teilstichprobe in Deutschland im Mittel über alle für den Zyklus 2015 neu entwickelten 30 Aufgaben (und acht Kodiererinnen und Kodierer) bei 92 Prozent, wobei die höchste Übereinstimmung der Beurteilung 98 Prozent und die niedrigste 74 Prozent (lediglich ein Item – sonst: 82 Prozent) beträgt. Für die 28 Trend-Aufgaben aus den Naturwissenschaften liegt die Beurteilerübereinstimmung im Durchschnitt bei 93 Prozent mit Werten von 99 Prozent für die höchste und 85 Prozent für die nied-

© Waxmann Verlag GmbH

Tabelle 12.3: Kodierdesign und mittlere prozentuale Übereinstimmung nach Kodiererin und Kodierer für den Bereich Naturwissenschaften

Kodiererin / Kodierer:	101*	102	103*	104	105	106	107	108
Itemgruppe 1. (7 neue Items; 7 Trend-Items)	x	x					x	x
Itemgruppe 2. (7 neue Items; 8 Trend-Items)	x	x			x	x		
Itemgruppe 3. (7 neue Items; 7 Trend-Items)			x	x	x	x		
Itemgruppe 4. (9 neue Items; 6 Trend-Items)			x	x			x	x
Prozentuale Übereinstimmung (insgesamt 30 neue Items)	0.94	0.94	0.90	0.91	0.92	0.91	0.93	0.93
Prozentuale Übereinstimmung (insgesamt 28 Trend-Items)	0.94	0.94	0.93	0.93	0.93	0.92	0.95	0.95

Anmerkung. Mit * markierte Kodiererinnen oder Kodierer sind bilingual (Deutsch / Englisch)

rigste Übereinstimmung bei der Beurteilung der offenen Aufgaben. Für den Kompetenzbereich Lesen beträgt die durchschnittliche prozentuale Übereinstimmung über die eingesetzten Kodiererinnen und Kodierer und 43 (Trend-)Items hinweg bei 95 Prozent (Minimum: 93 Prozent; Maximum: 96 Prozent). Für den ebenfalls in PISA 2015 als Nebendomäne erfassten Kompetenzbereich Mathematik fällt die mittlere Beurteilerübereinstimmung über die 18 Items und vier Kodiererinnen und Kodierer homogen aus und beträgt 97 Prozent. Damit fällt die Beurteilerübereinstimmung für den Kompetenzbereich Mathematik im Vergleich zu den anderen Kompetenzbereichen am besten aus.

12.5.2 Designeffekte, Stichprobenfehler und Standardfehler

Wie im Abschnitt 12.2 beschrieben wird bei PISA eine Stichprobe aus der Grundgesamtheit aller fünfzehnjährigen Schülerinnen und Schülern in Deutschland befragt. Zur Präzisierung der Schätzwerte wird eine zweistufige Stichprobenziehung mit Ziehung der Schulen sowie der Schüler innerhalb dieser Schulen vorgenommen. Es ist daher notwendig das Verfahren der Stichprobenziehung auch für die Schätzung der Genauigkeit bei der Schätzung der Populationsmerkmale (z. B. Kompetenzmittelwerte) zu berücksichtigen.

Durch die zweistufige Stichprobenziehung ergibt sich bei allen Analysen eine (statistische) Abhängigkeit der Schülerinnen und Schüler innerhalb der Schulen. Das bedeutet, dass die Merkmale der Jugendlichen (z. B. deren Kompetenzen) innerhalb einer Schule wahrscheinlich ähnlicher ausfallen als zwischen unterschiedlichen Schulen. So haben Jugendliche innerhalb einer Schulart ähnlichere Kompetenzwerte als Schülerinnen und Schüler unterschiedlicher Schularten. Des Weiteren unterscheiden sich diese Abhängigkeiten auch zwischen den einzelnen Teilnehmerstaaten (z. B. aufgrund unterschiedli-

© Waxmann Verlag GmbH

cher Bildungssysteme), was wiederum unterschiedliche Auswirkungen auf die Schätzgenauigkeit haben kann (vgl. OECD, 2012). Bei PISA und anderen Schulleistungsstudien wird dieser Effekt des Stichprobendesigns auf den Fehler der Populationsschätzungen (den sogenannten Standardfehler) als *Designeffekt* bezeichnet (Adams, 2005). Nach einer Definition von Kish (1965) kann der Designeffekt als Verhältnis der Varianz einer Messgröße auf der Basis einer komplexen, mehrstufigen Stichprobenziehung und der Varianz derselben Messgröße auf der Basis einer einfachen Zufallsstichprobe mit dem gleichen Umfang n beschrieben werden. Dieser Designeffekt hat typischerweise zwei Implikationen im Hinblick auf die Datenerhebung und Auswertung. Zum einen wird er verwendet, um die benötigten Stichprobengrößen bei *Large-Scale-Assessments* zu schätzen. Zum anderen wird er zur Verbesserung der Effizienz der Schätzwerte bei komplexen Stichprobenziehungen eingesetzt, wie sie bei PISA angewendet werden (Cochran, 1977; Rust & Rao, 1996).

Für einfache Zufallsstichproben kann zum Beispiel der Standardfehler des Mittelwerts als Stichprobenfehler über eine recht einfache, explizite Formel in Abhängigkeit des Stichprobenumfangs und der Varianz der Messgröße berechnet werden (z. B. Bortz, 2005). Für mehrstufige Stichprobendesigns lassen sich zwar prinzipiell ebenfalls explizite Formeln zur Berechnung der Standardfehler angeben (vgl. Rao, 1988), allerdings setzen diese voraus, dass die Varianzen innerhalb und zwischen den einzelnen Ziehungsstufen (Strata) und, damit der Designeffekt bekannt ist, für alle Populationen (Teilnehmerstaaten) gleich ist. Da die Größe des Designeffekts bei PISA im Vorfeld nicht eindeutig zu quantifizieren ist und darüber hinaus auch zwischen den einzelnen Teilnehmerstaaten unterschiedlich ausfällt, wird der Standardfehler in PISA seit der ersten Erhebung im Jahr 2000 mit sogenannten Replikationsmethoden auf Basis der empirischen Daten berechnet. Bei PISA kommt die *Balanced Repeated Replication* (BRR, z. B. Wolter, 1985) zur Anwendung, mit einer Erweiterung nach Fay (1989; vgl. dazu auch Judkins, 1990). Die statistische Herleitung der Schätzung des Standardfehlers sowie eine detailliertere Darstellung in Bezug auf die PISA-Studie ist dem *Technical Report* zu PISA 2012 (OECD, 2014) zu entnehmen und wird für PISA 2015 im *Technical Report* der OECD erwartet.

Durch den Einsatz von Replikationsmethoden werden die Varianzen zwischen und innerhalb der Schulen berücksichtigt und damit einer möglichen Unterschätzung der Standardfehler vorgebeugt (vgl. Heine et al., 2013; OECD, 2014). Bei der Datenauswertung wird die Berechnung der Varianz durch die Verwendung sogenannter Replikationsgewichte im Datensatz praktisch realisiert, welche bei der (wiederholten) Berechnung jeweils eingesetzt werden.

© Waxmann Verlag GmbH

12.6 Itemkalibrierung, Populationsmodell, Plausible Values und Kompetenzstufen

Eine wesentliche Zielsetzung der PISA-Erhebung besteht in einer international vergleichenden Darstellung des mittleren erreichten Kompetenzniveaus der Population fünfzehnjähriger Schülerinnen und Schüler einzelner Teilnehmerstaaten. Für die dazu eingesetzten Items muss zunächst deren Schwierigkeit bestimmt werden, was als Itemkalibrierung bezeichnet wird. Die Antworten der Schülerinnen und Schüler auf die kalibrierten Items mit bekannter Schwierigkeit dienen als Grundlage für die Bestimmung der erreichten Leistungen einzelner Teilnehmerstaaten.

Die Überführung der empirisch beobachteten Antworten der Schülerinnen und Schüler unter Einbezug der zuvor ermittelten psychometrischen Aufgabencharakteristiken (Itemparameter) in Messwerte des Kompetenzniveaus wird allgemein als *IRT-Skalierung* bezeichnet. Bei diesem Prozess sind bei PISA einige spezifische Besonderheiten hervorzuheben, die im Folgenden dargestellt werden. Der gesamte Prozess der IRT-Skalierung erfolgt in drei Schritten und wurde international für alle beteiligten Staaten durch das von der OECD beauftragte Konsortium durchgeführt. Nachdem zunächst die Itemparameter bestimmt wurden, werden im zweiten Teilschritt die Verteilungsparameter der A-posteriori-Verteilung der latenten Kompetenz(en) ermittelt. Dabei werden die fixierten Itemparameter, die Antworten der Schülerinnen und Schüler zu den Testaufgaben sowie deren Antworten im Schülerfragebogen über latente Regressionsmodelle berücksichtigt. Im letzten Schritt werden schließlich die Kovarianzen zwischen den latenten Kompetenzdomänen bestimmt sowie sogenannte *Plausible Values* als Realisationen der A-posteriori-Verteilung der latenten Kompetenz(en) gezogen.

Die Möglichkeit der Separierbarkeit des ersten Teilschritts (Kalibrierung der Items) von den beiden folgenden Schritten (latente Regressionsmodellierung und Ziehung von *Plausible Values*) ergibt sich aus den mathematischen Eigenschaften des Rasch-Modells und seinen Erweiterungen (vgl. Abschnitt 12.4). Für die praktische Umsetzung wurden unterschiedliche Softwarepakete eingesetzt. Im ersten Teilschritt wurden die Items mit der Software *mdltm* (von Davier, 2005) kalibriert. Zur Bestimmung der Kompetenzen der Schülerinnen und Schüler durch latente Regressionsmodellierung und Ziehung von *Plausible Values* diente die Software *DGROUP* (Rogers, Tang, Lin & Kandathil, 2010). Diese Analyseprozeduren entsprechen den bereits für das landesweit in den USA durchgeführte *National Assessment of Educational Progress* (NAEP) verwendeten Prozeduren zur Kompetenzmessung (von Davier, Sinharay & Oranje, 2006). Für eine detaillierte Darstellung sei auf umfangreiche Literatur zum NAEP-Projekt (z. B. Allen, Dondghue & Schoeps, 2001; U.S. Department of Education, National Center for Education Statistics, 1999), den technischen Bericht zu PIAAC 2012 (OECD, 2013) sowie auf den erwarteten internationalen *Technical Report* zu PISA 2015 der OECD verwiesen.

© Waxmann Verlag GmbH

12.6.1 Bestimmung der Aufgabenschwierigkeit – Itemkalibrierung

Die Bestimmung der Aufgabenschwierigkeit wurde durch eine sogenannte *Concurrent Item Calibration* realisiert (etwa: gleichzeitige Itemkalibrierung, vgl. Dorans, Pommerich & Holland, 2007; Kolen & Brennan, 2014; von Davier, 2016; Winkersky & Lord, 1984). Die Datengrundlage bildet dabei der aggregierte Datenbestand der PISA-Runden 2006 bis 2015. Die *Concurrent Item Calibration* verfolgt das Ziel, die Datenbasis über alle Teilnehmerstaaten und Erhebungsmodi zu maximieren und damit eine möglichst unverzerrte Schätzung der Itemcharakteristiken (z. B. deren Schwierigkeit) zu ermöglichen. Dabei wurden einerseits die Itemparameter für die gesamte aggregierte Datenbasis als auch für jeden Teilnehmerstaat frei geschätzt, wobei die Abweichungen der staatenspezifischen Schätzungen der Parameter quantifiziert werden können (ETS, 2016). Die dahinterstehende Annahme bezieht sich auf die Invarianz der Itemparameter über die unterschiedlichen Teilnehmerstaaten (und Erhebungsmodi) hinweg.

Ferner ist es möglich, die Itemparameter einerseits nach dem restriktiveren RM oder PCM und andererseits nach dem im Hinblick auf die Itemcharakteristik informativeren 2-PL-Modell (vgl. Abschnitt 12.3) in einem iterativen Prozess quasi gleichzeitig zu bestimmen. Die Anwendung dieses Prinzips zur Itemkalibrierung wurde dabei sowohl für die internationale Gesamtstichprobe als auch für die Teilstichproben auf der Ebene einzelner an PISA 2015 teilnehmender Staaten durchgeführt. Dabei ergaben sich (Abweichungs-)Koeffizienten zur Überprüfung möglicher Interaktionen zwischen einzelnen Teilnehmerstaaten und spezifischen Parametern einzelner Items. Zur Quantifizierung der Abweichungen der staatenspezifischen Itemparameter von den internationalen Parametern wurde, wie auch in den PIAAC-Studien (OECD, 2013), die *mittlere quadratische Differenz* (*Root Mean Square Deviation*, RMSD) berücksichtigt. Die RMSD ist ein standardisierter Koeffizient, welcher die Diskrepanz zwischen der beobachteten ICC und der modellbasierten ICC quantifiziert (vgl. Gleichung 12.2).

$$RMSD_{ic} = \sqrt{\sum_{h=0}^{H}\sum_{k=1}^{K}\frac{[f_{ihc}(\theta_k) - p(X_i = h|\theta_k)]^2 \cdot \omega_c(\theta_k)}{H}} \tag{12.2}$$

Dabei steht $p(X_i = h|\theta_k)$ für die erwartete Wahrscheinlichkeit einer Antwort in der Kategorie h (bei Item i), gegeben Fähigkeitsniveau θ_k und f_{ihc} (θ_k) entspricht dem Verhältnis von Schülerinnen und Schülern in Staat C, welche in der Kategorie h bei Item i antworten. Der Ausdruck $\omega_c(\theta_k)$ stellt einen Teilnehmerstaaten spezifischen Gewichtungsfaktor dar, der dem Verhältnis von Schülerinnen und Schülern in Staat C mit Fähigkeitsniveau θ_k entspricht. Die *RMSD* ist gleichermaßen sensitiv gegenüber Abweichungen bezüglich der Lage (differenzielle Itemschwierigkeiten) als auch der Steigung (differenzielle Trennschärfen) der ICCs der Items (OECD, 2013).

© Waxmann Verlag GmbH

Diese Itemanalysen wurden im Vorfeld der Skalierung der PISA-Daten für den Zyklus 2015 vom internationalen Konsortium durchgeführt. Die hier skizzierte Prozedur der Itemkalibrierung stellt insofern eine Abweichung von den Prozeduren früherer PISA-Runden dar, als dass vor PISA 2015 zur Kalibrierung der Items jeweils nur eine Teilstichprobe des internationalen Datensatzes der jeweils aktuellen PISA-Runde herangezogen wurde (z. B. OECD, 2014).

Der oben dargestellte Abweichungskoeffizient (*RMSD*) wurde nun im Rahmen der internationalen Kalibrierung der Items eingesetzt, um zu überprüfen, inwieweit die Itemcharakteristik für einzelne Items und Teilnehmerstaaten von denen auf Basis der aggregierten Gesamtstichprobe bestimmten Itemcharakteristiken abweichen. Items, welche für den Koeffizienten RMSD einen Wert von $RMSD > .20$ aufwiesen, wurden für die spätere Skalierung (Bestimmung der Personenparameter in dem jeweiligen Teilnehmerstaat) ausgeschlossen. Items mit einem Wert für den Koeffizienten $.10 < \text{RMSD} < .20$ wurden zwar für die Bestimmung der Kompetenzschätzung beibehalten, wobei allerdings für Teilnehmerstaaten spezifische Itemparameter verwendet wurden. Diese werden als *Unique Item Parameters* bezeichnet. Die Ergebnisse dieser Itemkalibrierung auf Basis der Teilstichprobe in Deutschland im Vergleich zur internationalen Gesamtstichprobe sind im Anhang B zusammengefasst. Insgesamt wird dabei deutlich, dass für die Teilstichprobe in Deutschland vergleichsweise wenige Items durch staatenspezifische Itemparameter beschrieben werden müssen. Es sind dies für den Bereich Naturwissenschaften acht Items (von 184, vgl. Tabelle B8), für den Bereich Mathematik lediglich drei Items (von 69; vgl. Tabelle B9), und für den Bereich Lesen 15 Items (von 88, vgl. Tabelle B10).

12.6.2 Hintergrundmodell als latente Regression für Populationsmodelle und Ziehung von Plausible Values als Messwerte für Kompetenzen

Zur Bestimmung der Kompetenz werden bei PISA neben den eigentlichen Antworten der Schülerinnen und Schüler auf die einzelnen Items auch zusätzliche Informationen aus dem Schülerfragebogen herangezogen (vgl. Kapitel 13). Das Ziel dieses Vorgehens besteht darin, mithilfe der zusätzlichen Informationen aus dem Schülerfragebogen die Schätzung der Parameter für die Kompetenzmodelle zu präzisieren (z. B. Mislevy, 1987; Mislevy & Sheehan, 1987, 1989). Dieses Vorgehen bei der Bestimmung der Personenparameter (Kompetenz) im Rahmen der IRT-Skalierung wird in der methodisch orientierten Literatur als *latente Regressionsmodellierung* bezeichnet (vgl. Adams, Wu & Carstensen, 2007; von Davier et al., 2006). Das Prinzip besteht darin, für bestimmte Merkmalskombinationen, welche im Schülerfragebogen erfasst werden (sogenannte *Hintergrundvariablen*), zunächst die Parameter einer A-posteriori-Verteilung der Kompetenz unter Berücksichtigung der Lösungswahrscheinlichkeiten der eigentlichen Items zu bestimmen (Mislevy, 1987). Diese Methodik hat Bezüge zu zwei Prinzipien aus dem

© Waxmann Verlag GmbH

Bereich der Datensammlung einerseits und deren Auswertung andererseits, welche zur Präzisierung von Parameterschätzungen (hier das Kompetenzniveau) von Populationen angewendet werden können.

Diese Methodik lässt sich in Anlehnung an die Stichprobentheorie zur Parameterschätzung von Populationen begründen. Danach sind stratifizierte, das heißt nach bestimmten Merkmalen geschichtete Stichproben (vgl. auch Abschnitt 12.2), geeignet, die Schätzung der Populationsparameter zu präzisieren, soweit Vorwissen über die Verteilung dieser Merkmale und deren Zusammenhang mit dem interessierenden Populationsparameter vorhanden ist (vgl. Bortz & Döring, 2006, S. 424 ff.). Nach diesem Prinzip lässt sich durch die Aufteilung der zu erzielenden Gesamtstichprobe in homogenere Subgruppen nach bestimmten Merkmalsvariablen die Genauigkeit der Parameterschätzung für die gesamte Population verbessern. Für die meisten sozialwissenschaftlichen Untersuchungen erweisen sich dabei Merkmalsvariablen zum soziodemografischen Hintergrund wie z. B. Alter, Geschlecht und Bildungsniveau als besonderes geeignet (Bortz & Döring, 2006, S. 425). Bei der IRT-Skalierung der einzelnen PISA-Domänen im Rahmen der *latenten Regressionsmodellierung* werden nun diese Informationen (Hintergrundvariablen) dazu genutzt, die Schätzung der Verteilung der Kompetenz in der betreffenden Zielpopulation zu präzisieren.

Ein weiterer Aspekt zur Begründung der bei PISA eingesetzten Methodik der *latenten Regressionsmodellierung* und anschließenden Ziehung von *Plausible Values* (PVs) leitet sich aus der Theorie der Imputation [dt. etwa: Ersetzung] fehlender Datenpunkte in unvollständig vorliegenden Datenmatrizen ab (Rubin, 1987; von Davier, 2013). Bei der Imputation solcher fehlenden Datenpunkte in unvollständigen Datenmatrizen werden ebenfalls auf Regressionen basierende Modelle aufgestellt, um die fehlenden Daten anhand von (Hintergrund-)Variablen zu bestimmen (Mislevy, 1991). Bei der Kompetenzschätzung für PISA werden mit dem Hintergrundmodell zunächst die Parameter der A-posteriori-Verteilung der Kompetenz bestimmt. Unter der Annahme normalverteilter Kompetenzen werden auf der Basis dieser Parameter dann aus dieser Verteilung einzelne Werte bestimmt (gezogen). Diese sind im Rahmen der zuvor bestimmten Verteilungsparameter plausibel – daher *Plausible Values* (PVs). Bei PISA 2015 werden, anders als in früheren PISA-Runden, für jeden der erfassten Kompetenzbereiche 10 Plausible Values aus der A-posteriori-Verteilung der Kompetenz einer jeden Schülerin, eines jeden Schülers gezogen. Diese PVs stehen dann in ihrer Gesamtheit für die Kompetenzverteilung in der jeweiligen Population, welche das Modell beschreibt. Bei der Anwendung solcher Modelle zur Generierung von Schätzwerten für die Population wird daher auch von einem *Populationsmodell* gesprochen (von Davier & Sinharay, 2013). Die „Schätzwerte" für die Population fallen dabei durch die Anwendung der latenten Regressionsmodelle genauer aus als die Kompetenzschätzungen auf der Ebene der einzelnen Untersuchungseinheiten (der Schülerinnen und Schüler) der Stichprobe (von Davier et al., 2006).

© Waxmann Verlag GmbH

12.6.3 Definition von Kompetenzstufen

Das bei PISA in den einzelnen Domänen erfasste Kompetenzniveau der Schülerinnen und Schüler wird als kontinuierliche latente Variable angesehen. Zur einfacheren Interpretation beim Vergleich des jeweiligen Kompetenzniveaus wird diese latente Variable jeweils auf einer standardisierten Skala abgetragen. Der *Mittelwert* dieser Skala wird bei der erstmaligen Erfassung der jeweiligen Domäne als Schwerpunkt über alle OECD-Staaten hinweg auf einen Wert von $M = 500$ bei einer *Standardabweichung* von $SD = 100$ festgesetzt. Bei normalverteiltem Kompetenzniveau befinden sich in einem Intervall, das nach oben und unten durch eine Standardabweichung begrenzt ist, ungefähr 68 Prozent aller teilnehmenden Fünfzehnjährigen (vgl. Kapitel 14). Gut zwei Drittel der untersuchten Jugendlichen erreichen also einen Kompetenzwert im Bereich von 400 bis 600 Punkten.

Zur inhaltlichen Bezugnahme und Interpretation dieser Kompetenzverteilung wird eine Unterteilung des Kompetenzkontinuums in sogenannte Kompetenzstufen vorgenommen. Die Beschreibung einzelner Kompetenzstufen ergibt sich aus den Inhalten derjenigen Aufgaben, welche aufgrund ihrer Schwierigkeit innerhalb einer Kompetenzstufe liegen und mit einer festgelegten Wahrscheinlichkeit von $p = .625$ (62.5 Prozent) von den Schülerinnen und Schülern gelöst werden können. Diese Lösungswahrscheinlichkeit entspricht dem unteren *Cut-off* (der Grenze) einer Kompetenzstufe. Am oberen Cut-off einer Kompetenzstufe werden die Aufgaben mit einer Wahrscheinlichkeit von etwa 70 Prozent richtig gelöst (OECD, 2012). Die Einteilung einzelner Gruppen von Schülerinnen und Schülern zu bestimmten Kompetenzstufen zeigt damit an, welche Aufgaben von diesen mit hinreichender Sicherheit gelöst werden können, und bietet so die Möglichkeit einer inhaltlichen Interpretation des erreichten Kompetenzniveaus. Für den Kompetenzbereich Lesen wurden in der PISA-Erhebungsrunde 2000 fünf Kompetenzstufen gebildet, welche 2009 erweitert wurden (OECD, 2001, 2004, 2007, 2012). Für die Bereiche Mathematik und Naturwissenschaften wurden jeweils sechs Kompetenzstufen in den PISA-Erhebungsrunden 2003 und 2006 definiert (OECD, 2004, 2007), wobei

Tabelle 12.4: Cut-off-Werte der Kompetenzstufen in PISA 2015

Kompetenzstufe	Lesen	Mathematik	Naturwissenschaften	Problemlösen
Stufe 6	698.32	669.30	707.93	683.14
Stufe 5	625.61	606.99	633.33	618.21
Stufe 4	552.89	544.68	558.73	553.28
Stufe 3	480.18	482.38	484.14	488.35
Stufe 2	407.47	420.07	409.54	423.42
Stufe 1	334.75 (1a) 262.04 (1b)	357.77	334.94 (1a) 260.54 (1b)	358.49

© Waxmann Verlag GmbH

diese für den Bereich Naturwissenschaften ebenfalls im unteren Bereich (erste Kompetenzstufe) nochmals differenziert wurden. Die Tabelle 12.4 gibt eine Übersicht über die für PISA 2015 gültigen Cut-off-Werte der Kompetenzstufen.

12.7 Betrachtung von Trends zwischen PISA 2000 und PISA 2015

Seit der ersten PISA-Erhebung im Jahre 2000 wird jeweils einer der drei klassischen Kompetenzbereiche Lesen, Mathematik und Naturwissenschaften als Schwerpunktdomäne erfasst (vgl. Abbildung 12.1). Diese Konsistenz in der Planung und Durchführung der PISA-Studien legt eine längsschnittliche Analyse und Interpretation von etwaigen Veränderungen nahe. Allerdings unterliegt die Analyse und Betrachtung von Trends einigen Limitationen (vgl. z. B. Carstensen, Prenzel & Baumert, 2009). Diese Limitationen ergeben sich zunächst aus dem Design der Studie. So handelt es sich bei den einzelnen PISA-Erhebungsrunden um ein Kohorten-Querschnitt-Design (Diekmann, 2014; Kromrey, 2002). Bei PISA werden die Untersuchungseinheiten (Schülerinnen und Schüler) über einen durch deren Geburt bestimmten Zeitraum als Mitglieder der zu untersuchenden Kohorte definiert – eine sogenannte *Alterskohorte*. Im Gegensatz zu einem echten längsschnittlichen Untersuchungsdesign werden nicht dieselben Personen einer definierten Kohorte zu späteren Messzeitpunkten wiederholt befragt und getestet.

Daneben bestehen beim Vergleich der Ergebnisse aus PISA 2015 mit anderen PISA-Runden zusätzliche limitierende Faktoren. Dies sind die Veränderungen bei der Auswertung der Daten (vgl. Abschnitt 12.3) und auch der Wechsel des Erhebungsmodus (erstmals Darbietung der Aufgaben durch Computer – vgl. Abschnitt 12.4).

Um überhaupt quasi-längsschnittliche Betrachtungen über die einzelnen Erhebungen zu ermöglichen, müssen die einzelnen Messzeitpunkte über eine gemeinsame (Teil-)Menge von Aufgaben (inhaltlich) miteinander verbunden sein. Diese sogenannten *Link-Aufgaben* werden den Jugendlichen über die einzelnen Kohorten hinweg in inhaltlich unveränderter Form vorgelegt. Die dahinterliegende Annahme ist, dass diese Link-Aufgaben bei allen Erhebungsrunden die gleiche Messcharakteristik (Itemparameter) aufweisen. Für die Erhebung im Jahre 2015 besteht nun ein limitierender Faktor für Vergleiche darin, dass der *Darbietungsmodus* der Aufgaben im Vergleich zu allen früheren Erhebungen über ein Computersystem realisiert wurde. Allein der Wechsel des Darbietungsmodus kann hier einen Einfluss auf die Messcharakteristik einzelner Aufgaben haben (vgl. Abschnitt 12.4).

Für die technische Durchführung einer Verlinkung zweier oder mehrerer querschnittlicher Untersuchungen steht eine ganze Reihe unterschiedlicher Verfahren zur Verfügung, wobei für einen Überblick auf Dorans et al. (2007) sowie auf Kolen und Brennan (2014) verwiesen sei. In PISA 2012 und davor wurde ein einfaches *Mean-Mean Linking* eingesetzt (Kolen & Brennan, 2014). Dabei wurden die Aufgaben auf Basis einer

© Waxmann Verlag GmbH

Teilstichprobe der damals aktuellen PISA-Runde kalibriert (OECD, 2014). Zum Linking der Kompetenzmessungen wurde die mittlere Abweichung der Itemparameter der Link-Items, zu denen aus der jeweils zu vergleichenden früheren PISA-Runde als Konstante berechnet. Bei der Transformation der Kompetenzschätzer auf die PISA-Metrik (M = 500, SD = 100) für die PISA Runde 2012 wurde diese Konstante dann einbezogen. Als Resultat des einfachen *Mean-Mean Linkings* lassen sich die Kompetenzschätzer verschiedener Runden auf derselben Metrik direkt miteinander vergleichen (Mazzeo & von Davier, 2013).

Bei der gleichzeitigen Kalibrierung und Verlinkung (*Concurrent Calibration and Equating*; vgl. Winkersky & Lord, 1984), wie sie in PISA 2015 eingesetzt wurde, werden die Antworten der Schülerinnen und Schüler aus den verschiedenen Kohorten zu einer Datei kombiniert. Durch die in allen Kohorten in gleicher Weise eingesetzten Aufgaben werden die Tests (deren Items) in den einzelnen Kohorten gemeinsam kalibriert. Als Ergebnis lassen sich die geschätzten Itemparameter (z. B. deren Schwierigkeiten) auf einer gemeinsamen Skala abbilden. Dies bildet dann die Grundlage für die (eigentliche) Skalierung, d. h. die Bestimmung der Personenparameter als Schätzer für die Kompetenz. Dieses hier skizzierte Verfahren zum Linking wurde für die PISA-Runde 2015 eingesetzt. Darüber hinaus wurde bei PISA 2015 im Gegensatz zu allen früheren Runden ein erweitertes Skalierungsmodell verwendet, welches in seinen Grundzügen bereits in den Abschnitten 12.3.2 und 12.3.3 beschrieben wurde.

Wie in Abbildung 12.1 dargestellt liegen inzwischen sechs Erhebungen mit Stichproben fünfzehnjähriger Schülerinnen und Schüler vor. Aufgrund der zyklisch wechselnden Schwerpunktsetzungen ist eine quasi-längsschnittliche Vergleichbarkeit der Ergebnisse der einzelnen Erhebungsrunden zunächst immer zwischen denjenigen Erhebungsrunden sinnvoll, in denen der jeweilige Kompetenzbereich (erstmalig) den Schwerpunkt bildete oder aber (in den beiden zu vergleichenden Runden) jeweils als Nebendomäne erfasst wurde. Daraus ergibt sich, dass beispielsweise für den in PISA 2015 als Schwerpunkt erhobenen Kompetenzbereich *Naturwissenschaften* die zentrale Referenz die PISA-Erhebung im Jahre 2006 ist (vgl. Prenzel et al., 2007). Grundsätzlich ist es für die Betrachtung von quasi-längsschnittlichen Trends möglich, jede andere PISA-Runde einzubeziehen. Allerdings vergrößert sich dann aufgrund der reduzierten Anzahl gemeinsamer Aufgaben die Unsicherheit eines solchen Vergleichs. Ein Maß für diese Unsicherheit ist der sogenannte *Link-Fehler*, welcher beim Vergleich der Punktschätzer für die mittlere Kompetenz zweier PISA-Runden zu berücksichtigen ist. Tabelle 12.5 zeigt eine Übersicht dieser Link-Fehler, für alle vergangenen sowie die aktuelle PISA-Runde 2015, jeweils einzeln für die erfassten Kompetenzbereiche. Zur detaillierten Darstellung der formalen Grundlage und der Berechnungen des Link-Fehlers in der PISA-Runde 2015 sei etwa auf den Aufsatz von Rousseeuw und Croux (1993) hingewiesen, für die vergangenen Zyklen auf den technischen Bericht der PISA-Erhebungsrunde 2012 (OECD, 2014).

Für den Vergleich des durchschnittlichen Kompetenzniveaus zweier PISA-Runden im Hinblick auf eventuelle Veränderungen reicht demzufolge der Blick auf die beiden

© Waxmann Verlag GmbH

Tabelle 12.5: Robuster Link-Fehler für Trendbetrachtungen zwischen PISA 2015 und früheren Runden

Vergleich	Mathematik	Lesen	Naturwissenschaften
PISA 2000 mit PISA 2015	–	6.8044	–
PISA 2003 mit PISA 2015	5.6080	5.3907	–
PISA 2006 mit PISA 2015	3.5111	6.6064	4.4821
PISA 2009 mit PISA 2015	3.7853	3.4301	4.5016
PISA 2012 mit PISA 2015	3.5462	5.2535	3.9228

Anmerkungen: Link-Fehler Berechnung auf Basis der paarweisen Differenzstatistik *Sn* nach Rousseeuw und Croux (1993).

Punktschätzer nicht aus. Vielmehr muss unter Einbezug der jeweiligen Messunsicherheit (vgl. Abschnitt 12.5) und des Link-Fehlers berechnet werden, ob sich der numerische Unterschied zwischen den beiden Werten (den beiden Punktschätzern) zufallskritisch gegen Null absichern lässt. Formal geht es darum, den Standardfehler (*SE*) der anhand der beiden Stichproben beobachteten Differenz der Punktschätzer zu bestimmen. Über die Standardfehler der jeweiligen Punktschätzer und den in Tabelle 12.5 wiedergegebenen Link-Fehler (*LE*) ist dies vergleichsweise einfach möglich (vgl. Gleichung 12.3).

$$SE(\Delta_{2015-t}) = \sqrt{SE_{2015}{}^2 + SE_t{}^2 + LE_{2015,t}{}^2} \tag{12.3}$$

Der Index (Δ_{2015-t}) steht dabei für die Differenz zwischen der PISA-Runde 2015 und einer damit zu vergleichenden anderen Runde *t* (vgl. Tabelle 12.5). Mit dem so berechneten Standardfehler kann die Differenz zwischen zwei Punktschätzern statistisch gegen den Wert Null abgesichert werden. Dazu wird die beobachtete Differenz durch deren nach Gleichung 12.3 berechneten Standardfehler dividiert. Die errechnete Prüfgröße folgt approximativ der Standardnormalverteilung und kann mit den bekannten Signifikanzgrenzen ($-1.96 < z < +1.96$) bewertet werden. Für die Domäne *naturwissenschaftliche Kompetenz* ergibt sich zum Beispiel für den Vergleich der beiden Runden 2006 und 2015 für Deutschland zunächst eine anhand der Stichproben beobachtete Differenz von $\Delta_{2015\text{-}2006} = 509\text{-}516 = -7$ mit einem Standardfehler von $SE_{\Delta 2015\text{-}2006} = 6.5$[20]. Die Division -7 / 6.5 = -1.08 resultiert in einem Wert für die Prüfgröße *z*, welcher innerhalb des Bereiches $-1.96 > z > +1.96$ liegt, sodass sich die beobachtete Differenz von $\Delta_{2015\text{-}2006} = -7$ statistisch nicht signifikant vom Wert $\Delta_{2015\text{-}2006} = 0$ für die Population absichern lässt.

20 Der Standardfehler der Differenz errechnet sich in diesem Beispiel gemäß Gleichung 12.3 wie folgt:

$$SE(\Delta_{2015-2006}) = \sqrt{SE_{2015}{}^2 + SE_{2006}{}^2 + LE_{2015,2006}{}^2} = \sqrt{2{,}7^2 + 3{,}8^2 + 4{,}48^2} \approx 6{,}5$$

© Waxmann Verlag GmbH

12.8 Zusammenfassung und Ausblick

Aus den in diesem Kapitel beschriebenen Gründen ist bei der inhaltlichen Interpretation von Entwicklungstrends in Bezug zu PISA 2015 besondere Vorsicht angezeigt. Neben den bei allen PISA-Runden vorhandenen Unsicherheitsfaktoren, welche über entsprechende *Standardfehler* und *Link-Fehler* berücksichtigt wurden, gibt es, bezogen auf die Runde 2015, Gesichtspunkte, die zusätzlich bei der Interpretation von Entwicklungstrends im Vergleich zu früheren Runden zu berücksichtigen sind. So wurden in PISA 2015 mehrere methodische Neuerungen bei der Testauswertung und Datenerhebung eingeführt. Dies sind zum Beispiel der Wechsel vom Rasch-Modell zum 2-PL-Modell bei der Auswertung der PISA-Testaufgaben (vgl. Abschnitt 12.3), die Einbeziehung von historischen und aktuellen Daten bei der Kalibrierung der Items – im Gegensatz zur Verwendung von Teilstichproben zur Itemkalibrierung in vergangenen Runden – sowie nicht zuletzt der Wechsel des Erhebungsmodus von papierbasierter Erhebung zu computerbasierter Erhebung (vgl. Abschnitt 12.4).

Was eine abschließende Bewertung der in PISA durchgeführten Analysen eines möglichen Moduseffekts betrifft, fehlen noch detailliertere Informationen zu (in-)varianten Items in Feldtest und Hauptstudie (je Domäne und teilnehmendem Staat) inklusive der jeweiligen Klassifikationskriterien. Mit dem angekündigten *Technical Report* der OECD wird diese Lücke voraussichtlich zumindest zum Teil geschlossen werden können. Auf dieser Informationsgrundlage kann insgesamt die Robustheit des (neuen) methodischen Vorgehens weiter beurteilt und mit alternativen Möglichkeiten der Behandlung von Moduseffekten (etwa der expliziten Modellierung oder einem Linking) verglichen werden.

Zusammenfassend ist festzuhalten, dass die beschriebenen Veränderungen in der sechsten PISA-Erhebungsrunde letztlich mit Blick auf die Zukunft der Studie und deren Anspruch, Kompetenzen nah an der Lebenswelt fünfzehnjähriger Schülerinnen und Schüler zu messen, eingeführt wurden. Die damit verbundenen Vorteile bringen eine Reihe von Einschränkungen mit sich, die bei der Interpretation der Befunde zu berücksichtigen sind. Insofern musste bei der Planung und Auswertung von PISA 2015 ein schwieriger Abwägungsprozess zwischen der Einführung neuer, verbesserter psychometrischer Methoden einerseits und der Sicherung der Vergleichbarkeit historischer PISA-Ergebnisse andererseits gefunden werden. Die letztlich getroffene Entscheidung zur Implementierung zahlreicher technischer und inhaltlicher Neuerungen in PISA 2015 verbessert sicherlich die psychometrische Qualität der Messung (vgl. z. B. Oliveri & von Davier, 2011, 2014). Insofern erscheinen mit Blick auf die Zukunft der PISA-Studien die Einschränkungen (z. B. vergrößerte Link-Fehler) beim längsschnittlichen Vergleich hinnehmbar zu sein.

© Waxmann Verlag GmbH

Literatur

Adams, R. (2011). *Comments on Kreiner 2011: Is the foundation under PISA solid? A critical look at the scaling model underlying international comparisons of student attainment.* Zugriff am 11.10.2016. Verfügbar unter http://www.oecd.org/pisa/47681954.pdf

Adams, R. J. (2005). Reliability as a measurement design effect. *Studies in Educational Evaluation*, *31* (2–3), 162–172. http://dx.doi.org/10.1016/j.stueduc.2005.05.008

Adams, R. J., Wu, M. L. & Carstensen, C. H. (2007). Application of multivariate Rasch models in international large-scale educational assessments. In M. von Davier & C. H. Carstensen (Hrsg.), *Multivariate and mixture distribution Rasch models: Extensions and applications* (1. Aufl., S. 271–280). New York: Springer US. http://dx.doi.org/10.1007/978-0-387-49839-3_17

Aitkin, I. & Aitkin, M. (2011). *Statistical modeling of the national assessment of educational progress*. New York: Springer. Zugriff am 18.10.2016 auf http://link.springer.com/10.1007/978-1-4419-9937-5

Allen, N. L., Dondghue, J. R. & Schoeps, T. L. (Hrsg.). (2001). *The NAEP 1998 technical report*. Zugriff am 11.10.2016. Verfügbar unter https://nces.ed.gov/nationsreportcard/pdf/main1998/2001509.pdf

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, *43* (4), 561–573.

Balakrishnan, N. & Rao, C. R. (2001). *Handbook of statistics (Bd. 20): Advances in reliability*. Oxford: Elsevier.

Bechger, T. M., Maris, G. & Hsiao, Y. P. (2010). Detecting halo effects in performance-based examinations. *Applied Psychological Measurement*, *34* (8), 607–619. http://dx.doi.org/10.1177/0146621610367897

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick (Hrsg.), *Statistical theories of mental test scores* (S. 395–479). Reading, MA.: Addison-Wesley.

Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6. Aufl.). Heidelberg: Springer.

Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.

Bortz, J. & Schuster, C. (2010). *Statistik für Human- und Sozialwissenschaftler* (7. Aufl.). Berlin: Springer.

Brown, R. S. (2010). Sampling. In P. P. McGaw (Hrsg.), *International encyclopedia of education* (3. Aufl., S. 142–146). Oxford: Elsevier.

Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3. aktual. u. erw. Aufl.). München: Pearson.

Bürger, S., Kröhne, U. & Goldhammer, F. (2016). *The transition to computer-based testing in large-scale assessments: Investigating (partial) measurement invariance between modes*. Manuscript submitted for publication.

Carstensen, C. H., Prenzel, M. & Baumert, J. (2009). Trendanalysen in PISA: Wie haben sich die Kompetenzen in Deutschland zwischen PISA 2000 und PISA 2006 entwickelt? In M. Prenzel & J. Baumert (Hrsg.), *Vertiefende Analysen zu PISA 2006*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Cochran, W. G. (1977). *Sampling techniques* (3. Aufl.). New York: John Wiley.

© Waxmann Verlag GmbH

Daniel, J. (2012). *Sampling essentials – practical guidelines for making sampling choices*. Thousand Oaks: Sage.

DeMars, C. (2010). *Item response theory*. Oxford, New York: Oxford University Press.

Diekmann, A. (2014). *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen* (9. Aufl.). Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.

Dorans, N. J., Pommerich, M. & Holland, P. W. (Hrsg.). (2007). *Linking and aligning scores and scales*. New York: Springer.

ETS = Educational Testing Service. (2015a). *PISA 2015 Field trial analysis report: Outcomes of the cognitive assessment* (interner Bericht). Princeton, NJ.

ETS = Educational Testing Service. (2015b). *PISA 2015 main study coding procedures – CBA Ccountries* (interner Bericht). Princeton, NJ.

ETS = Educational Testing Service. (2016). *Main survey – notes on IRT scaling and analysis* (interner Bericht). Princeton, NJ.

Faulbaum, F. (2014). Total Survey Error. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer VS.

Fay, R. E. (1989). Theory and application of replicate weighting for variance calculations. In *Proceedings of the American Statistical Association* (S. 212–217). Washington, D.C.: Alexandria, VA, American Statistical Association.

Fischer, G. H. (1974). *Einführung in die Theorie psychologischer Tests*. Bern: Huber.

Fischer, G. & Molenaar, I. (Hrsg.). (1995). *Rasch models: Foundations, recent developments, and applications*. New York: Springer.

Fisher, R. A. & Yates, J. F. (1963). *Statistical tables for biological, agricultural and medical research*. Edinburgh: Oliver and Boyd.

Frey, A. (2012). Adaptives Testen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2. aktual. u. überarb. Aufl., S. 275-293). Berlin, Heidelberg: Springer.

Frey, A. & Bernhardt, R. (2012). On the importance of using balanced booklet designs in PISA. *Psychological Test and Assessment Modeling*, *54* (4), 397–417. Zugriff am 11.10.2016. Verfügbar unter http://www.psychologie-aktuell.com/fileadmin/download/ptam/4-2012_20121224/05_Frey.pdf

Frey, A., Hartig, J. & Rupp, A. A. (2009). An NCME instructional module on booklet designs in large-scale assessments of student achievement: Theory and practice. *Educational Measurement: Issues and Practice*, *28* (3), 39–53. http://dx.doi.org/10.1111/j.1745-3992.2009.00154.x

Gelman, A. (2005). Analysis of Variance: Why It Is More Important than Ever. *The Annals of Statistics*, *33* (1), 1–31. http://doi.org/10.2307/3448650

Goldhammer, F., Naumann, J., Rölke, H., Stelter, A. & Tóth, K. (im Druck). 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*. Heidelberg: Springer.

Groves, R. M. (2005). *Survey errors and survey costs*. New York: Wiley & Sons.

Häder, M. & Häder, S. (2014). Stichprobenziehung in der quantitativen Sozialforschung. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 283–297). Wiesbaden: Springer VS.

Häder, M. (2015). *Empirische Sozialforschung – eine Einführung* (3. Aufl.). Wiesbaden: Springer VS.

© Waxmann Verlag GmbH

Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory: principles and applications*. Boston: Hingham, MA: Kluwer-Nijhoff Pub.

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. 309–345). Münster: Waxmann.

Johnson, M. C. & Lord, F. M. (1958). An empirical study of the stability of a group mean in relation to the distribution of test items among students. *Educational and Psychological Measurement*, *18* (2), 325–329. http://doi.org/10.1177/001316445801800209

Judkins D. R. (1990). Fay's method for variance estimation. *Journal of Official Statistics*, *6*, 223–239. Zugriff am 11.10.2016. Verfügbar unter http://www.jos.nu/Articles/abstract.asp?article=63223

Kalton, G. (1983). *Introduction to survey sampling*. Newbury Park: Sage.

Kish, L. (1965). *Survey sampling*. New York: Wiley & Sons.

Kish, L. (1995). *Survey sampling*. New York: Wiley & Sons.

KMK = Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2015). *Definitionenkatalog zur Schulstatistik 2015* (2). Zugriff am 18.07.2016. Verfügbar unter http://www.kmk.org/statistik/schule/statistische-veroeffentlichungen/definitionenkatalog-zur-schulstatistik.html.

Kolen, M. J. & Brennan, R. L. (2014). *Test equating, scaling, and linking*. New York, NY: Springer.

Krebs, D. & Menold, N. (2014). Gütekriterien quantitativer Sozialforschung. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 425–438). Wiesbaden: Springer VS.

Kreiner, S. (2011). *Is the foundation under PISA solid? A critical look at the scaling model underlying international comparisons of student attainment.* University of Copenhagen, Department of Biostatistics.

Kreiner, S. & Christensen, K. B. (2014). Analyses of model fit and robustness. A new look at the PISA scaling model underlying ranking of countries according to reading literacy. *Psychometrika*, *79* (2), 210–231. http://dx.doi.org/10.1007/s11336-013-9347-z

Kröhne, 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), 169–186. http://dx.doi.org/10.1007/s11618-011-0185-4

Kromrey, H. (2002). *Empirische Sozialforschung* (10. Aufl.). Opladen: Leske + Budrich.

Lee, Y.-H. & Chen, H. (2011). A review of recent response-time analyses in educational testing. *Psychological Test and Assessment Modeling*, *53* (3), 359–379. Zugriff am 11.10.2016. Verfügbar unter http://p16277.typo3server.info/fileadmin/download/ptam/3-2011_20110927/06_Lee.pdf

Levy, P. S. & Lemeshaw, S. (2008). *Sampling of populations – methods and applications* (4. Aufl.). Hoboken, NJ: Wiley & Sons.

Lienert, G. A. & Raatz, U. (1994). *Testaufbau und Testanalyse*. Weinheim: Beltz, PVU.

Lord, F. M. (1962). Estimating Norms by Item-Sampling. *Educational and Psychological Measurement*, *22* (2), 259–267. http://dx.doi.org/10.1002/j.2333-8504.1961.tb00103.x

Lord, F. M. (1965). Item sampling in test theory and in research design. *ETS Research Bulletin Series*, *1965* (2), i-39. http://dx.doi.org/10.1002/j.2333-8504.1965.tb00968.x

Luke, D. A. (2004). *Multilevel modeling*. Thousand Oaks, CA: SAGE.

© Waxmann Verlag GmbH

Macaskill, G. (2008). *Alternative scaling models and dependencies.* Präsentiert auf dem TAG Meeting, Sydney Australia. Zugriff am 11.10.2016. Verfügbar unter https://www.acer.edu.au/files/macaskill_alternativescalingmodelsdependenciespisa(1).pdf

Masters, G. N. (1982). A rasch model for partial credit scoring. *Psychometrika*, *47* (2), 149–174. http://dx.doi.org/10.1007/bf02296272

Masters, G. N. (1988). Item discrimination: When more is worse. *Journal of Educational Measurement*, *25* (1), 15–29. http://dx.doi.org/10.1111/j.1745-3984.1988.tb00288.x

Mazzeo, J. & Von Davier, M. (2013). Linking scales in international large-scale assessments. In L. Rutkowski, M. von Davier & D. Rutkowski (Hrsg.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (S. 229–258). Boca Raton, FL: CRC Press.

Miller, M. D. (2010). Classical test theory reliability. In P. P. McGaw (Hrsg.), *International encyclopedia of education* (3. Aufl., S. 27–30). Oxford: Elsevier. http://dx.doi.org/10.1016/b978-0-08-044894-7.00235-9

Mislevy, R. (1987). Exploiting auxiliary infornlation about examinees in the estimation of item parameters. *Applied Psychological Measurement*, *11* (1), 81–91. http://dx.doi.org/10.1177/014662168701100106

Mislevy, R. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, *56* (2), 177–196. http://dx.doi.org/10.1007/bf02294457

Mislevy, R. J. & Sheehan, K. M. (1987). Marginal estimation procedures. In A. E. Beaton (Hrsg.), *Implementing the new design: The NAEP 1983-84 technical report* (S. 293–360). Princeton, N.J: National Assessment of Educational Progress, ETS.

Mislevy, R. J. & Sheehan, K. M. (1989). The role of collateral information about examinees in item parameter estimation. *Psychometrika*, *54* (4), 661–679. http://dx.doi.org/10.1007/bf02296402

Moosbrugger, H. & Kelava, A. (Hrsg.). (2012). *Testtheorie und Fragebogenkonstruktion* (2., aktual. u. überarb. Aufl). Berlin: Springer.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, *16* (2), 159–176. http://dx.doi.org/10.1177/014662169201600206

Murphy, K. R. & Davidshofer, C. O. (2005). *Psychological testing: Principles and applications* (6. Aufl.). New Jersey: Pearson Education International.

OECD. (2001). *Knowledge and skills for life*. Paris: OECD.

OECD. (2004). *Learning for tomorrow's world. First results from PISA 2003* (Bd. 1). Paris: OECD.

OECD. (2007). *PISA 2006. Science competencies for tomorrow's world* (Bd. 2). Paris: OECD.

OECD. (2012). *PISA 2009 technical report*. Paris: OECD. Zugriff am 11.10.2016. Verfügbar unter https://www.oecd.org/pisa/pisaproducts/50036771.pdf

OECD. (2013). *Technical report of the survey of adult skills (PIAAC)*. Paris: OECD. Zugriff am 11.10.2016. Verfügbar unter http://www.oecd.org/skills/piaac/_Technical%20Report_17OCT13.pdf

OECD. (2014). *PISA 2012 technical report*. Paris: OECD. Zugriff am 11.10.2016. Verfügbar unter https://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf

OECD. (2016a). *PISA 2015 assessment and analytical framework. Science, reading, mathematic and financial literacy*. Paris: OECD. http://dx.doi.org/10.1787/9789264255425-9-en

OECD. (2016b). PISA 2015 integrated design. (interner Bericht) Zugriff am 11.10.2016. Verfügbar unter https://www.oecd.org/pisa/pisaproducts/2015-Integrated-Design.pdf

© Waxmann Verlag GmbH

Oliveri, M. E., Olson, B. F., Ercikan, K. & Zumbo, B. D. (2012). Methodologies for investigating item- and test-level measurement equivalence in international large-scale assessments. *International Journal of Testing*, *12* (3), 203–223. http://dx.doi.org/10.1080/15305058.2011.617475

Oliveri, M. E. & von Davier, M. (2011). Investigation of model fit and score scale comparability in international assessments. *Psychological Test and Assessment Modeling*, *53* (3), 315–333. Zugriff am 11.10.2016. Verfügbar unter http://www.psychologie-aktuell.com/fileadmin/download/ptam/3-2011_20110927/04_Oliveri.pdf

Oliveri, M. E. & von Davier, M. (2014). Toward increasing fairness in score scale calibrations employed in international large-scale assessments. *International Journal of Testing*, *14* (1), 1–21. http://dx.doi.org/10.1080/15305058.2013.825265

Osterlind, S. J. (1990). Toward a uniform definition of a test item. *Educational Research Quarterly*, *14* (4), 2–5.

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, statistics for social and behavorial sciences* (S. 215–230). http://dx.doi.org/10.1007/978-0-387-85461-8_11

Pommerich, M. (2004). Developing computerized versions of paper-and-pencil tests: Mode effects for passage-based tests. *The Journal of Technology, Learning, and Assessment, 2* (6), 3–44. Zugriff am 11.10.2016. Verfügbar unter https://ejournals.bc.edu/ojs/index.php/jtla/article/viewFile/1666/1508

Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E. & Pekrun, R. (Hrsg.). (2007). *PISA 2006: Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster: Waxmann.

Pugh, R. C. (1971). Empirical evidence on the application of Lord's sampling technique to Likert items. *Journal of Experimental Education*, *39*, 54–56. http://dx.doi.org/10.1080/00220973.1971.11011265

Rao, J. N. K. (1988). 17 Variance estimation in sample surveys. In C. R. Rao & Krishnaiah, P.R. (Hrsg.), *Handbook of statistics* (Bd. 6, S. 427–447). Amsterdam, New York: Elsevier. http://dx.doi.org/10.1016/s0169-7161(88)06019-5

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Copenhagen: Danmarks pædagogiske Institut.

Rogers, A., Tang, C., Lin, M.-J. & Kandathil, M. (2010). DGROUP (computer software). Princeton, NJ: Educational Testing Service.

Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion* (2., vollst. überarb. erw. Aufl.). Bern: Huber.

Rousseeuw, P. J. & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, *88* (424), 1273. http://doi.org/10.2307/2291267

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Hoboken, NJ: Wiley.

Rust, K. & Rao, J. (1996). Variance estimation for complex surveys using replication techniques. *Statistical Methods in Medical Research*, 5 (3), 283–310. http://dx.doi.org/10.1177/096228029600500305

Rutkowski, L., Gonzalez, E., von Davier, M. & Zhou, Y. (2013). Assessment design for international large-scale assessments. In L. Rutkowski, M. von Davier & D. Rutkowski (Hrsg.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (S. 75–95). Boca Raton, FL: CRC Press.

© Waxmann Verlag GmbH

Scalise, K. (2012). Creating innovative assessment items and test forms. In R. W. Lissitz & H. Jiao (Hrsg.), *Computers and their impact on state assessment: Recent history and predications for the future* (S. 134–156). Charlotte, NC: Information Age Publisher.

Schmidt, F. L. & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods, 1* (2), 199–223. http://dx.doi.org/10.1037/1082-989x.1.2.199

Sijtsma, K. & Hemker, B. T. (2000). A taxonomy of IRT models for ordering persons and items using simple sum scores. *Journal of Educational and Behavioral Statistics*, *25* (4), 391–415. http://dx.doi.org/10.2307/1165222

Sijtsma, K. & Molenaar, I. W. (2002). *Introduction to Nonparametric Item Response Theory*. Thousand Oaks: SAGE.

Stanat, P., Böhme, K., Schipolowski, S. & Haag, N. (2016). *IQB-Bildungstrend 2015. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich*. Münster: Waxmann.

Statistisches Bundesamt & DESTATIS. (2014). *Bildung und Kultur: Allgemeinbildende Schulen. Schuljahr 2013/2014 (Fachserie 11, Reihe 1)*. Wiesbaden: Statistisches Bundesamt. Zugriff am 11.10.2016. Verfügbar unter https://www.destatis.de/DE/Publikationen/Thematisch/BildungForschungKultur/Schulen/AllgemeinbildendeSchulen2110100147004.pdf?__blob=publicationFile

Thompson, S. K. (2012). *Sampling* (3. Aufl.). Hoboken, NJ: Wiley & Sons.

U.S. Department of Education, National Center for Education Statistics. (1999). *Directory of NAEP Publications. NCES 1999–489*. Zugriff am 11.10.2016. Verfügbar unter http://eric.ed.gov/?id=ED429128

van der Linden, W. J. & Hambleton, R. K. (Hrsg.). (2010). *Handbook of modern item response theory* (2. Aufl.). New York: Springer.

von Davier, M. (2005). *A general diagnostic model applied to language testing data*. Princeton, NJ: Educational Testing Service.

von Davier, M. (2013). Imputing profciency data under planned missingness in population models. In L. Rutkowski, M. von Davier & D. Rutkowski (Hrsg.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (S. 175–201). Boca Raton, FL: CRC Press.

von Davier, M. (2016). *Persönliche Kommunikation während des ETS Workshops*. Princeton: NJ.

von Davier, M. & Sinharay, S. (2013). Analytics in international large-scale assessments: Item response theory and population models. In L. Rutkowski, M. von Davier & D. Rutkowski (Hrsg.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (S. 155–174). Boca Raton, FL: CRC Press.

von Davier, M., Sinharay, S. & Oranje, A. (2006). The statistical procedures used in national assessment of educational progress: Recent developments and future directions. In C. R. Rao & S. Sinharay (Hrsg.), *Handbook of statistics 26: Psychometrics* (S. 1039–1055). Amsterdam, New York: Elsevier.

Winkersky, M. S. & Lord, F. M. (1984). An investigation of methods for reducing sampling error in certain IRT procedures. *Applied Psychological Measurement*, *8*, 347–364. http://dx.doi.org/10.1002/j.2330-8516.1983.tb00028.x

Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe.

© Waxmann Verlag GmbH

Wolter, K. M. (1985). *Introduction to variance estimation*. New York: Springer.
Wright, B. D. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement*, *14* (2), 97–116. http://dx.doi.org/10.1111/j.1745-3984.1977.tb00031.x

© Waxmann Verlag GmbH

13 Die Erhebung von Lern- und Lehrkontexten, Hintergrundmerkmalen und nichtleistungsbezogenen Lernergebnissen in PISA 2015

Nina Jude, Eckhard Klieme, Susanne Kuger & Fabian Zehner

Die sogenannten *Hintergrund-Fragebögen* haben in den PISA-Erhebungen wachsende Bedeutung. Die Auskünfte von Lernenden, Lehrkräften, Schulleitungen und Eltern werden benötigt, um den sozialen Hintergrund, die Zuwanderungsgeschichte und den Bildungsverlauf zu rekonstruieren sowie Lehr-Lernprozesse, schulische Rahmenbedingungen und die Steuerung des Schulsystems zu beschreiben. Neben diesen Kontextfaktoren werden auch wichtige Lernergebnisse wie Interessen und Lernmotivation erfasst. Das Kapitel skizziert die Rahmenkonzeption der international eingesetzten Fragebögen, für die erstmals ein Team aus Deutschland verantwortlich zeichnete. Abschließend diskutiert das Kapitel Fragen der Qualitätssicherung, die Vergleichbarkeit von Daten über Staaten hinweg, den Computereinsatz und den Datenschutz sowie die Grenzen der Aussagekraft von Fragebögen.

PISA hat die Aufgabe, Informationen über die Bildungssysteme der jeweiligen Teilnehmerstaaten zu sammeln, zu analysieren und bereitzustellen. Wesentliche Ergebnisse beziehen sich auf Kompetenzen, die Schülerinnen und Schüler gegen Ende der Pflichtschulzeit in bestimmten Domänen erreicht haben und die anhand von Tests gemessen werden. Darüber hinaus spielen in PISA wie in allen Large-Scale-Assessments sogenannte *Kontextmerkmale* eine wesentliche Rolle. Diese Kontextmerkmale beschreiben Aspekte des Lehr-Lernkontexts auf verschiedenen Ebenen des Bildungssystems, die aus Sicht der Lehr-Lernforschung und der Bildungspolitik relevant sind. Informationen über Kontextmerkmale können dazu beitragen, Unterschiede in den erreichten Kompetenzen zu erklären. Darüber hinaus liefern sie Wissen über Schulen, etwa über deren technische Ausstattung, die Ausbildung der Lehrpersonen, die Lernatmosphäre für Schülerinnen und Schüler und die Frage, wie stark Eltern in der Schule einbezogen werden. Auf Grundlage dieses Wissens können Politik und Schulleitungen entsprechende Maßnahmen einleiten, und die Bildungsforschung kann auf aktuelle Problemfelder aufmerksam machen. Wenn in international vergleichenden Schulleistungsstudien Kontextmerkmale erfasst werden, können empirische Bildungsforschung und Bildungspolitik miteinander verknüpft werden.

© Waxmann Verlag GmbH

Kontextmerkmale werden in PISA traditionell über Fragebögen erhoben. Anders als bei Kompetenztests geht es hier um Selbstberichte der Akteure im Bildungssystem. Fragebögen haben den Vorteil, viele verschiedene Aspekte des Lehr-Lernkontextes standardisiert und effizient bei unterschiedlichen Zielgruppen und damit aus verschiedenen Perspektiven erfragen zu können. In PISA sind dies in der Regel die teilnehmenden Schülerinnen und Schüler sowie die Schulleitungen der entsprechenden Schulen. Darüber hinaus bietet PISA den Teilnehmerstaaten optional an, Fragebögen für weitere Zielgruppen einzusetzen. In Deutschland wurden für PISA 2015 die Eltern der Schülerinnen und Schüler unter anderem nach den häuslichen Lernumgebungen befragt (vgl. Kapitel 10). Darüber hinaus nahmen – eine Neuerung für PISA auf internationaler Ebene – Lehrpersonen an einer Befragung teil. Sie wurden speziell nach ihrer Sicht auf den Fachunterricht und die Schulqualität, aber auch nach dem eigenen Ausbildungsweg gefragt. Schülerinnen und Schüler beantworteten ausgewählte Fragen der Optionen „Educational Career“ und „ICT Familiarity“. Diese Fragebögen beleuchteten zusätzlich den Bildungsweg der Schülerinnen und Schüler, etwa, ob sie bislang Nachhilfeunterricht erhalten haben, sowie ihre Erfahrungen und Nutzungsgewohnheiten mit digitalen Geräten, also mit Informations- und Kommunikationstechnologien.

Abgesehen davon, dass mithilfe von Fragebögen Kontextmerkmale erhoben werden, dienen sie auch dazu, die Lernergebnisse im motivationalen und affektiven Bereich zu erfassen: Hierzu gehören zum Beispiel das Interesse am Fach oder fachbezogene Einstellungen sowie weitere psychosoziale Merkmale wie etwa Wohlbefinden, Selbstkonzept und Lernstrategien. Diese werden im vorliegenden Buch vornehmlich in den Kapiteln 5, 6 und 7 vorgestellt.

Insgesamt enthält der internationale PISA-2015-Datensatz 112 Skalen und Indizes, die aus den Fragebögen gewonnen wurden. Die Erkenntnisse aus den Fragebögen sind somit ein wesentlicher Bestandteil der Berichterstattung von PISA. Über die bislang sechs PISA-Erhebungen hinweg hat sich die Bedeutung der Befragungen erhöht (Jude, im Druck). Dabei ist zu berücksichtigen, dass, vergleichbar mit den Kompetenztests, auch die Fragebögen für jede PISA-Erhebung intensiv überarbeitet werden. Für PISA 2015 lag die Verantwortung für die Entwicklung der theoretischen Rahmenkonzeption (Framework) und der Fragebögen beim *Deutschen Institut für Internationale Pädagogische Forschung* (DIPF), das hierzu von der OECD nach einer wettbewerblichen Ausschreibung beauftragt wurde. Das DIPF wurde durch die international und interdisziplinär zusammengesetzte *Questionnaire Expert Group* (QEG) beraten und arbeitete bei der Entwicklung von neuen Fragen speziell für die Bereiche der Naturwissenschaften im Rahmen des *Zentrums für internationale Bildungsvergleichsstudien* (ZIB) eng mit der Technischen Universität München zusammen.

In diesem Kapitel werden einleitend die theoretischen Grundlagen der PISA-Fragebogenentwicklung dargestellt. Darauf aufbauend werden das technische Vorgehen bei der Entwicklung der Fragebögen für PISA 2015 sowie die Auswertungsroutinen (Skalierung) erläutert. Abschließend wird auf die Besonderheiten der computerbasierten Erfassung und die Befragung der Lehrpersonen eingegangen.

© Waxmann Verlag GmbH

13.1 Konzeptionelle Grundlagen der Fragebögen für PISA

Alle in PISA eingesetzten Erhebungsinstrumente, Fragebögen und Testverfahren basieren auf theoretischen Rahmenkonzeptionen, den sogenannten *Frameworks*. Diese Rahmenkonzeptionen werden von den verschiedenen internationalen Expertengruppen im Vorlauf eines neuen Erhebungszyklus entwickelt und dokumentieren die Grundlagen der Test- und Fragebogenkonstruktion.

Während die Rahmenkonzeptionen der Testverfahren in PISA seit der ersten Erhebung des Jahres 2000 detailliert dokumentiert wurden (vgl. Kapitel 12), fanden die damals noch als „Hintergrundfragebogen" bezeichneten Fragebogeninstrumente nur am Rande Erwähnung. Obwohl bereits im Jahr 2000 verschiedene Indikatoren der Lehr-Lernumgebung erhoben wurden, lag der Schwerpunkt der Auswertung auf den Merkmalen der sozialen Herkunft und des Zuwanderungshintergrunds sowie den Geschlechtsunterschieden in Bezug auf die erzielten Kompetenzmittelwerte.

Eine separate theoretische Herangehensweise an die Kontextfragebögen, die ihre Grundlage in der empirischen Schul- und Unterrichtforschung hat, wurde in PISA zum ersten Mal für die Erhebung im Jahr 2009 dokumentiert (OECD, 2009). Verantwortliche Autoren waren die Schulforscher Jaap Scheerens und Henry Levin. Sie erläuterten in einem eigenen Kapitel dieser Rahmenkonzeption insbesondere die Mehrebenenstruktur der Lehr-Lernumgebungen. Hervorgehoben wurde dabei vor allem, dass es wichtig ist, Kontextfaktoren auf der Ebene des Bildungssystems, der Schule, auf Unterrichtsebene sowie auf Individualebene zu erheben. Entsprechend sollte die Fragebogenkonstruktion berücksichtigen, über welche Informationen Schülerinnen und Schüler Auskunft geben können und welche bei den Schulleitungen erfragt werden müssen.

PISA 2009 war auch der erste Erhebungszyklus, der die Kontextmerkmale umfassend in der Berichtlegung berücksichtigte. Die internationale Berichterstattung der OECD ging in separaten Publikationen auf die Bedeutung des sozialen Hintergrunds, individuelle Lernstrategien und Motivation sowie Kontextfaktoren auf der Schulebene ein (vgl. etwa den detaillierten Bericht zu Schulmerkmalen; OECD, 2010). Die zentralen Indikatoren für diese Berichte wurden mit den Kontextfragebögen für Schülerinnen und Schüler sowie Schulleitungen erhoben.

Um Entwicklungen im Lehr-Lernkontext auf den verschiedenen Ebenen über die Zeit zu dokumentieren, so die Autoren des Frameworks zu PISA 2009, sei es wichtig, politisch relevante Merkmale (Kernindikatoren) in regelmäßigen Abständen zu erheben. Da die Rahmenkonzeptionen für jeden Erhebungszyklus überarbeitet werden, bedarf es dazu einer Definition der Kernindikatoren und ihrer Einbettung im jeweiligen theoretischen Rahmenkonzept.

Diese Forderung wurde im Framework für PISA 2012 (Klieme et al., 2013) systematisch umgesetzt (vgl. auch Klieme & Vieluf, 2013). Eine leitende These war die folgende: Mithilfe der PISA-Fragebögen sollen kontinuierlich politisch relevante Indikatoren auf Systemebene zu jedem Messzeitpunkt erfasst und zum anderen je nach Schwerpunkt des Zyklus spezifische Indikatoren auf Schul- und Individualebene ergänzt werden. Darüber

© Waxmann Verlag GmbH

hinaus unterscheidet das Framework zu PISA 2012 noch deutlicher zwischen Ausgangsmerkmalen, Prozessmerkmalen und Ergebnisindikatoren und bezieht sich dabei auf entsprechende Modelle der Schul- und Unterrichtsforschung.

Für die Konzipierung der Fragebögen in PISA 2015 wurden die bereits existierenden Rahmenkonzepte aus 2009 und 2012 weiterentwickelt (Klieme & Kuger, 2016). Ausgangspunkt war dabei zum einen die Maßgabe, Kernindikatoren beizubehalten, um Veränderungen messen zu können. Zum anderen wurden die Teilnehmerstaaten dazu befragt, welche Inhaltsbereiche und Indikatoren für PISA 2015 als besonders relevant empfunden wurden und daher ergänzt werden sollten.

Die in PISA erhobenen Konstrukte wurden dafür 19 politisch relevanten Bereichen zugeordnet und entsprechend in Modulen strukturiert. Diese lassen sich analog zu PISA 2009 und PISA 2012 sowohl hinsichtlich der Mehrebenenstruktur des Bildungssystems als auch in Bezug auf Ausgangsmerkmale, Prozesse und Ergebnisse beschreiben. Die in PISA 2015 berücksichtigten Module sind in Abbildung 13.1 dargestellt. Sie sollen im Folgenden kurz skizziert werden. Umfassendere theoretische Konzeptionen und die Gesamtheit der jeweils zu diesen Modulen entwickelten Fragen werden zeitgleich zum vorliegenden Bericht in einer Buchpublikation mit zugehöriger Online-Datenbank veröffentlicht (Kuger, Klieme, Jude & Kaplan, im Druck).

Da bei PISA 2015 Naturwissenschaften im Fokus standen, ist ein wesentlicher Teil der Module darauf ausgelegt, das naturwissenschaftliche Lehren und Lernen im Detail abzubilden. Hierbei werden zunächst die Ressourcen erfasst, also die materielle und personelle Ausstattung der Schulen, der Aufbau des Curriculums und die vorgesehene Lernzeit in der Schule, aber auch die außerhalb der Schule verfügbaren Lernerfahrungen.

Ein besonderer Schwerpunkt liegt auf der *Qualifikation*, den *professionellen Einstellungen* und den *Handlungen* der Lehrkräfte in den Fächern Physik, Chemie und Biologie. Hierzu wurden vor allem die Lehrkräfte selbst befragt (siehe Abschnitt 4 unten), aber auch die Schulleitungen. Das *Lehr- und Lerngeschehen im Unterricht* wurde aus der Perspektive der Lehrenden und der Lernenden erhoben, wobei darauf geachtet wurde, die große Bandbreite von Unterrichtspraktiken in verschiedenen Ländern darzustellen: von traditionellen lehrergesteuerten Praktiken bis zu schüleraktivierenden Praktiken wie etwa Schülerexperimenten, um sowohl den kognitiven Anregungsgehalt als auch die sozioemotionale Unterstützung zu erfassen (Müller et al., im Druck).

So werden auch vielfältige Ergebnismerkmale erhoben: Bei einem wichtigen Modul geht es darum, *Interessen, Einstellungen und Motivation in Bezug auf Naturwissenschaften* zu erheben. Das betrifft sowohl die Lernmotivation in der Schule als auch das Interesse an naturwissenschaftlichen Themen und das Bewusstsein um deren gesellschaftliche Bedeutung, insbesondere in Bezug auf Umweltfragen (Schiepe-Tiska et al., im Druck). Da ökologisches Problembewusstsein und Einstellungen bereits im Rahmen von PISA 2006 erfragt worden waren, lassen sich nun Veränderungen über neun Jahre hinweg darstellen. Im Bereich des *naturwissenschaftlichen Lehrens und Lernens* knüpfen die

© Waxmann Verlag GmbH

	Ausgangsmerkmale		Prozessmerkmale			Ergebnismerkmale
	Familie	Bildungserfahrungen	Akteure	Kernprozessse	Ressourcen	
Themen bezogen auf die Naturwissenschaften			Lehren und Lernen			Interesse, Einstellungen und Motivation bezogen auf die Naturwissenschaften
		Erfahrungen mit den Naturwissenschaften außerhalb der Schule	Qualifikation und professionelles Handeln der Lehrkräfte	Unterrichtspraktiken in den Naturwissenschaften	Lernzeit und Curriculum	
			Schulebene			
					Lernbedingungen für die Naturwissenschaften auf Schulebene	
Fachunabhängige Themen	Sozioökonomischer Hintergrund	Frühkindliche Bildung	Elterliche Einbindung	Schulklima	Personelle und materielle Ressourcen	Allgemeine Einstellungen, Lernmotivation und Selbstwahrnehmung
	Vielfalt und Zuwanderung		Schulleitung und Schulmanagement			Dispositionen zum kollaborativen Problemlösen
			Steuerung des Schulsystems			
			Entscheidungshoheit im Schulsystem	Leistungsmessung, Evaluation und Rechenschaftslegung	Schulwahl	Bildungsaspirationen

Abbildung 13.1: Die modulare Struktur der PISA-2015-Kontextfragebögen. In Anlehnung an Abb. 6.2 in OECD (2016), Übersetzung und Überarbeitung durch die Autoren

Fragebögen in PISA 2015 demnach an frühere PISA-Erhebungsrunden an, greifen aber auch neue Entwicklungen in der Unterrichtsforschung und den Fachdidaktiken auf.

Die Vertreterinnen und Vertreter der Teilnehmerstaaten im PISA Governing Board haben diesem Bereich höchste Priorität beigemessen, obwohl sie sich der Tatsache bewusst waren, dass PISA selbst über die Wirksamkeit des Unterrichts und die Entstehung von Einstellungen kaum etwas aussagen kann. Denn es handelt sich um eine reine Querschnittserhebung und die befragten Fünfzehnjährigen stammen innerhalb ein und derselben Schule aus unterschiedlichen Jahrgängen, Klassen und Kursen. PISA versteht sich nicht als Ersatz für fachbezogene Lehr-Lernforschung, sondern verwendet bewährte Befragungsskalen der empirischen Bildungsforschung, um repräsentative Aussagen zum Stand der naturwissenschaftlichen Bildung zu machen. Dieselbe Logik liegt auch den Fragen zugrunde, die sich mit fachunabhängigen Themen beschäftigen: Sie wurden auf der Basis der Pädagogischen Psychologie und der Schulforschung entwickelt beziehungsweise ausgewählt, um repräsentative Aussagen zu Strukturen und pädagogischen Prozes-

© Waxmann Verlag GmbH

sen, zu Qualität, Effektivität und Chancengerechtigkeit in Schulsystemen formulieren zu können.

Ein Markenzeichen von PISA bleibt auch in der Erhebung 2015 die differenzierte Erfassung von *individuellen Ausgangsmerkmalen der Schülerinnen und Schüler* (familiärer Status im Sinne von sozialem und kulturellem Kapital sowie ökonomischem Wohlstand), *Zuwanderung* (Migration der Schüler bzw. ihrer Eltern, Sprachnutzung) und *Bildungswegen* (z. B. Besuch des Kindergartens, verzögerte oder verkürzte Schullaufbahn). Die in früheren Jahren genutzten Indikatoren haben sich bewährt und werden weiterhin genutzt, auch weil verschiedene innovative Erhebungsmethoden, die in der Felderprobung 2014 getestet wurden, keine Verbesserung der Datenqualität bei der Messung des sozialen Hintergrunds erbrachten (Watermann, Maaz, Bayer & Roczen, 2016). Ein besonderes Anliegen war es, den *Umgang mit Vielfalt und Heterogenität* an den Schulen zu erfassen, beispielsweise Diskriminierungserfahrungen und die Achtung gegenüber vielfältigen kulturellen Identitäten.

In Bezug auf die Qualität schulischer Prozesse bildet PISA 2015 praktisch alle wesentlichen Bereiche ab, die aus der internationalen Schulforschung bekannt sind und auch in den Qualitätsrahmen der deutschen Bundesländer vorkommen. Neben Themen der *Lehrerausbildung, -weiterbildung und -kooperation*, zu denen neben den Naturwissenschaftslehrkräften auch Lehrpersonen anderer Fächer befragt wurden, gehören hierzu *Schulleitungshandeln, Einbeziehung der Eltern, geteilte Werte und Normen* und *Umgangsformen* (Schulklima) sowie die *Nutzung von Evaluierungen* für die Weiterentwicklung von Schule und Unterricht (vgl. etwa Bryk, Sebring, Allensworth, Easton & Luppescu, 2010; Klieme, 2013). Zudem erfasst auch PISA 2015 mögliche Einschränkungen wegen mangelnder Ressourcen. Dies ist ein Bereich, der sich in den wohlhabenden OECD-Staaten praktisch nicht auf die Leistungsfähigkeit von Schulen auswirkt, wohl aber in Entwicklungsländern.

Für die bildungspolitische Diskussion besonders relevant sind die Befragungsmodule, die sich mit der *Struktur und Steuerung* von Bildungssystemen befassen. Bereits seit 15 Jahren kann man mithilfe von PISA beobachten, wie sich Schularten, Schulwahl und Zuweisung zu weiterführenden Schulen entwickeln und wer für welche Entscheidungen (z. B. über Zuweisung von Ressourcen oder Lehrpläne) zuständig ist, wie weit also die Autonomie der Schulen bzw. der einzelnen Lehrkräfte reicht. Mit hoher Priorität widmet sich PISA 2015 nun auch der Frage, wie in den Schulsystemen mit Tests und Evaluierungen umgegangen wird. Hierzu wurden Schulleitungen, Lehrkräfte und Lernende befragt. Damit kann PISA belastbare Daten liefern, um Thesen etwa über eine wachsende Testbelastung zu prüfen, wie sie in der PISA-kritischen Öffentlichkeit formuliert werden.

Schließlich beschäftigten sich insgesamt drei Module mit *fachunabhängigen Ergebnissen von Bildung und Erziehung*, etwa allgemeiner Leistungsmotivation, körperlicher Aktivität sowie Lebenszufriedenheit und Wohlbefinden in der Schule. Damit wird PISA einem Bildungsverständnis gerecht, das nicht nur intellektuelle Leistung und fachliches Lernen berücksichtigt, sondern breitere Aspekte der psychosozialen Entwicklung.

© Waxmann Verlag GmbH

Soweit verfügbar sieht die Rahmenkonzeption für die Fragebögen in PISA 2015 vor, Fragen zu verwenden, die bereits in früheren PISA-Erhebungen verwendet wurden. Damit soll sichergestellt werden, dass Kernindikatoren über die Zeit hinweg erfasst und Trends analysiert werden können. Hierzu gehören beispielsweise Veränderungen beim Schulklima, bei der Ausstattung der Schulen und bei der Nutzung von Tests.

13.2 Entwicklung der Fragebögen für PISA 2015

Während in den vorherigen Erhebungen die Fragebogenentwicklung zeitlich *nach* der Fertigstellung der Rahmenkonzeption erfolgte, wurden beide Prozesse für PISA 2015 integriert: Internationale Expertengruppen bearbeiteten spezifische Inhalte der Fragebögen, dokumentierten die dafür relevanten theoretischen Grundlagen und entwickelten darauf aufbauend konkrete Vorschläge für die Inhalte der Fragebögen. Die Koordination lag auch hierfür beim DIPF.

Der Fokus der Entwicklung lag sowohl auf inhaltlichen als auch auf methodischen Innovationen. Inhaltlich musste vor allem der Bezug der Fragebögen und der Lernkontextbedingungen zu den Naturwissenschaften, der Hauptdomäne in PISA 2015, ausgebaut werden. Zudem sollten die Fragebögen in PISA 2015 auch aus methodischer Sicht die Konstrukte noch genauer abbilden können. Unter anderem wurden dafür neue Frage- und Antwortformate implementiert, die dabei helfen sollen, internationale Unterschiede bei der Interpretation von Fragen und Einschätzungsdimensionen auszugleichen. Hierzu zählen unter anderem sogenannte Ankervignetten. Der folgende Abschnitt stellt die methodischen Maßnahmen vor, die ergriffen wurden, um die Qualität der über die Fragebögen erfassten Daten zu verbessern.

Qualitätssicherung bei der Fragebogenentwicklung

PISA muss als international vergleichende Studie sicherstellen, dass alle Testinstrumente und Fragebögen über die teilnehmenden Staaten hinweg miteinander vergleichbar sind. Das betrifft zum einen Sprachfeinheiten, also typische Übersetzungsproblematiken, die in der einen, nicht aber unbedingt in der anderen Sprache auftreten können. Beispielsweise müssen Antwortkategorien, die unterschiedliche Grade von Zustimmung erfassen, sehr sorgfältig übersetzt werden. Sprachliche Feinheiten können ausschlaggebend dafür sein, dass sich eine Schülerin oder ein Schüler eher für die Kategorie der stärksten Zustimmung („stimme völlig zu“) oder für eine etwas schwächer formulierte Kategorie („stimme eher zu“) entscheidet.

Gleichzeitig müssen inhaltliche Besonderheiten berücksichtigt werden, die nur für einzelne oder einen Teil der an PISA teilnehmenden Staaten relevant sind. Ein Beispiel dafür ist die Sonderstellung des Referendariats in der Ausbildung von Lehrkräften in Deutschland. Fragen, die sich auf die Ausbildungsphase der befragten Lehrkräfte beziehen, müssen daher für die deutschen Lehrerfragebögen entsprechend angepasst wer-

© Waxmann Verlag GmbH

den. Es sollte klar festgelegt werden, ob die Befragten das Referendariat etwa für die Einschätzung der Schwerpunkte in ihrer Ausbildung miteinbeziehen oder ob sich die Antworten nur auf das Studium beziehen sollen. Entsprechend sorgsam muss die Übersetzung und Anpassung der Fragebögen erfolgen, um sowohl nationale Besonderheiten abzubilden als auch die staatenübergreifende Vergleichbarkeit sicherzustellen. Die Sicherung der Qualität der Fragebögen und damit auch der Informationen, die auf ihrer Basis gewonnen werden, erfolgt in PISA in mehreren Schritten: Rückmeldungen der Teilnehmerstaaten, kognitive Interviews, Übersetzungsproben sowie staatenspezifische Anpassungen.

Zunächst begutachteten die nationalen Projektmanager die ersten Fragebogenentwürfe (vgl. Kapitel 1). Kriterien für die Begutachtung waren unter anderem die Relevanz der Fragen für das eigene Bildungssystem, ethische oder datenschutzrechtliche Bedenken oder zu erwartende Probleme bei der Übersetzung der Fragen. Ein Großteil der in PISA 2015 eingesetzten Fragen und Konstrukte war bereits in früheren PISA-Erhebungsrunden verwendet worden und diente als breit angelegte Grundlage dafür, die Trends als Entwicklungen über die Zeit hinweg zu beschreiben.

Nach der Begutachtung der geplanten Fragebögen erfolgte eine Erprobung der für PISA 2015 neu entwickelten Fragen. Dazu beantworteten Schülerinnen und Schüler sowie Lehrkräfte in verschiedenen Ländern die Fragebögen im Entwurfsstadium (sogenannte *Cognitive Labs* oder kognitive Interviews). Dabei wurde zunächst erfasst, wie lange die Beantwortung der geplanten Fragen tatsächlich dauerte. Insbesondere wurde im Anschluss an die Bearbeitung in einer vertiefenden Diskussionsrunde danach gefragt, ob die Intention der jeweiligen Frage klar erkannt wurde, ob sie verständlich formuliert war und welche Aspekte darüber hinaus noch berücksichtigt werden sollten. Diese Erprobung der Instrumente fand auf Englisch, Französisch und Spanisch statt.

Anschließend wurden die sogenannten Quellversionen (vgl. Kapitel 1) der Fragebögen für die Übersetzung in die jeweilige Testsprache der Teilnehmerstaaten zur Verfügung gestellt. Um Probleme bei der Übersetzung in die mehr als 100 Zielsprachen vorwegzunehmen und zu reduzieren, wurden die neu entwickelten Materialien vorab zur Probe in je eine Sprache verschiedener Sprachgruppen übersetzt. Dazu gehörten eine asiatische, eine slawische, eine indogermanische, eine romanische sowie eine arabische Sprache. Eventuell auftretende Schwierigkeiten oder Unklarheiten bei der Übersetzung wurden an die Fragebogenentwickler zurückgemeldet. Häufig fielen an dieser Stelle Sprachidiome auf, die sich zum Beispiel bildlicher Sprache bedienen und daher eventuell sowohl über die Kulturen als auch über die Sprachen hinweg schwierig zu übersetzen sind.

Teilweise müssen für ein Bildungssystem auch spezifische nationale Begriffe gewählt werden, die sich in keinem anderen Staat wiederfinden. Beispiele hierfür sind Klassenstufen, Schularten oder Ausbildungsgänge. So enthält die Frage nach der Schulart, welche die deutschen Schülerinnen und Schüler besuchen, in den deutschen Fragebögen 16 Antwortoptionen, welche sich auf 33 verschiedene Schularten beziehen, um die stark differenzierten Bildungssysteme in Deutschland adäquat abzubilden.

© Waxmann Verlag GmbH

Fragebogen für Schülerinnen und Schüler (computerbasiert)	
(*n* = 6504 in Deutschland)	
(Bearbeitungszeit ca. 30 Minuten)	
Optional: Educational Career Questionnaire (nur auszugsweise in Deutschland)	
Optional: ICT Familiarity Questionnaire (nur auszugsweise in Deutschland)	
Fragebogen für die Schulleitung (internetbasiert)	
(*n* = 253 in Deutschland)	
(Bearbeitungszeit ca. 45 Minuten)	
Fragebogen für Lehrpersonen (internetbasiert)	
(*n* = 1862 in Naturwissenschaften, 3313 in anderen Fächern in Deutschland)	
(10 Lehrerinnen und Lehrer der Naturwissenschaften sowie 15 Lehrerinnen und Lehrer anderer Fächer pro Schule)	
(Bearbeitungszeit ca. 30 Minuten)	
Allgemeine Hintergrundfaktoren (5 Minuten)	
Fragebogen für die Lehrerinnen und Lehrer der Naturwissenschaften (25 Minuten)	Fragebogen für die Lehrerinnen und Lehrer anderer Fächer (25 Minuten)
Fragebogen für Eltern (papierbasiert)	
(*n* = 3385 in Deutschland)	
(Bearbeitungszeit ca. 30 Minuten)	

Abbildung 13.2: Internationales Erhebungsdesign der Fragebögen in PISA 2015

Der Umfang und die Erhebungszeit der Fragebogeninstrumente variieren je nach Adressatengruppe. In der für die Schülerinnen und Schüler vorgesehenen Zeit von 30 Minuten wurden etwa 70 Fragen beantwortetet. Der Fragebogen für die Eltern umfasst insgesamt 50 Fragen, was ebenfalls einer maximalen Erhebungszeit von 30 Minuten entspricht. Für den Schulfragebogen sind circa 45 Minuten Bearbeitungszeit vorgesehen. Da hier viele Fragen nach statistischen Informationen zur Anzahl der Ressourcen einer Schule gestellt werden, ist dieser Fragebogen sehr umfangreich und besteht aus 240 Einzelitems. Die Lehrkräfte schließlich benötigten für ihren Fragebogen circa 30 Minuten; insgesamt wurden über 110 Angaben jeweils von Naturwissenschaftslehrkräften und sonstigen Lehrkräften erfragt. Das Erhebungsdesign ist in Abbildung 13.2 dargestellt.

© Waxmann Verlag GmbH

13.3 Skalierung der Fragebögen: Vorbereitung der Datenauswertung

Die Fragebögen in PISA sind typischerweise so aufgebaut, dass eine Frage zu einem Thema mehrere dazugehörige Items (Teilfragen, Teilaussagen) umfasst. Dabei sind diese Items in der Regel so gewählt und formuliert, dass sie gemeinsam betrachtet einen Indikator für ein bildungsrelevantes Thema bilden. Eine lange Tradition in PISA hat beispielsweise die Skala *School Belonging*, also das individuelle Gefühl der Zugehörigkeit zur eigenen Schule (vgl. auch Kapitel 5). Anhand von sechs Items geben die Schülerinnen und Schüler Auskunft zu verschiedenen Aspekten des *School Belonging*, etwa ob sie an ihrer Schule leicht neue Freunde finden oder ob sie sie sich unbehaglich und fehl am Platz fühlen. Einzeln betrachtet liefern solche Items Informationen zu bestimmten Ausschnitten von Indikatoren, die jedoch nicht für sich genommen interpretiert werden sollten. Dies liegt daran, dass die Zusammenstellung von Fragen zu einer Skala vor dem Einsatz in Studien wie PISA umfangreiche Erprobungen durchlaufen hat, um einschätzen zu können, welche Aspekte zusammengefasst das beste Abbild des betreffenden Indikators ergeben und wie sich die dazugehörigen Einstellungen in der befragten Personengruppe typischerweise verteilen (vgl. auch Kapitel 12).

Statistisch gibt es verschiedene Möglichkeiten, solche Skalen zu bilden. PISA 2015 verwendet hierzu drei verschiedene Arten: erstens sogenannte einfache Indikatoren, die durch eine arithmetische Transformation zusammengefasst werden, zweitens abgeleitete Variablen, die auf einer Skalierung gemäß der Item Response Theory beruhen (vgl. Kapitel 12), sowie drittens zusammengesetzte Indikatoren aus mehreren Fragebogenskalen. Ein Beispiel für einen solchen zusammengesetzten Indikator ist der sogenannte ESCS-Index (Index of Economic, Social and Cultural Status), der in Kapitel 10 in diesem Band detailliert beschrieben wird. Der internationale technische Bericht der OECD stellt detailliert dar, welche Items aus den Fragebögen jeweils in einer Skala zusammengefasst werden.

Für PISA 2015 wurde in Anlehnung an die Skalierung der Testaufgaben ein gegenüber früheren Erhebungsrunden neues Skalierungsmodell gewählt (vgl. Kapitel 12). Während bei den PISA-Tests für Naturwissenschaften, Mathematik und Lesen im Rahmen der ersten Erfassung als Schwerpunkt eine Festlegung des Mittelwerts auf 500 und der Standardabweichung auf 100 erfolgt, wird der OECD-Mittelwert bei Fragebogenskalen immer auf $M = 0$ fixiert. Dadurch entsprechen negative Werte einer unterdurchschnittlichen, positive Werte einer überdurchschnittlichen und Werte nahe 0 einer durchschnittlichen Ausprägung der Skala. Die Standardabweichung dieser Werte ist auf $SD = 1$ festgelegt, sodass beispielsweise ein Messwert von -0.15 auf der Skala *School Belonging* bedeutet, dass eine Schülerin oder ein Schüler mit ihrem oder seinem Gefühl der Zugehörigkeit zur eigenen Schule um 0.15 Standardabweichungen unter dem OECD-Durchschnitt liegt.

© Waxmann Verlag GmbH

Kontextvariablen, die in PISA 2015 erhoben wurden und auch bereits in PISA 2006 erfragt worden waren, können im Trend analysiert werden. In diesen Fällen wurde der OECD-Durchschnitt für 2015 nicht erneut auf 0 sowie die Standardabweichung auf 1 gesetzt, sondern die Werte sind im Vergleich zu 2006 zu interpretieren. Das heißt, falls ein Teilnehmer etwa auf der Skala für *School Belonging* einen Wert von 0.12 erhält, bedeutet dies, dass diese Person um 0.12 Standardabweichungen über dem OECD-Durchschnitt für 2006 liegt. Wie der Wert außerdem im Vergleich zum OECD-Durchschnitt im Jahr 2015 liegt, wird dann durch die Differenz von 0.12 zum neuen OECD-Durchschnitt (zum Beispiel 0.04) ersichtlich.

13.4 Innovationen 2015: Computerbasierte Erhebung und ein internationaler Lehrerfragebogen

Computerbasierte Erhebung

Mit PISA 2015 wurde bereits die sechste Erhebungsrunde und zugleich der zweite Zyklus der Studie (2009, 2012 und 2015) abgeschlossen. Im Rahmen dieser sechsten Erhebung wurde eine Reihe von Neuerungen eingeführt (vgl. Kapitel 1 und 12). Die markanteste ist, dass (mit Ausnahme des optionalen Fragebogens für die Eltern) alle Erhebungsinstrumente ausschließlich computerbasiert dargeboten wurden. Bereits in früheren Erhebungsrunden bestand die Möglichkeit Zusatztests computerbasiert auf nationaler Ebene durchzuführen oder den Schulfragebogen online anzubieten. In PISA 2015 aber wurde die technologiegestützte Datenerhebung zur Regel. Man spricht trotzdem weiterhin von Fragebögen. Lediglich einige wenige OECD-Partnerstaaten, die eine vollständig computerbasierte Testung nicht ermöglichen konnten, führten PISA 2015 ohne Einsatz neu entwickelter Aufgaben und Fragen papierbasiert durch.

Während mögliche Auswirkungen des neuen Erhebungsmodus bei den Kompetenztests für die Naturwissenschaften, Mathematik und Lesen durch eine sogenannte *Modus-effekt-Studie* im Rahmen des internationalen Feldtests geprüft wurden (vgl. Kapitel 12), konnte sich der Vorschlag, eine entsprechende Studie auch für die Fragebögen durchzuführen, bei der OECD nicht durchsetzen. Um eventuelle Effekte des Testmodus (Papier versus Computer) möglichst gering zu halten und damit die Interpretation von Trends als Entwicklungen von Indikatoren über mehrere PISA-Runden hinweg zu stützen, wurden die Kernindikatoren computergestützt in der Regel im selben Antwortformat wie ursprünglich auf Papier erfasst. Auch wurde der Feldtest dafür genutzt, einzelne technologiebasierte Innovationen im Fragebogen empirisch zu überprüfen. So wurden für eine Frage das traditionelle Antwortformat der offenen Texteingabe mit dem neuen Format eines Schiebereglers verglichen und die Bearbeitungszeiten analysiert.

Die computerbasierte Erhebung ermöglicht darüber hinaus noch weitere technische Neuerungen, etwa den Einsatz neuer Antwortformate wie Drop-down-Menüs oder

© Waxmann Verlag GmbH

automatisierte Kontrollen während der Beantwortung – beispielsweise durch Warnhinweise an die Schülerinnen und Schüler, wenn sie einen unerwartet hohen oder niedrigen Wert für die Anzahl der wöchentlichen Unterrichtsstunden in naturwissenschaftlichen Fächern angegeben hatten. Auch die gesteuerte Vorgabe von Fragen an bestimmte Zielgruppen, beispielsweise nur an Schülerinnen und Schüler mit Zuwanderungshintergrund, ist ein innovatives Element, das eine genaue Passung von Fragebögen an bestimmte Teilgruppen der PISA-Kohorte erlaubt. Außerdem reduziert die elektronische Erhebung den Arbeitsaufwand bei der Datenverarbeitung und -speicherung sowie bei der Datenbereinigung, indem die Antworten der Schülerinnen und Schüler bereits in elektronischem Format vorliegen und nicht mehr von Hand eingegeben werden müssen.

Datenschutz

Aus der computerbasierten Erhebung in PISA 2015 ergeben sich neue und erweiterte Ansprüche an den Schutz der erhobenen Daten. In bisherigen Erhebungsrunden lagen diese vor ihrer Digitalisierung ausschließlich in Papierform vor, konnten und mussten also physisch vernichtet werden. In PISA 2015 bestanden die Daten bereits bei der Eingabe durch die Befragten in digitaler Form und bedurften somit auch neuer Formen des Schutzes der Vertraulichkeit (vgl. auch Kapitel 1 und 12).

Entsprechend ist von Interesse, welchen Weg die erfassten, pseudonymisierten Daten nach der Bearbeitung durch die Teilnehmerinnen und Teilnehmer an PISA 2015 nehmen. Hier müssen drei Formen unterschieden werden: Übertragung von Papier in digitales Format, Verarbeitung über einen USB-Stick sowie Online-Verarbeitung.

Erstens wurde der Elternfragebogen wie in bisherigen Erhebungsrunden papierbasiert administriert und verarbeitet. Er wird den Schülerinnen und Schülern vor dem PISA-Testtag für ihre Eltern mit der Bitte mitgegeben, diesen ausfüllen zu lassen und ihn entweder am PISA-Testtag ausgefüllt mitzubringen oder ihn innerhalb einer Woche an das IEA Data Processing and Research Center zurückzusenden. Die Namen der Eltern oder ihres Kindes sind zu keinem Zeitpunkt bekannt, werden nicht erfragt und die abgegebenen Elternfragebögen werden automatisiert eingelesen und weiterverarbeitet. Dieses Verfahren hat sich gegenüber früheren PISA-Erhebungsrunden nicht verändert.

Zweitens wird der Fragebogen für Schülerinnen und Schüler direkt im Anschluss an den Kompetenztest über denselben USB-Stick administriert wie die Daten des Kompetenztests. Die eingegebenen Daten werden auf diesem Daten-Stick gespeichert und lokal wieder ausgelesen.

Eine dritte Form des Datenflusses ist für die Schulleiter- und Lehrerfragebögen vorgesehen. Diese wurden *online* administriert, das heißt, Lehrkräfte und Schulleitungen erhielten eine Internetadresse, die zum Onlinefragebogen führte. Über zufällig generierte, pseudonymisierte Kennworte bekamen sie eine gesicherte Verbindung, um den Fragebogen zu bearbeiten. Die eingegebenen Antworten wurden dann auf diesem Server gespeichert, bis sie mit Abschluss des Test- und Befragungszeitraums von der nationalen

© Waxmann Verlag GmbH

Studienleitung erneut über eine sichere Verbindung heruntergeladen und in die Datenbank mit den Daten der Schülerinnen und Schüler integriert wurden.

Lehrerfragebogen

Während auf nationaler Ebene in Deutschland bereits seit PISA 2003 ein Lehrerfragebogen eingesetzt wurde, führte die OECD diese Neuerung erst für die Erhebungsrunde 2015 als optionale Komponente ein. 18 Staaten nahmen international an der Befragung der Lehrerinnen und Lehrer teil (vgl. Kapitel 5). Jeweils 25 Lehrpersonen pro Schule wurden zufallsbasiert in die Stichprobe gezogen (vgl. Kapitel 12), die Teilnahme an der Befragung war freiwillig. Der Fragebogen wurde in zwei Versionen eingesetzt: (a) als Fragebogen für Lehrkräfte, die ein naturwissenschaftliches Fach bei der PISA-Kohorte der Fünfzehnjährigen unterrichten, oder (b) als Fragebogen für Lehrkräfte, die Unterricht in anderen, nicht naturwissenschaftlichen Fächern in dieser Altersgruppe erteilen.

Themen des Lehrerfragebogens waren unter anderem allgemeine Informationen zur Stellensituation, der eigenen Berufswahl und den aktuellen Arbeitsbedingungen an ihrer Schule. Besonders detailliert erfasst wurden die *Eingangsvoraussetzungen* wie soziodemographische Merkmale und die Aus-, Fort- und Weiterbildung der Lehrkräfte, *Prozessmerkmale* (z. B. Kooperation im Kollegium und Handeln im Unterricht) sowie *Ergebnismerkmale* wie die Zufriedenheit der Lehrkräfte allgemein mit ihrem Beruf und speziell mit dem Unterrichten an ihrer Schule.

13.5 Ausblick

Die Kontextfragebögen haben sich in PISA über die bisherigen sechs Erhebungsrunden hinweg zu einem wesentlichen Bestandteil der Studie entwickelt. Basierend auf der theoretischen Rahmenkonzeption ermöglichen sie es, Bedingungen, Prozesse und Ergebnismerkmale auf verschiedenen Ebenen des Bildungssystems und bei verschiedenen Akteuren zu erfassen. Viele der dadurch gewonnenen Indikatoren können über die Zeit verfolgt werden, sodass sich insbesondere Veränderungen oder Entwicklungstrends gut beschreiben lassen. Darüber hinaus bieten die erhobenen Informationen eine Grundlage für die Analyse komplexer Modelle von Lehr-Lernzusammenhängen auch in Bezug auf die Kompetenzen, die Schülerinnen und Schülern erreicht haben. Dabei werden die Fragebogeninstrumente analog zu den in PISA eingesetzten Testinstrumenten kontinuierlich weiterentwickelt, sowohl hinsichtlich ihrer Inhalte, die zunehmend über internationale Optionen erweitert werden, aber auch hinsichtlich der Auswertungs- und Skalierungsmethoden. Aktuelle Publikationen beschäftigen sich beispielsweise auch mit der Frage, inwieweit die erhobenen Skalen über alle Teilnehmerstaaten hinweg vergleichbar sind (van de Vijver & He, 2016). Somit bietet PISA mittlerweile eine der umfassendsten internationalen Datenquellen zu Kontextfaktoren des Lernens.

© Waxmann Verlag GmbH

Bei der Verwendung dieser Datenquelle ist jedoch zu beachten, welche Schlussfolgerungen aus einer beschreibenden Querschnittstudie wie PISA gezogen werden können und wo solche Schlussfolgerungen an Grenzen stoßen. Aussagen über Veränderungen von Kontextfaktoren auf Systemebene, beispielsweise bildungspolitische oder bildungsadministrative Entscheidungen und Veränderungen, können nicht mit den PISA-Leistungen der Schülerinnen und Schülern kausal verknüpft werden (vgl. Kapitel 1 und 12). Folglich sind in Zusammenhang mit den PISA-Daten Ursache-Wirkungs-Aussagen in nahezu allen Fällen unzulässig. Dies liegt zum einen daran, dass PISA als Querschnittstudie (*yield study*) keinerlei Daten über die Leistungsentwicklung oder das Vorwissen derselben Schülerinnen und Schüler sammelt. Darüber hinaus spielen sich zahlreiche bildungspolitische Reformen auf unterschiedlichen Ebenen des Bildungssystems ab, die sich mehr oder weniger direkt auf die Individualebene der Schülerinnen und Schüler auswirken können.

Bei den Analysen der Lehr-Lernkontexte sind immer deren Komplexität und das Zusammenspiel unterschiedlicher Einflussfaktoren über die Zeit zu berücksichtigen. Die Wirkrichtung oder auch zwischengeschaltete indirekte Effekte sind daher nicht aus den Daten ableitbar. Vielmehr liegt die Stärke der PISA-Daten darin, als wiederholte Momentaufnahme ein sehr umfassendes und repräsentatives Bild über die Umstände und Begebenheiten der untersuchten Bildungssysteme zu zeichnen. Außerdem können auch systematisch miteinander auftretende Merkmale identifiziert werden. Anschlussuntersuchungen erlauben dann einen auf diese Weise gefundenen Zusammenhang auf zugrunde liegende Wirkmechanismen zu prüfen. Ein Beispiel hierfür ist die in Deutschland mit PISA 2012 verknüpfte Längsschnittstudie PISA Plus. Ein weiteres ist die geplante internationale TALIS-Videostudie, die das in Deutschland mehrfach erprobte Prinzip, Lernentwicklungen über eine vorgegebene Unterrichtseinheit hinweg zu betrachten, mit aus PISA übernommenen Befragungen verbindet und so zur Validierung der Fragebögen auch im internationalen Vergleich beitragen wird.

Literatur

Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q. & Luppescu, S. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago: University of Chicago Press.

Jude, N. (im Druck). The assessment of learning contexts in PISA. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing contexts of learning world-wide*. Berlin: Springer.

Klieme, E. (2013). The role of large-scale assessments in research on educational effectiveness and school development. In M. v. 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.

Klieme, E., Backhoff, E., Blum, W., Buckley, J., Hong, Y., Kaplan, D., Levin, H., Scheerens, J., Schmidt, W., van de Vijver, A. J. R & Vieluf, S. (2013). PISA 2012 context questionnaires framework. In OECD (Hrsg.), *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy* (S. 167–258). Paris: OECD.

© Waxmann Verlag GmbH

Klieme, E. & Kuger, S. (2016). PISA 2015 context questionnaires framework. In OECD (Hrsg.), *PISA 2015 assessment and analytical framework: Science, reading, mathematic and financial literacy* (S. 101–127). Paris: OECD. http://dx.doi.org/10.1787/9789264255425-7-en

Klieme, E. & Vieluf, S. (2013). Schulische Bildung im internationalen Vergleich. Ein Rahmenmodell für Kontextanalysen in PISA. *Zeitschrift für Pädagogik*, *Beiheft 59*, 229–246.

Kuger, S., Klieme, E., Jude, N. & Kaplan, D. (Hrsg.). (im Druck). *Assessing contexts of learning world-wide.* Berlin: Springer.

Müller, K., Prenzel, M., Seidel, T., Schiepe-Tiska, A. & Kjærnsli, M. (im Druck). Science teaching and learning in schools: Theoretical and empirical foundations for investigating classroom-level processes. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.). *Assessing contexts of learning: An international perspective.* Berlin: Springer.

OECD. (2009). *PISA 2009 assessment framework – key competencies in reading, mathematics and science.* Paris: OECD.

OECD. (2010). *PISA 2009 results: What makes a school successful* (Volume IV). Paris: OECD.

OECD. (2016). *PISA 2015 assessment and analytical framework: Science, reading, mathematic and financial literacy.* Paris: OECD. http://dx.doi.org/10.1787/9789264255425-en

Schiepe-Tiska, A., Roczen, N., Müller, K., Prenzel, M. & Osborne, J. (im Druck). 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*. Berlin: Springer.

van de Vijver & He (2016). Bias assessment and prevention in non-cognitive outcome measures in PISA questionnaires. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing contexts of learning: An international perspective*. Berlin: Springer.

Watermann, R., Maaz, K., Bayer, S. & Roczen, N. (2016). Social background. In S. Kuger, E. Klieme, N. Jude & D. Kaplan (Hrsg.), *Assessing contexts of learning: An international perspective*. Berlin: Springer.

© Waxmann Verlag GmbH

© Waxmann Verlag GmbH

14 Glossar zentraler Begriffe zu den PISA-Ergebnissen

Christine Sälzer, Kristina Reiss & Jörg-Henrik Heine

Nach mittlerweile sechs abgeschlossenen Erhebungsrunden haben sich in der Berichterstattung zu PISA einige zentrale Fachbegriffe und Arten der Ergebnisdarstellung etabliert. Diese werden in den folgenden Abschnitten vorgestellt.

Antwortformate (offen, geschlossen, teiloffen). In PISA kommen üblicherweise Testaufgaben mit drei unterschiedlichen Antwortformaten zum Einsatz: Mehrfachwahl (*Multiple Choice*), Kurzantwort (*Short Answer*) oder offenes Antwortformat (*Extended Response*). Bei ungefähr 50 Prozent der Aufgaben müssen die Schülerinnen und Schüler aus mehreren Antwortmöglichkeiten eine oder mehrere auswählen, die ihrer Meinung nach korrekt sind. Die andere Hälfte der Aufgaben verlangt eine selbst formulierte, mehr oder weniger ausführliche Antwort. Hier können die Jugendlichen entweder Kurzantworten beziehungsweise eine Zahl eintragen oder aber Antworten, die aus mehreren Sätzen, einer Skizze oder Zeichnung bestehen (offenes Antwortformat). Die Auswertung der Multiple-Choice-Antworten erfolgt automatisiert und digital.

Booklet-Design. Im Rahmen groß angelegter Schulleistungsvergleichsstudien wie PISA kommen in der Regel deutlich mehr Aufgaben zum Einsatz als ein einzelner Schüler oder eine Schülerin realistisch bearbeiten kann. Deshalb wird jedem Studienteilnehmer jeweils nur eine Auswahl von Aufgaben vorgelegt. Dabei muss sichergestellt werden, dass einerseits jede vorhandene Aufgabe von genügend (und möglichst gleich vielen) Schülerinnen und Schülern bearbeitet wird und andererseits jeder Studienteilnehmer eine Auswahl von Aufgaben bekommt, die fair und nicht systematisch verzerrt ist (also beispielsweise nur sehr schwierige Aufgaben). Dies wird meist mit einem sogenannten Booklet-Design gelöst: Die Aufgaben einer Kompetenzdomäne (Naturwissenschaften, Mathematik, Lesen) werden in sogenannte Cluster gruppiert und diese Cluster wiederum in unterschiedlichen Booklets zusammengestellt (auch Testheft oder bei Testung am Computer *Testform* genannt). In PISA umfasst ein Booklet jeweils vier Cluster, wobei die Schülerinnen und Schüler für die Bearbeitung eines Clusters etwa 30 Minuten benötigen. Ein Cluster ist dabei immer sortenrein, d. h. es enthält ausschließlich Aufgaben aus einer Domäne.

Effektstärke. Die Effektstärke zeigt die praktische Relevanz statistisch signifikanter Ergebnisse aus sozialwissenschaftlichen Studien an. Sie ist ein Maß für Unterschiede, das die Streuung der Werte berücksichtigt. Grob gesprochen wird sie bei einer breiten Streu-

© Waxmann Verlag GmbH

ung der Messergebnisse eher kleiner und bei einer geringen Streuung eher größer. Speziell im Kontext von Bildungsvergleichsstudien ist die Umwandlung solcher Effektstärken in bestimmte Entwicklungszeiträume wie ein Schuljahr sehr anschaulich.

Grundgesamtheit. Die interessierende Grundgesamtheit in PISA sind Fünfzehnjährige, die in den teilnehmenden Staaten eine Schule besuchen. Von der in PISA getesteten Stichprobe wird entsprechend auf diese Grundgesamtheit geschlossen.

Indikator, manifester (auch: *manifeste Variable*). Bei manifesten Indikatoren oder Variablen handelt es sich im Gegensatz zu → *latenten Variablen* um Merkmale, die im Rahmen sozialwissenschaftlicher Untersuchungen direkt beobachtet oder erfasst werden können (z. B. das Alter einer Person). Manifeste Indikatoren werden im Rahmen sozialwissenschaftlicher Untersuchungen auch dazu eingesetzt, → *latente Variablen* zu operationalisieren und damit indirekt zu beobachten und zu messen.

Item. Als Item (im Rahmen von PISA auch: → *Teilaufgabe*) werden in der sozialwissenschaftlichen Literatur die einzelnen (kleinsten) Elemente eines Tests oder Fragebogens bezeichnet. Im Kontext von PISA ist damit auch eine Teilaufgabe gemeint. Dabei können bei PISA im Wesentlichen drei unterschiedliche Arten von Items in Bezug auf ihr → *Antwortformat* (offen, geschlossen, teiloffen) unterschieden werden. Für die Datenauswertung und → *Skalierung* im Rahmen der → *Item Response Theory* bei PISA werden die Antworten der Schülerinnen und Schüler zu einem Item anhand eines zwei- oder mehrstufigen Kategoriensystems kodiert (z. B. 0 = „falsch", 1 = „teilweise richtig" und 2 = „richtig").

Item Response Theory (IRT). Die *Item Response Theory* liefert den theoretischen und methodischen Rahmen für die Untersuchung des Zusammenhangs zwischen manifest beobachteten Variablen und den als dafür ursächlich angenommenen, dahinterliegenden latenten Merkmalen. Die manifesten Variablen (→ *Indikatoren*) wie etwa die Antworten von Jugendlichen auf PISA-Testaufgaben werden mit latenten Merkmalen (→ *latente Variable*) wie etwa der naturwissenschaftlichen Kompetenz der Schülerinnen und Schüler verbunden. Im Gegensatz zur → *Klassischen Testtheorie* (KTT) ermöglicht die IRT die explizite Prüfung der Gültigkeit des angenommenen (ursächlichen) Zusammenhangs zwischen den Indikatoren und der latenten Variable im Rahmen der Modelltestung (→ *Modellparameter*).

Klassische Testtheorie (KTT). Die KTT ist im Wesentlichen eine Theorie zum sogenannten Messfehler. Aufbauend auf einer Reihe von Axiomen (theoriebasierten Grundannahmen) besteht der Kerngedanke der KTT darin, dass sich die *gemessene Merkmalsausprägung X* aus zwei grundlegenden Komponenten zusammensetzt. Diese sind der Fehler *E* (engl. *Error*) sowie der theoretische wahre Wert der Merkmalsausprägung *T* (*True Score*). Im Gegensatz zur → *Item Response Theory* werden im Rahmen der KTT keinerlei (überprüfbare) Annahmen zum Antwortprozess (bei Tests und Fragebogen), oder allgemein, zum Entstehungsprozess der jeweiligen Messwerte getroffen.

© Waxmann Verlag GmbH

Kodierung. Ein Teil der Antworten, welche die Fünfzehnjährigen in ihre PISA-Testform eingetragen haben, kann automatisch und digitalisiert ausgewertet werden (*Multiple-Choice*-Antworten; vgl. → *Antwortformate*). Frei formulierte Antworten der Schülerinnen und Schüler müssen jedoch von intensiv geschulten und fachkundigen Kodiererinnen und Kodierern beurteilt werden. Unter Beachtung detaillierter Kodiervorschriften weisen diese Kodiererinnen und Kodierer den Schülerantworten einen Code zu, der die Antwort als „richtig", „teilweise richtig" oder „falsch" klassifiziert.

Kohorte. Eine Kohorte ist eine Gruppe von Personen, die eine in der Regel unveränderliche, ereignisbezogene Eigenschaft gemeinsam hat. Beispiele sind das Jahr der Geburt (Geburts- oder Alterskohorte) oder des Eintritts in das Erwerbsleben (Berufseintrittskohorte). In PISA wird die Kohorte fünfzehnjähriger Schülerinnen und Schüler untersucht, es handelt sich also um eine Alterskohorte mit dem zusätzlichen Merkmal des Schulbesuchs. Diese Kohorte definiert zugleich die → *Grundgesamtheit*, auf die anhand der Daten aus der PISA-Studie geschlossen wird.

Kompetenzstufen. Die in PISA berechneten Mittelwerte geben jeweils das durchschnittliche Niveau auf einer Kompetenzskala an, also beispielsweise der Skala der naturwissenschaftlichen Kompetenz. Diese Skalen werden in Kompetenzstufen unterteilt, die auf statistischer Ebene einem Wertebereich mit einem definierten Minimum und Maximum entsprechen. Auf der inhaltlich-fachdidaktischen Ebene ist eine entsprechend zugewiesene Beschreibung der Kompetenzen, über die Jugendliche typischerweise auf den einzelnen Kompetenzstufen verfügen, ein zweites grundlegendes Merkmal. Darin enthalten sind auch die definierten kognitiven Anforderungen, die von den Schülerinnen und Schülern gemeistert werden müssen, um die der entsprechenden Kompetenzstufe zugeordneten Aufgaben lösen zu können. Formal bedeutet das Folgendes: Fünfzehnjährige, die eine für eine Stufe mittlere Kompetenz aufweisen, können eine für diese Stufe mittelschwierige Aufgabe mit einer festgelegten Wahrscheinlichkeit von 62 Prozent lösen. Schülerinnen und Schüler auf dieser Kompetenzstufe werden Aufgaben, deren Schwierigkeit einer höheren Kompetenzstufe entspricht, mit einer deutlich geringeren Wahrscheinlichkeit lösen können. Für Aufgaben, die für eine niedrigere Kompetenzstufe typisch sind, gilt eine entsprechend höhere Lösungswahrscheinlichkeit.

Large-Scale-Assessment. Der Begriff *Large-Scale-Assessment* hat sich als Bezeichnung für groß angelegte Bildungsvergleichsstudien etabliert. Gemeint ist damit typischerweise eine Stichprobe von mehreren Tausend Teilnehmenden, deren Kompetenz und weitere Merkmale systematisch miteinander verglichen werden.

Mittelwerte und Streuungsmaße (Standardabweichungen). Die bekanntesten und am häufigsten berichteten Maßzahlen der PISA-Studien sind Mittelwerte. Mittelwerte geben an, welches Niveau in einem Kompetenztest durchschnittlich erreicht wurde; auf ihrer Basis können Rangreihen gebildet werden. Je nach Verteilung können die Leistungen der Schülerinnen und Schüler aber unterschiedlich um diesen Mittelwert streuen, also homogener oder heterogener sein. Eine solche Kennzahl für die Streuung von Leistungs-

© Waxmann Verlag GmbH

daten um ihren Mittelwert ist die Standardabweichung. Bei einem Mittelwert *M* (*Mean*) und einer Standardabweichung *SD* (*Standard Deviation*) liegen nach Definition etwas mehr als 68 Prozent aller gemessenen Werte in einem Intervall zwischen *M* + *SD* und *M* – *SD*. Ein Mittelwert von 100 und eine Standardabweichung von 20 bedeuten also, dass gut 68 Prozent aller gemessenen Werte zwischen 80 und 120 liegen. Dabei nimmt man eine sogenannte → *Normalverteilung* der Messwerte an.

Zur Erleichterung von Vergleichen zwischen den OECD-Staaten und den Partnerstaaten sowie für Vergleiche über mehrere PISA-Runden hinweg wurden die Mittelwerte und Standardabweichungen normiert. Konkret wurden diese Kennzahlen für die jeweils erste Erhebung einer Hauptdomäne (also Lesen 2000, Mathematik 2003, Naturwissenschaften 2006) statistisch so transformiert, dass für die Gruppe der OECD-Staaten der *Mittelwert* (*M*) bei 500 Punkten lag und die *Standardabweichung* (*SD, Standard Deviation*) bei 100 Punkten. Diese Normierung ist dann auch der Bezugspunkt für nachfolgende PISA-Erhebungen: Wenn sich die Leistungen der OECD-Staaten insgesamt verschlechtert haben, sinkt der OECD-Mittelwert unter 500 (was tatsächlich passiert ist). Änderungen der Standardabweichung zeigen hingegen eine Vergrößerung oder Verkleinerung der Leistungsstreuung an. Ein internationaler Mittelwert von *M* = 491 und eine Standardabweichung von *SD* = 98 in den Naturwissenschaften bedeuten also im Vergleich zu 2006 einen um 9 Punkte kleineren Mittelwert und eine minimal geringere Streuung der Leistungen.

Unter Annahme einer → *Normalverteilung* der Kompetenzen lagen die Punktzahlen (jeweils bei der ersten Erhebung einer Domäne) von etwa zwei Dritteln (genauer: 68.3 Prozent) der Schülerinnen und Schüler aus den OECD-Staaten im Bereich von ± 1 (einer) Standardabweichung um den Mittelwert, also im Bereich zwischen 400 und 600 Punkten. Dies bedeutet, dass die Kompetenzwerte von 15.9 Prozent der Jugendlichen mehr als eine Standardabweichung unter dem Durchschnitt lagen und entsprechend die Werte von weiteren 15.9 Prozent [(100 – 68.3) : 2] erheblich – um mindestens eine Standardabweichung – über dem Mittelwert. Innerhalb von zwei Standardabweichungen um den Mittelwert, also im Bereich von 300 bis 700 Punkten, liegen 95.4 Prozent der Kompetenzwerte. Damit befinden sich fast alle Schülerinnen und Schüler im Punktebereich von 300 bis 700 Punkten, lediglich jeweils 2.3 Prozent [(100 – 95.4) : 2] erreichen weniger als 300 oder mehr als 700 Punkte.

Modellparameter. Bei der Analyse von Beobachtungsdaten aus Tests oder Fragebögen im Rahmen der → *Item Response Theory* (IRT) werden Modelle für den Antwortprozess formuliert. In solchen Modellen werden Annahmen getroffen, wie die Antwort auf eine Frage mit Eigenschaften der befragten Personen zusammenhängen kann. Durch ein einfaches Modell werden zum Beispiel die Antwortdaten durch die → *Schwierigkeit* der einzelnen Aufgaben und die (unterschiedliche) Fähigkeit (→ *latente Variable*, Fähigkeitsdimension) der Personen erklärt. Das Modell enthält so in diesem Beispiel zwei Parameter (Fähigkeit der Personen und Schwierigkeit der → *Items*). Im Rahmen der Überprü-

© Waxmann Verlag GmbH

fung des Zusammenhangs zwischen beobachteten und latenten Variablen werden diese Modellparameter mittels geeigneter statistischer Verfahren bestimmt.

Normalverteilung. Auch Gauß-Verteilung genannt (nach dem Mathematiker Carl Friedrich Gauß). Eine Normalverteilung von Werten kann man sich entsprechend der Normalverteilungskurve in Abbildung 14.1 so vorstellen: Die meisten Werte liegen nahe am Durchschnitt (daher hat die Kurve dort ihr Maximum) und je weiter ein Wert im Randbereich liegt, desto seltener kommt er vor. Normalverteilungen zeigen sich in sozialwissenschaftlichen Befragungen oder Studien häufig dann, wenn die Stichproben sehr groß sind; dies ist auch in Large-Scale-Assessments wie PISA der Fall. Die von den Schülerinnen und Schülern erzielten Punkte auf den Kompetenzskalen folgen entsprechend auch einer Normalverteilung, für welche die oben beschriebene Standardabweichung und die prozentualen Anteile derjenigen, deren Punktzahl sich im Bereich von $M - SD$ bis $M + SD$ befindet, gelten.

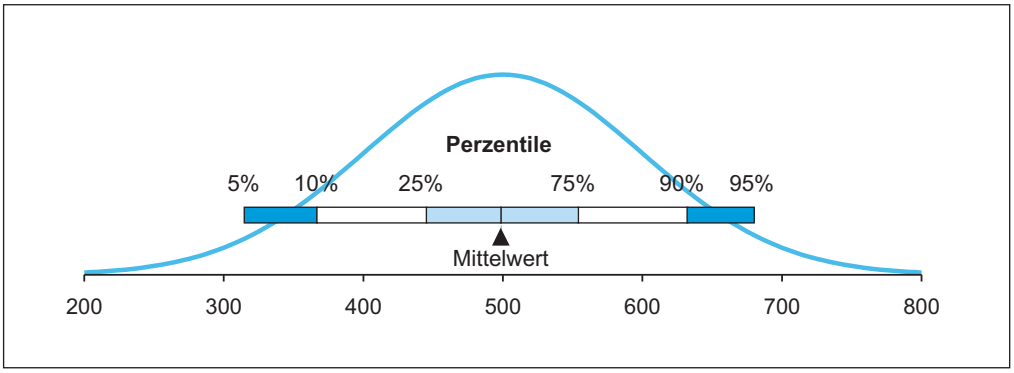


Abbildung 14.1: Perzentilband und Normalverteilungskurve

Oversampling. PISA zielt als Untersuchung auf die Grundgesamtheit der fünfzehnjährigen Schülerinnen und Schüler in den beteiligten Bildungssystemen ab. Teilweise verfolgen die PISA-Teilnehmerstaaten neben den in PISA erfassten Indikatoren noch weitere Themen, die für sie besonders interessant sind (etwa das Kompetenzniveau bestimmter Teilpopulationen wie Zuwanderer, Schülerinnen und Schüler mit besonderem Förderbedarf, Schülerinnen und Schüler in Klasse 9). Meist ist der Anteil dieser spezifischen Gruppen von Jugendlichen an der Gesamtpopulation jedoch so gering, dass man anhand der regulären Stichprobe keine repräsentativen Aussagen über sie machen kann. Um diese Teilpopulationen repräsentativ abbilden zu können, muss die Stichprobe deshalb entsprechend erweitert werden. Eine solche Stichprobenerweiterung heißt auch *Oversample* und bedeutet, dass gezielt Mitglieder der interessierenden Teilgruppe (z. B. Zuwanderer) in größerem Umfang für die Studie ausgewählt werden. Neben der üblichen Stichprobe wird also zusätzlich eine ergänzende Stichprobe von Teilnehmerinnen und Teilnehmern gezogen, die ein bestimmtes Merkmal wie Zuwanderungshintergrund oder sonderpädagogischen Förderbedarf haben. In Deutschland gibt es in PISA regelmäßig ein Oversample von Schülerinnen und Schülern, die eine 9. Klasse besuchen.

© Waxmann Verlag GmbH

Perzentile und Perzentilbänder. In vielen Abbildungen für den internationalen Vergleich in PISA, aber etwa auch für Vergleiche zwischen Schularten in Deutschland, werden Angaben über sogenannte *Perzentile* gemacht. Ein *Perzentilwert* trennt einen Bereich einer Verteilung vom restlichen Bereich; für diese Bereiche werden diejenigen Prozente der Stichprobe angegeben, die unterhalb dieses Wertes bzw. 100 minus dem Prozentsatz liegen. So wird beispielsweise am 95. Perzentil derjenige Kompetenzwert berichtet, ab dem die oberen 5 Prozent der Verteilung beginnen. Ein Perzentilwert ist nützlich, wenn man ihn in Verbindung mit einem Leistungswert interpretiert. So kann man etwa angeben, dass in Deutschland im Bereich der Lesekompetenz das 95. Perzentil bei 664 Punkten beginnt, in Schweden hingegen bei 655. Folglich erreichen die besten 5 Prozent in Deutschland höhere Werte als die besten 5 Prozent in Schweden. Grafisch werden Perzentile auch als *Perzentilbänder* dargestellt. Diese Bänder bilden die Breite einer Leistungsverteilung ab und damit die Distanz zwischen den leistungsschwächsten und den leistungsstärksten Schülerinnen und Schülern in den Teilnehmerstaaten. Abbildung 14.1 veranschaulicht ein solches Perzentilband und die Normalverteilungskurve (vgl. → *Normalverteilung*).

1-PL-Modell. Das 1-PL-Modell (*1-parametrisches logistisches Modell*) ist ein psychometrisches Modell im Kontext der → *Item Response Theory*. Es geht davon aus, dass sich die Kompetenz einer Person anhand ihrer Fähigkeit sowie der → *Schwierigkeit* der gestellten Testaufgaben eindimensional beschreiben lässt. Das bedeutet, dass das 1-PL-Modell die Fähigkeit der Person und die Schwierigkeit der bearbeiteten Testaufgaben auf einer gemeinsamen Dimension („Kompetenz") abbildet. Der eine Parameter, welcher bei diesem Modell frei geschätzt wird, ist die Position (die Schwierigkeit) eines → *Items* auf der Skala der latenten Eigenschaft (Kompetenz). Das 1-PL-Modell kann als ein Spezialfall des → *2-PL-Modells* angesehen werden.

2-PL-Modell. Das 2-PL-Modell (*2-parametrisches logistisches Modell)* ist etwas komplexer als das 1-PL-Modell und nutzt als zweiten Parameter neben der Position eines Items auf der Kompetenzskala die sogenannte Trennschärfe einer Aufgabe. Die Trennschärfe beschreibt die Beziehung zwischen der Antwort auf eine Testaufgabe und der interessierenden → *latenten Variablen* (z. B. Kompetenz). Sie drückt also das Potenzial einer Aufgabe aus, tatsächlich zwischen kompetenteren und weniger kompetenten Studienteilnehmern zu unterscheiden.

Plausible Values (PVs): *Plausible Values* (dt. plausible Werte) repräsentieren bei PISA die Messwerte für die erfassten Kompetenzdomänen (*Naturwissenschaften*, *Mathematik* und *Lesen*). Im Gegensatz zu einfachen Punktschätzern für jede Person in der Stichprobe (d. h. ein Zahlenwert für jede Person) wird das erreichte Kompetenzniveau für jede Person durch mehrere *Plausible Values* abgebildet. Die grundlegende Idee im Rahmen der Kompetenzschätzung über Modelle (→ *Modellparameter*) aus der → *Item Response Theory* ist, dass zuerst die Kennwerte einer (individuellen) Verteilung der Kompetenz ermittelt werden. Für diese (individuelle) Kompetenzverteilung jeder Person wird (approximativ) eine → *Normalverteilung* angenommen. Aus dieser Verteilung können

© Waxmann Verlag GmbH

nun theoretisch beliebig viele Werte gezogen werden. Bei PISA 2015 werden 10 dieser plausiblen Werte (daher deren Bezeichnung) gezogen. Der Vorteil des Verfahrens besteht im Wesentlichen darin, dass die jeweils auf der Ebene der einzelnen Personen bestehende Messungenauigkeit (→ *Reliabilität*) bei der Bestimmung der Kennwerte der → *Grundgesamtheit* angemessen berücksichtigt wird.

Population. Siehe → *Grundgesamtheit*.

Psychometrie. Ein Fachgebiet der Psychologie, das sich mit Theorie und Methodik des psychologischen Messens beschäftigt. Dazu gehören beispielsweise psychologische Testverfahren (Intelligenztests, Kompetenz- oder Persönlichkeitstests), aber auch Methoden zur Verarbeitung und Auswertung von Daten, wie sie unter anderem in PISA erhoben werden.

Querschnitt. Als Querschnitt werden wissenschaftliche Untersuchungen und Erhebungen bezeichnet, die einen einzigen Messzeitpunkt haben. Das heißt, die untersuchten Personen (bzw. Fälle, Zellen etc.) nehmen genau einmal an einer Datenerhebung teil. Im Gegensatz dazu beinhalten Längsschnittstudien mehrere Messzeitpunkte bei denselben Fällen und erheben mehrmals Daten über einen längeren Zeitraum hinweg. Auch PISA ist eine Querschnittstudie, da die altersbasierte Untersuchungskohorte (fünfzehnjährige Schülerinnen und Schüler, → *Kohorte*) jeweils nur einmal an PISA teilnimmt und in der nächsten Runde die dann fünfzehnjährigen Schülerinnen und Schüler die Grundgesamtheit bilden. Die Betrachtung mehrerer PISA-Erhebungsrunden im Vergleich erlaubt allerdings – wenn auch eingeschränkt – die Abbildung von Trends, was oftmals auch als Quasi-Längsschnitt bezeichnet wird; es können jedoch keine Entwicklungen bei den Schülerinnen und Schülern nachgezeichnet werden, sondern ausschließlich Veränderungen auf der Ebene der Bildungssysteme (dazu gehören Indikatoren wie das Alter bei der Einschulung oder die Quote von Klassenwiederholungen). Dies setzt voraus, dass die verwendeten Aufgaben entweder dieselben sind oder auf derselben Dimension abgebildet werden können (→ *Skalierung*).

Regression, latente. Mit *Regression* wird in der empirischen sozialwissenschaftlichen Forschung ein (statistisches) Verfahren bezeichnet, bei dem eine Variable (das sogenannte Kriterium) durch die Kenntnis mindestens einer weiteren Variablen (Prädiktoren) vorhergesagt wird. Das Kriterium geht also sozusagen auf andere Variablen zurück („Regression"). Kausale Interpretationen hinsichtlich der Beziehung zwischen dem Kriterium und den Prädiktoren sind allerdings nicht unbedingt zulässig. Mit einer *latenten Regression* werden als Erweiterung Verfahren bezeichnet, bei denen das Kriterium als → *latente Variable* über geeignete Modelle (vgl. → *Modellparameter*) beispielsweise im Rahmen der → *Item Response Theory* operationalisiert wird. Dabei wird das Kriterium gleichzeitig durch eine Reihe von Prädiktoren ‚vorhergesagt'. Anhand einer latenten Regression soll die Messgenauigkeit (→ *Reliabilität*) der → *latenten Variablen* durch die Prädiktoren verbessert, d. h. um Messfehler bereinigt, werden. Als zusätzliche Erweiterung können bei latenten Regressionsmodellen die Zusammenhänge zwischen dem

© Waxmann Verlag GmbH

Kriterium und weiteren Variablen unter angemessener Berücksichtigung der Messgenauigkeit sozusagen latent bestimmt werden.

Reliabilität. Mit dem Begriff *Reliabilität* (dt. Zuverlässigkeit) werden in der sozialwissenschaftlichen Forschung und Literatur Maße für die Messgenauigkeit bei der Erfassung von (latenten) Variablen bezeichnet. In der Psychometrie wird die Reliabilität einer Messung theoretisch ausgedrückt als Quotient aus der → *Varianz* der „wahren" Merkmalsausprägung (vgl. auch → *latente Variable*) und den potenziell messfehlerbehafteten beobachteten Werten. Für den eher theoretischen Fall einer perfekten Messung ohne Fehler würde dieser Quotient und damit die Reliabilität den Wert $r = 1$ annehmen. Die theoretische untere Grenze der Reliabilität ist durch den Wert $r = 0$ definiert und repräsentiert eine maximal ungenaue Messung ohne prädiktiven (oder sonstigen) Mehrwert. Die theoretische Definition der Reliabilität als Varianzverhältnis aus wahrem und beobachtetem Wert macht deutlich, dass in der Praxis (unterschiedliche) Methoden zur Abschätzung der Reliabilität gefunden werden müssen, da die „wahre" Merkmalsausprägung ja nicht bekannt ist und daher erst gemessen werden soll. Eine Gemeinsamkeit der unterschiedlichen Methoden der Reliabilitätsschätzung besteht darin, dass dazu meist unterschiedliche Quellen der → *Varianz* zueinander ins Verhältnis gesetzt werden.

Schwierigkeit, psychometrische. Die psychometrische Schwierigkeit ist neben der → *psychometrischen Trennschärfe* ein zentrales Merkmal von Testaufgaben (→ *Item*), das im Rahmen der psychometrischen Datenauswertung empirisch bestimmt werden kann. Die → *Klassische Testtheorie* (*KTT*) definiert die Schwierigkeit vereinfacht ausgedrückt durch den Anteil derjenigen Personen einer → *Stichprobe*, die eine Testaufgabe korrekt lösen. Wird eine Testaufgabe beispielsweise von 90 Prozent der Testteilnehmer gelöst, so weist die Aufgabe eine psychometrische Schwierigkeit von $\sigma = 0.9$ auf. Der Wertebereich der psychometrischen Schwierigkeit liegt dadurch im (theoretischen) Bereich von $\sigma = 0$ (sehr schwierige Aufgaben) bis $\sigma = 1.0$ (sehr einfache Aufgaben).

Signifikanz. Beim Vergleich von Mittelwerten zwischen zwei oder mehr Stichproben muss beurteilt werden können, ob Unterschiede statistisch aussagekräftig sind oder nur zufällig durch die Auswahl der Stichproben zustande gekommen sind. Für diese Beurteilung werden sogenannte *Signifikanztests* durchgeführt. Mit einer vorab festgelegten Irrtumswahrscheinlichkeit (also der Wahrscheinlichkeit, dass die Ergebnisse durch Zufall zustande kommen; oft 5 oder besser noch 1 Prozent) wird geprüft, ob sich *kein* Unterschied in der Grundgesamtheit finden lässt. Wenn also beispielsweise in der PISA-Stichprobe Unterschiede in der durchschnittlichen Kompetenz von Jungen und Mädchen gefunden werden, zeigt ein Signifikanztest an, ob diese Unterschiede dafür oder dagegen sprechen, dass in der Grundgesamtheit keine Unterschiede bestehen. Bei einem Signifikanzniveau von 5 Prozent wird man bei fünf von hundert Messungen fälschlicherweise von einem Unterschied ausgehen, bei einem Signifikanzniveau von 1 Prozent betrifft dies entsprechend eine Messung von hundert. Signifikanztests werden üblicherweise beim Vergleich statistischer Kennwerte durchgeführt (etwa bezogen auf Mittel-

© Waxmann Verlag GmbH

werte oder Korrelationen). Bei diesen Tests spielt neben der Größe des Unterschieds in den Kennwerten oder dem Betrag des Korrelationskoeffizienten auch die Größe der Stichprobe eine zentrale Rolle. Bei einer sehr umfangreichen Stichprobe kann ein Unterschied statistisch signifikant werden, der bei einer kleineren Stichprobe als nicht signifikant beurteilt worden wäre. Insofern sagt die statistische Signifikanz eines Unterschieds manchmal wenig über die Größenordnung und praktische Relevanz dieses Unterschieds aus. Häufig werden als Maß für die praktische Relevanz von Unterschieden sogenannte → *Effektstärken* angegeben.

Skalierung. Die Kompetenzen der in PISA untersuchten Schülerinnen und Schüler werden anhand von mehreren Aufgabenclustern in zahlreichen Testformen erfasst. Jede Schülerin und jeder Schüler bearbeitet demnach nur eine Auswahl aller eingesetzten Aufgaben (vgl. → *Booklet-Design*). Die Abbildung der mit unterschiedlichen Aufgaben gemessenen Kompetenzen auf einer gemeinsamen Skala wird als Skalierung bezeichnet. Weil die Aufgaben in den Testformen unterschiedlich schwierig sind, ist es dabei nicht ausreichend, einfach die Anzahl der gelösten beziehungsweise nicht gelösten Aufgaben für die Ermittlung der Kompetenz der Fünfzehnjährigen heranzuziehen. Für einen objektiven und fairen Vergleich der Kompetenzen zwischen Gruppen von Schülerinnen und Schülern (z. B. in verschiedenen Staaten, an unterschiedlichen Schulformen oder zwischen Mädchen und Jungen) muss auch die jeweilige Schwierigkeit der unterschiedlichen Aufgaben berücksichtigt werden. Einfache Modelle gehen bei der Skalierung davon aus, dass die Schwierigkeit der Aufgaben und die Fähigkeit der Testpersonen auf einer gemeinsamen Skala und damit eindimensional abgebildet werden können. Dabei ist dann anzunehmen, dass ein Schüler, der viele Aufgaben gelöst hat, eine hohe Fähigkeit besitzt und eine Aufgabe, die häufig (also von vielen Schülern) gelöst wurde, eine geringe Schwierigkeit aufweist.

Standardfehler. Eines der grundlegenden Ziele der PISA-Studien sind Schlussfolgerungen von in Stichproben erhobenen Kennwerten auf entsprechende Kennwerte in der Grundgesamtheit (also z. B. in PISA aller fünfzehnjährigen Schülerinnen und Schüler gegen Ende ihrer Pflichtschulzeit). Entsprechende statistische *Schätzungen* sind immer mit einer gewissen Unsicherheit und Messungenauigkeit verbunden. Diese Unsicherheit kann jedoch quantifiziert und damit in den Analysen berücksichtigt werden. Ein aussagekräftiges Maß dafür ist der sogenannte *Standardfehler*. Je kleiner der Standardfehler ist, desto präziser bildet ein Stichprobenkennwert die Grundgesamtheit ab. Der Standardfehler gibt an, in welchem Wertebereich der entsprechende Kennwert für die Grundgesamtheit mit einer bestimmten *Irrtumswahrscheinlichkeit* liegt. Über die Höhe dieser Irrtumswahrscheinlichkeit gibt es Konventionen, oft wird ein Wert von $p = 0.05$ festgelegt. Bei maximal 5 Prozent Irrtumswahrscheinlichkeit entspricht die Spannweite des Wertebereichs für den „wahren Wert“ ± 1.96 Standardfehler.

© Waxmann Verlag GmbH

Stichprobe. Eine Stichprobe bezeichnet die Auswahl einzelner Elemente aus der Grundgesamtheit (vgl. → *Grundgesamtheit*), die in einer wissenschaftlichen Studie untersucht werden soll. Idealerweise wird die Stichprobe so gezogen, dass sie repräsentativ für die Grundgesamtheit ist. Das bedeutet, dass bestimmte Merkmale der Grundgesamtheit entsprechend ihrem Anteil auch in der Stichprobe abgebildet sind. Wenn beispielsweise unter allen Schulen der Sekundarstufe I in Deutschland 30 Prozent Gymnasien sind, so sollte auch eine Schulstichprobe, die auf die Schulen der Sekundarstufe verallgemeinert werden soll, 30 Prozent Gymnasien enthalten. Für die Stichprobenziehung stehen zahlreiche Verfahren zur Verfügung, die unter anderem die Repräsentativität sichern können. Im Rahmen von PISA verfolgt die Stichprobe den Anspruch, repräsentativ für fünfzehnjährige Schülerinnen und Schüler in den teilnehmenden Bildungssystemen zu sein.

Stratifizierung. Stratifizierung bedeutet, dass man die Liste aller zur Grundgesamtheit gehörenden möglichen Studienteilnehmer (z. B. Schulen oder Schüler) vor der Stichprobenziehung unterteilt. Jeder Teil der Grundgesamtheit bildet ein Stratum, also eine Schicht. Aus jedem Stratum wird dann eine definierte Anzahl von Elementen gezogen (in PISA: Schulen, daraus dann im nächsten Schritt Schülerinnen und Schüler). Eine Stratifizierung stellt so beispielsweise sicher, dass für eine gesamtdeutsche Schulstichprobe aus jedem Bundesland Schulen aufgenommen werden. Jedes der 16 Bundesländer bildet für die Stichprobe zu PISA in Deutschland ein Stratum. Förderschulen und berufliche Schulen bilden ein eigenes Stratum, da sie im internationalen Vergleich relativ selten und über die Bundesländer in Deutschland ungleichmäßig verteilt sind. Aus jedem der so entstehenden 16 + 1 + 1 = 18 Strata werden so viele Schulen für die Teilnahme an PISA ausgewählt, wie es dem Anteil an fünfzehnjährigen Schülerinnen und Schülern im Stratum entspricht. Innerhalb der Bundesland-Strata wird noch eine weitere Schichtung vorgenommen, um der Sekundarschullandschaft in Deutschland gerecht zu werden: Alle Bundesländer haben Gymnasien, wobei die nicht gymnasialen Schularten sowohl in Bezug auf die Ausgestaltung als auch auf die Anzahl dieser Schularten zwischen den Bundesländern variieren. Deshalb wird jedes Bundesland-Stratum in Gymnasien und nicht gymnasiale Schulformen untergliedert, und aus beiden Gruppen wird eine festgelegte Anzahl Schulen für die Teilnahme an PISA gezogen.

Streuobstwiese. Süddeutsche Form der Landschaft, die durch zweidimensionale → *Normalverteilungen* („Apfelbäume“) gekennzeichnet ist. Die einzelnen Objekte („Äpfel“) können darüber hinaus eindimensional auf einer optischen Skala angeordnet werden („Anzahl der Wurmstiche“), die sich wiederum auf einer eindimensionalen Qualitätsskala abbilden lässt („von Würmern bevorzugter Geschmack“). Dieses sogenannte 1-PL-Modell ermöglicht eine eher einfache Beschreibung der Beziehung zwischen den Konstrukten und lässt entsprechende Rückschlüsse auf die Objekte zu. Ein sogenanntes 2-PL-Modell berücksichtigt zusätzlich die Größe der Objekte. In dieses Modell geht damit die → *Trennschärfe* in Bezug auf den Weg ein, den der Wurm bis zur Oberfläche zurückgelegt hat, was die Aussagekraft der Messung beträchtlich verbessert.

© Waxmann Verlag GmbH

Teilaufgabe. Siehe → *Item*.

Testform. Im Rahmen von Schulleistungsvergleichsstudien werden häufig so viele Aufgaben eingesetzt, dass jede Schülerin und jeder Schüler lediglich eine Auswahl aller Aufgaben bearbeiten kann (vgl. → *Booklet-Design*). Bei Tests, die wie bei PISA 2000 bis PISA 2012 als Papier-und-Bleistift-Tests durchgeführt wurden, spricht man in aller Regel von unterschiedlichen *Testheften*, die den Schülerinnen und Schülern vorgelegt wurden. Diese Testhefte enthalten typischerweise mehrere Aufgabengruppen (sogenannte Cluster), die wiederum aus mehreren Aufgaben (Units) mit Teilaufgaben (Items) bestehen. Im Kontext von computerbasierten Tests, wie bei PISA 2015, werden diese Zusammenstellungen von Aufgaben meist Testformen genannt. Jede Schülerin und jeder Schüler bearbeitet also eine Testform am Computer, die mehrere Cluster von Aufgaben umfasst.

Trennschärfe, psychometrische. Die psychometrische Trennschärfe ist neben der → *psychometrischen Schwierigkeit* ein zentrales Merkmal von Testaufgaben (→ *Item*), welche im Rahmen der psychometrischen Datenauswertung empirisch bestimmt werden kann. Die Trennschärfe ist dabei ein → *Modellparameter* und kann im theoretischen Bereich von $\alpha = -1$ (negative Trennschärfe) über $\alpha = 0$ (niedrige Trennschärfe) bis $\alpha = 1$ (hohe Trennschärfe) variieren. Die Trennschärfe einer Aufgabe bzw. eines Items quantifiziert damit beispielsweise, wie gut eine Testaufgabe dazu geeignet ist, zwischen Personen mit hoher und niedriger Kompetenz in Naturwissenschaften zu unterscheiden. Weisen Testaufgaben eine sehr niedrige (positive) Trennschärfe auf (d. h. deren Wert geht gegen $\alpha = 0$), so können sowohl Personen mit hoher als auch niedriger Kompetenz die betreffende Aufgabe gleichermaßen gut lösen. Negative Trennschärfen ($\alpha < 0$) weisen hingegen darauf hin, dass zwischen der → *latenten Variablen* und der Testaufgabe als → *manifestem Indikator* ein negativer Zusammenhang besteht. Die Testaufgabe ist in diesem Fall also ungeeignet als Indikator für die gewünschte latente Eigenschaft (z. B. Kompetenz). Weil eine negative Trennschärfe dem angenommenen positiven Zusammenhang zwischen der → *latenten Variablen* und dem → *manifesten Indikator* widerspricht, wird die betreffende Aufgabe in der Regel aus dem Test ausgeschlossen.

Variable, latente *(*auch: *latente Fähigkeitsdimension).* Als latente Variable, latente Fähigkeitsdimension oder auch latentes Merkmal werden in den Sozialwissenschaften Eigenschaften bezeichnet, die zwischen Personen variieren, allerdings nicht direkt empirisch beobachtet werden können. Latenten Fähigkeitsdimensionen (z. B. der naturwissenschaftlichen Kompetenz in PISA) liegen in der Regel explizite Zusammenhangsannahmen zu beobachtbarem Verhalten zugrunde, welche die latente Fähigkeitsdimension anhand manifester → *Indikatoren* messbar und damit der Forschung zugänglich machen.

Varianz. In der statistischen Datenanalyse ist die Varianz ein Streuungsmaß, das die Verteilung der Werte einer (beobachteten) Variable beschreibt. Sie vermittelt so einen Eindruck über die Heterogenität einer Verteilung, also etwa, wie breit die Mittelwerte der naturwissenschaftlichen Kompetenz streuen. Die Varianz für diskrete Zufallsvariab-

© Waxmann Verlag GmbH

len ist definiert durch die Summe aller quadratischen Abweichungen der Beobachtungen von deren Erwartungswert (Mittelwert). Die Quadratwurzel aus der Varianz bildet die → *Standardabweichung*.

© Waxmann Verlag GmbH

15 Anhang

© Waxmann Verlag GmbH

Tabelle A1: Mittelwerte, Streuungen und Perzentile naturwissenschaftlicher Kompetenz

OECD-Staaten					Perzentile					
	M	*(SE)*	*SD*	*(SE)*	5%	10%	25%	75%	90%	95%
Japan	538	(3.0)	93	(1.6)	375	412	475	605	655	683
Estland	534	(2.1)	89	(1.1)	384	416	473	597	648	677
Finnland	531	(2.4)	96	(1.3)	364	402	466	599	651	681
Kanada	528	(2.1)	92	(0.9)	369	404	465	593	644	674
Korea	516	(3.1)	95	(1.5)	352	388	451	584	636	665
Neuseeland	513	(2.4)	104	(1.4)	341	374	439	588	647	682
Slowenien	513	(1.3)	95	(1.1)	354	386	445	581	636	667
Australien	510	(1.5)	102	(0.9)	336	372	438	583	639	672
Vereinigtes Königreich	509	(2.6)	100	(1.0)	345	377	438	581	638	670
Deutschland	509	(2.7)	99	(1.5)	342	376	439	580	636	669
Niederlande	509	(2.3)	101	(1.5)	341	372	434	583	638	668
Schweiz	506	(2.9)	100	(1.5)	339	373	433	580	632	662
Irland	503	(2.4)	89	(1.3)	356	387	441	565	618	648
Belgien	502	(2.3)	100	(1.2)	332	364	429	577	629	657
Dänemark	502	(2.4)	90	(1.1)	351	383	440	565	617	648
Polen	501	(2.5)	91	(1.3)	354	384	437	565	619	650
Portugal	501	(2.4)	92	(1.1)	349	379	435	568	620	649
Norwegen	498	(2.3)	96	(1.3)	338	370	432	566	622	655
Vereinigte Staaten	496	(3.2)	99	(1.4)	336	368	425	567	626	658
Österreich	495	(2.4)	97	(1.3)	335	365	424	565	621	652
Frankreich	495	(2.1)	102	(1.4)	322	355	421	571	623	652
Schweden	493	(3.6)	102	(1.4)	322	357	421	567	625	658
Tschechische Republik	493	(2.3)	95	(1.4)	338	367	424	561	618	650
Spanien	493	(2.1)	88	(1.1)	344	374	432	556	605	633
Lettland	490	(1.6)	82	(1.1)	355	382	432	548	596	623
Luxemburg	483	(1.1)	100	(1.1)	323	351	407	556	615	649
Italien	481	(2.5)	91	(1.4)	328	359	415	547	599	626
Ungarn	477	(2.4)	96	(1.6)	319	347	406	547	601	630
Island	473	(1.7)	91	(1.2)	324	354	408	538	593	622
Israel	467	(3.4)	106	(1.6)	295	327	389	544	606	640
Slowakische Republik	461	(2.6)	99	(1.5)	296	329	391	532	588	621
Griechenland	455	(3.9)	92	(1.8)	305	333	388	522	575	604
Chile	447	(2.4)	86	(1.3)	308	336	385	509	560	589
Türkei	425	(3.9)	79	(1.9)	301	325	368	482	532	560
Mexiko	416	(2.1)	71	(1.1)	301	325	366	464	510	535
OECD-Durchschnitt	493	(0.4)	94	(0.2)	336	368	426	561	615	645
OECD-Partnerstaaten										
Singapur	556	(1.2)	104	(0.9)	373	412	485	631	683	712
Chinesisch Taipeh	532	(2.7)	100	(1.9)	358	395	465	603	655	685
Macau (China)	529	(1.1)	81	(1.0)	389	420	474	586	630	656
Vietnam	525	(3.9)	77	(2.3)	404	428	470	576	624	655
Hongkong (China)	523	(2.5)	81	(1.4)	379	413	473	579	622	646
B-S-J-G (China)*	518	(4.6)	103	(2.5)	341	377	445	595	649	677
Russische Föderation	487	(2.9)	82	(1.1)	352	379	428	544	595	623
Litauen	475	(2.7)	91	(1.4)	329	357	410	540	597	626
Kroatien	475	(2.5)	89	(1.2)	332	360	411	538	593	624
Malta	465	(1.6)	118	(1.5)	273	310	382	548	618	656
Bulgarien	446	(4.4)	102	(2.1)	283	313	370	521	578	611
Vereinigte Arabische Emirate	437	(2.4)	99	(1.1)	284	312	364	505	571	608
Uruguay	435	(2.2)	87	(1.3)	301	326	372	496	552	583
Rumänien	435	(3.2)	79	(1.7)	309	334	379	488	539	570
Republik Moldau	428	(2.0)	86	(1.4)	290	318	367	488	541	570
Albanien	427	(3.3)	78	(1.5)	301	328	373	481	530	558
Republik Trinidad und Tobago	425	(1.4)	94	(1.1)	279	306	356	491	551	585
Thailand	421	(2.8)	78	(1.6)	301	324	365	473	528	559
Costa Rica	420	(2.1)	70	(1.2)	310	332	370	466	514	541
Katar	418	(1.0)	99	(0.7)	268	295	344	486	554	589
Kolumbien	416	(2.4)	80	(1.3)	291	315	357	471	524	554
Republik Montenegro	411	(1.0)	85	(0.9)	277	304	352	468	526	558
Georgien	411	(2.4)	91	(1.3)	267	297	348	471	531	566
Jordanien	409	(2.7)	84	(1.6)	268	299	351	468	517	544
Indonesien	403	(2.6)	68	(1.6)	296	319	356	447	493	522
Brasilien	401	(2.3)	89	(1.3)	265	291	337	460	522	558
Peru	397	(2.4)	77	(1.4)	278	301	342	448	500	529
Libanesische Republik	386	(3.4)	90	(1.8)	249	276	322	446	511	545
Tunesien	386	(2.1)	65	(1.6)	287	306	341	428	472	500
Republik Mazedonien	384	(1.2)	85	(1.3)	248	277	325	440	496	528
Kosovo	378	(1.7)	71	(1.1)	266	289	328	426	474	501
Algerien	376	(2.6)	69	(1.5)	268	291	329	419	465	496
Dominikanische Republik	332	(2.6)	72	(1.8)	224	244	281	376	429	461

*B-S-J-G (China) bezieht sich auf vier Provinzen in China, die an der PISA-Studie teilgenommen haben: Peking, Shanghai, Jiangsu und Guangdong.

signifikant über dem OECD-Durchschnitt

nicht signifikant verschieden vom OECD-Durchschnitt

signifikant unter dem OECD-Durchschnitt

© Waxmann Verlag GmbH

Tabelle A2: Mittelwerte, Streuungen und Perzentile der Teilskala *Phänomene naturwissenschaftlich erklären*

OECD-Staaten					Perzentile					
	M	*(SE)*	*SD*	*(SE)*	5%	10%	25%	75%	90%	95%
Japan	539	(3.3)	97	(1.7)	370	408	473	607	660	691
Finnland	534	(2.4)	95	(1.2)	371	407	470	601	653	684
Estland	533	(2.0)	90	(1.2)	384	415	470	596	651	680
Kanada	530	(2.1)	93	(1.2)	373	406	467	595	647	679
Slowenien	515	(1.5)	96	(1.2)	357	389	448	583	640	672
Neuseeland	511	(2.6)	106	(1.5)	337	370	435	587	649	685
Deutschland	511	(2.8)	100	(1.5)	345	378	440	582	640	674
Australien	510	(1.6)	105	(0.9)	335	371	436	585	645	679
Korea	510	(3.4)	98	(1.8)	346	381	442	580	636	668
Vereinigtes Königreich	509	(2.7)	103	(1.6)	341	374	435	584	643	676
Niederlande	509	(2.5)	99	(1.6)	347	377	437	581	637	668
Irland	505	(2.5)	91	(1.3)	357	388	443	568	623	655
Schweiz	505	(3.1)	102	(1.8)	335	369	431	580	635	665
Norwegen	502	(2.3)	99	(1.3)	338	371	433	571	629	663
Dänemark	502	(2.7)	94	(1.2)	347	379	437	567	623	656
Polen	501	(2.8)	95	(1.4)	349	380	434	566	624	658
Österreich	499	(2.7)	101	(1.5)	333	366	427	571	629	663
Belgien	499	(2.4)	102	(1.3)	327	361	425	574	629	659
Portugal	498	(2.5)	94	(1.4)	345	374	430	566	622	654
Schweden	498	(3.7)	104	(1.7)	324	360	426	572	633	666
Tschechische Republik	496	(2.5)	97	(1.5)	341	371	426	564	623	657
Spanien	494	(2.2)	92	(1.3)	340	373	431	558	611	642
Vereinigte Staaten	492	(3.4)	102	(1.6)	327	360	418	565	625	659
Frankreich	488	(2.2)	102	(1.4)	316	351	414	564	618	648
Lettland	488	(1.8)	84	(1.1)	349	378	428	547	597	625
Luxemburg	482	(1.1)	105	(1.2)	317	346	403	558	622	657
Italien	481	(2.7)	96	(1.6)	320	352	412	550	604	635
Ungarn	478	(2.5)	98	(1.6)	319	348	408	548	604	637
Island	468	(2.0)	93	(1.5)	316	348	403	533	589	619
Slowakische Republik	464	(2.7)	102	(1.6)	295	330	392	535	596	630
Israel	463	(3.5)	109	(1.9)	284	319	383	541	605	641
Griechenland	454	(3.9)	94	(1.7)	300	329	385	522	577	608
Chile	446	(2.6)	90	(1.5)	300	330	383	511	564	595
Türkei	426	(4.2)	84	(2.0)	295	320	365	485	538	567
Mexiko	414	(2.3)	74	(1.4)	295	320	363	464	509	536
OECD-Durchschnitt	493	(0.5)	97	(0.3)	334	366	424	562	618	650
OECD-Partnerstaaten										
Singapur	553	(1.5)	109	(1.0)	365	405	478	632	688	721
Chinesisch Taipeh	536	(2.8)	102	(2.0)	357	396	468	610	662	693
Macau (China)	528	(1.4)	86	(1.1)	383	415	471	588	636	664
Hongkong (China)	524	(2.6)	83	(1.5)	378	413	471	582	628	652
B-S-J-G (China)*	520	(4.7)	107	(2.6)	337	376	445	598	655	685
Russische Föderation	486	(3.2)	84	(1.2)	350	377	427	545	596	626
Litauen	478	(2.7)	93	(1.6)	329	359	411	542	601	634
Kroatien	476	(2.4)	89	(1.7)	333	362	412	537	593	626
Bulgarien	449	(4.5)	106	(2.1)	277	310	371	526	587	621
Vereinigte Arabische Emirate	437	(2.5)	104	(1.5)	275	306	362	509	578	617
Uruguay	434	(2.3)	88	(1.5)	297	323	370	495	551	584
Costa Rica	420	(2.3)	72	(1.2)	306	330	369	467	516	544
Thailand	419	(2.9)	82	(1.7)	293	317	361	472	528	561
Katar	417	(1.2)	105	(1.1)	257	286	339	491	561	598
Kolumbien	412	(2.6)	82	(1.3)	283	309	353	468	522	554
Republik Montenegro	411	(1.3)	86	(1.3)	277	304	352	469	526	559
Brasilien	403	(2.7)	95	(1.4)	258	286	335	465	531	571
Peru	392	(2.6)	81	(1.7)	266	291	336	446	500	531
Tunesien	385	(2.3)	69	(1.7)	277	299	337	430	476	504
Dominikanische Republik	332	(2.6)	75	(1.9)	217	241	281	379	432	466
Albanien	m	m	m	m	m	m	m	m	m	m
Algerien	m	m	m	m	m	m	m	m	m	m
Georgien	m	m	m	m	m	m	m	m	m	m
Indonesien	m	m	m	m	m	m	m	m	m	m
Jordanien	m	m	m	m	m	m	m	m	m	m
Kosovo	m	m	m	m	m	m	m	m	m	m
Libanesische Republik	m	m	m	m	m	m	m	m	m	m
Malta	m	m	m	m	m	m	m	m	m	m
Republik Mazedonien	m	m	m	m	m	m	m	m	m	m
Republik Moldau	m	m	m	m	m	m	m	m	m	m
Republik Trinidad und Tobago	m	m	m	m	m	m	m	m	m	m
Rumänien	m	m	m	m	m	m	m	m	m	m
Vietnam	m	m	m	m	m	m	m	m	m	m

*B-S-J-G (China) bezieht sich auf vier Provinzen in China, die an der PISA-Studie teilgenommen haben: Peking, Shanghai, Jiangsu und Guangdong.

m: fehlende Angaben

signifikant über dem OECD-Durchschnitt | nicht signifikant verschieden vom OECD-Durchschnitt | signifikant unter dem OECD-Durchschnitt

© Waxmann Verlag GmbH

Tabelle A3: Mittelwerte, Streuungen und Perzentile der Teilskala *naturwissenschaftliche Forschung bewerten und Untersuchungen planen*

OECD-Staaten					Perzentile					
	M	*(SE)*	*SD*	*(SE)*	5%	10%	25%	75%	90%	95%
Japan	536	(3.3)	98	(2.1)	366	405	472	605	656	686
Estland	535	(2.6)	93	(1.6)	374	411	473	600	652	681
Kanada	530	(2.7)	102	(1.6)	355	394	462	601	657	690
Finnland	529	(2.9)	102	(1.6)	352	393	461	601	657	689
Neuseeland	517	(3.1)	114	(1.9)	326	365	436	598	663	699
Korea	515	(3.3)	100	(2.0)	342	380	448	587	641	671
Australien	512	(2.0)	111	(1.7)	324	364	435	591	652	689
Slowenien	511	(2.0)	99	(1.5)	344	378	442	582	639	669
Niederlande	511	(2.5)	109	(1.6)	331	364	430	592	652	684
Vereinigtes Königreich	508	(2.8)	107	(1.5)	330	366	433	583	644	679
Belgien	507	(2.5)	105	(1.4)	328	363	431	586	639	668
Schweiz	507	(3.5)	108	(2.6)	324	361	430	586	644	675
Deutschland	506	(2.9)	103	(1.7)	329	367	435	580	636	668
Dänemark	504	(2.6)	95	(1.8)	344	380	441	570	623	655
Vereinigte Staaten	503	(3.6)	105	(1.8)	330	365	429	578	639	673
Portugal	502	(2.7)	98	(1.5)	335	371	434	574	626	655
Polen	502	(3.0)	94	(1.8)	346	380	436	568	621	652
Irland	500	(2.6)	95	(1.7)	342	377	436	566	620	651
Frankreich	498	(2.5)	109	(2.2)	310	350	421	578	634	666
Norwegen	493	(2.6)	99	(1.4)	326	362	426	562	619	652
Schweden	491	(4.0)	108	(2.0)	310	348	416	568	628	663
Lettland	489	(2.0)	86	(1.2)	346	377	431	549	599	629
Spanien	489	(2.7)	91	(1.6)	334	368	427	553	605	633
Österreich	488	(2.6)	102	(1.9)	317	352	417	561	618	649
Tschechische Republik	486	(2.8)	96	(1.7)	326	359	419	554	611	643
Luxemburg	479	(1.7)	105	(1.5)	309	340	401	555	614	648
Italien	477	(2.7)	98	(1.7)	312	347	409	548	602	632
Island	476	(2.5)	94	(1.5)	320	352	410	543	597	626
Ungarn	474	(2.8)	101	(1.9)	307	339	401	547	604	635
Israel	471	(3.8)	113	(1.9)	287	324	389	554	620	656
Slowakische Republik	457	(3.2)	105	(2.2)	277	319	387	531	589	622
Griechenland	453	(4.2)	99	(2.3)	290	324	383	525	581	612
Chile	443	(2.9)	93	(1.6)	290	322	378	510	563	594
Türkei	428	(4.0)	83	(2.1)	296	322	368	487	538	565
Mexiko	415	(2.9)	83	(1.6)	279	309	359	471	522	552
OECD-Durchschnitt	493	(0.5)	100	(0.3)	324	360	423	564	620	652
OECD-Partnerstaaten										
Singapur	560	(1.4)	110	(1.5)	365	409	489	639	694	725
Macau (China)	525	(1.9)	85	(1.3)	379	412	470	584	631	657
Chinesisch Taipeh	525	(3.1)	104	(2.2)	342	384	457	598	653	684
Hongkong (China)	524	(3.0)	83	(1.6)	374	411	474	581	624	648
B-S-J-G (China)*	517	(5.1)	110	(2.5)	328	369	441	598	656	686
Russische Föderation	484	(3.3)	88	(1.4)	341	370	421	545	599	629
Litauen	478	(2.9)	98	(1.5)	317	350	408	548	605	636
Kroatien	473	(2.9)	95	(1.7)	319	349	406	539	596	628
Bulgarien	440	(4.8)	105	(2.3)	271	303	363	516	579	615
Uruguay	433	(2.9)	91	(1.9)	288	316	367	497	554	587
Vereinigte Arabische Emirate	431	(2.7)	105	(1.5)	265	299	356	503	572	611
Thailand	423	(3.5)	88	(1.9)	285	312	362	481	539	573
Costa Rica	422	(2.7)	78	(2.0)	298	324	368	473	525	555
Kolumbien	420	(2.9)	89	(1.8)	279	307	356	480	538	572
Katar	414	(1.5)	103	(1.2)	255	287	341	484	553	593
Republik Montenegro	408	(1.6)	95	(1.5)	255	287	342	473	532	566
Peru	399	(3.1)	81	(1.7)	271	297	342	453	506	536
Brasilien	398	(2.9)	97	(2.0)	247	277	330	462	527	566
Tunesien	379	(2.6)	71	(2.0)	268	291	331	425	472	502
Dominikanische Republik	324	(3.5)	85	(2.3)	192	219	265	378	435	472
Albanien	m	m	m	m	m	m	m	m	m	m
Algerien	m	m	m	m	m	m	m	m	m	m
Georgien	m	m	m	m	m	m	m	m	m	m
Indonesien	m	m	m	m	m	m	m	m	m	m
Jordanien	m	m	m	m	m	m	m	m	m	m
Kosovo	m	m	m	m	m	m	m	m	m	m
Libanesische Republik	m	m	m	m	m	m	m	m	m	m
Malta	m	m	m	m	m	m	m	m	m	m
Republik Mazedonien	m	m	m	m	m	m	m	m	m	m
Republik Moldau	m	m	m	m	m	m	m	m	m	m
Republik Trinidad und Tobago	m	m	m	m	m	m	m	m	m	m
Rumänien	m	m	m	m	m	m	m	m	m	m
Vietnam	m	m	m	m	m	m	m	m	m	m

*B-S-J-G (China) bezieht sich auf vier Provinzen in China, die an der PISA-Studie teilgenommen haben: Peking, Shanghai, Jiangsu und Guangdong.

m: fehlende Angaben

signifikant über dem OECD-Durchschnitt

nicht signifikant verschieden vom OECD-Durchschnitt

signifikant unter dem OECD-Durchschnitt

© Waxmann Verlag GmbH

Tabelle A4: Mittelwerte, Streuungen und Perzentile der Teilskala *Daten und Evidenz naturwissenschaftlich interpretieren*

OECD-Staaten					Perzentile					
	M	*(SE)*	*SD*	*(SE)*	5%	10%	25%	75%	90%	95%
Japan	541	(3.1)	97	(2.1)	371	410	477	609	661	691
Estland	537	(2.7)	95	(1.5)	375	409	472	603	658	689
Finnland	529	(2.8)	104	(1.5)	349	391	460	602	659	692
Kanada	525	(2.7)	97	(0.9)	358	395	460	592	647	679
Korea	523	(3.2)	100	(1.5)	350	389	457	594	648	678
Neuseeland	512	(2.5)	104	(1.5)	337	373	438	587	645	681
Slowenien	512	(2.0)	102	(1.4)	340	377	442	584	643	674
Vereinigtes Königreich	509	(2.9)	103	(1.3)	338	373	436	582	641	674
Deutschland	509	(3.0)	105	(1.6)	332	369	436	583	643	678
Australien	508	(1.8)	104	(1.1)	332	369	436	583	641	674
Niederlande	506	(2.5)	109	(1.8)	323	359	429	585	644	676
Schweiz	506	(3.0)	102	(1.8)	335	369	432	581	637	666
Belgien	503	(2.5)	103	(1.3)	328	362	430	580	633	663
Portugal	503	(2.6)	96	(1.2)	341	375	435	573	625	656
Polen	501	(2.6)	91	(1.4)	350	382	437	565	618	649
Frankreich	501	(2.5)	107	(1.6)	318	354	423	582	637	666
Irland	500	(2.7)	91	(1.7)	349	382	438	564	618	648
Dänemark	500	(2.6)	93	(1.1)	345	378	436	565	617	649
Norwegen	498	(2.8)	101	(1.5)	330	365	429	568	627	661
Vereinigte Staaten	497	(3.5)	99	(1.7)	336	367	426	567	626	657
Lettland	494	(1.7)	88	(1.3)	348	378	432	555	607	637
Tschechische Republik	493	(2.8)	102	(1.8)	326	359	421	565	625	660
Spanien	493	(2.4)	93	(1.3)	335	369	429	559	611	641
Österreich	493	(2.6)	101	(1.6)	325	358	421	565	623	655
Schweden	490	(3.7)	107	(1.6)	311	349	415	565	627	661
Luxemburg	486	(1.8)	100	(1.3)	324	354	412	559	617	650
Italien	482	(2.9)	97	(1.6)	319	352	413	552	606	637
Island	478	(2.1)	97	(1.3)	319	351	409	547	604	634
Ungarn	476	(2.7)	102	(1.7)	307	339	402	551	607	639
Israel	467	(3.7)	109	(1.8)	290	325	388	547	609	645
Slowakische Republik	459	(2.9)	103	(1.6)	284	322	388	532	590	624
Griechenland	454	(4.1)	95	(2.0)	298	330	386	523	576	607
Chile	447	(2.7)	89	(1.6)	303	333	384	510	564	595
Türkei	423	(4.2)	83	(2.0)	291	318	364	482	534	562
Mexiko	415	(2.3)	74	(1.4)	295	320	363	465	512	539
OECD-Durchschnitt	493	(0.5)	98	(0.3)	329	363	424	564	619	651
OECD-Partnerstaaten										
Singapur	556	(1.4)	103	(1.4)	374	414	487	630	683	712
Chinesisch Taipeh	533	(2.9)	104	(2.1)	351	392	464	606	660	692
Macau (China)	532	(1.3)	84	(1.0)	387	421	477	591	636	662
Hongkong (China)	521	(2.7)	85	(1.5)	369	405	468	580	625	650
B-S-J-G (China)*	516	(4.8)	107	(2.7)	334	372	441	596	652	683
Russische Föderation	489	(3.0)	87	(1.4)	346	376	427	550	603	634
Kroatien	476	(2.7)	95	(1.5)	322	352	408	544	600	631
Litauen	471	(3.0)	95	(1.6)	316	348	404	539	595	627
Bulgarien	445	(4.6)	104	(2.3)	278	310	368	521	581	614
Vereinigte Arabische Emirate	437	(2.8)	102	(1.3)	278	308	363	508	574	611
Uruguay	436	(2.4)	93	(1.5)	291	318	368	502	561	595
Thailand	422	(3.1)	80	(1.7)	299	323	366	474	529	562
Katar	418	(1.0)	101	(1.0)	262	290	343	488	555	592
Kolumbien	416	(2.5)	81	(1.6)	289	315	359	471	523	553
Costa Rica	415	(2.6)	73	(1.3)	299	322	364	464	513	540
Republik Montenegro	410	(1.7)	90	(1.2)	265	295	347	471	529	563
Peru	398	(2.7)	80	(1.4)	273	298	341	452	505	537
Brasilien	398	(2.6)	91	(1.8)	257	285	333	457	521	558
Tunesien	390	(2.7)	69	(1.9)	283	305	342	434	480	511
Dominikanische Republik	330	(2.9)	78	(2.2)	211	236	277	378	434	467
Albanien	m	m	m	m	m	m	m	m	m	m
Algerien	m	m	m	m	m	m	m	m	m	m
Georgien	m	m	m	m	m	m	m	m	m	m
Indonesien	m	m	m	m	m	m	m	m	m	m
Jordanien	m	m	m	m	m	m	m	m	m	m
Kosovo	m	m	m	m	m	m	m	m	m	m
Libanesische Republik	m	m	m	m	m	m	m	m	m	m
Malta	m	m	m	m	m	m	m	m	m	m
Republik Mazedonien	m	m	m	m	m	m	m	m	m	m
Republik Moldau	m	m	m	m	m	m	m	m	m	m
Republik Trinidad und Tobago	m	m	m	m	m	m	m	m	m	m
Rumänien	m	m	m	m	m	m	m	m	m	m
Vietnam	m	m	m	m	m	m	m	m	m	m

*B-S-J-G (China) bezieht sich auf vier Provinzen in China, die an der PISA-Studie teilgenommen haben: Peking, Shanghai, Jiangsu und Guangdong.

m: fehlende Angaben

signifikant über dem OECD-Durchschnitt

nicht signifikant verschieden vom OECD-Durchschnitt

signifikant unter dem OECD-Durchschnitt

© Waxmann Verlag GmbH

Tabelle A5: Mittelwerte, Streuungen und Perzentile mathematischer Kompetenz

OECD-Staaten					Perzentile					
	M	*(SE)*	*SD*	*(SE)*	5%	10%	25%	75%	90%	95%
Japan	532	(3.0)	88	(1.7)	381	416	474	594	643	672
Korea	524	(3.7)	100	(1.8)	353	391	458	594	649	681
Schweiz	521	(2.9)	96	(1.6)	358	394	455	590	641	671
Estland	520	(2.0)	80	(1.1)	386	415	464	576	623	650
Kanada	516	(2.3)	88	(1.1)	368	400	456	577	627	657
Niederlande	512	(2.2)	92	(1.5)	356	390	449	579	627	655
Dänemark	511	(2.2)	81	(1.2)	376	405	457	567	614	639
Finnland	511	(2.3)	82	(1.3)	372	404	456	568	614	642
Slowenien	510	(1.3)	88	(1.3)	363	394	449	572	622	651
Belgien	507	(2.4)	97	(1.5)	341	374	438	579	630	657
Deutschland	506	(2.9)	89	(1.4)	356	389	445	568	620	650
Polen	504	(2.4)	88	(1.7)	363	391	443	565	617	649
Irland	504	(2.1)	80	(1.4)	371	400	450	559	606	633
Norwegen	502	(2.2)	85	(1.1)	359	391	444	561	610	638
Österreich	497	(2.9)	95	(1.8)	337	370	431	564	618	648
Neuseeland	495	(2.3)	92	(1.3)	342	375	431	560	613	646
Australien	494	(1.6)	93	(1.2)	339	371	430	559	613	645
Schweden	494	(3.2)	90	(1.7)	342	376	433	557	609	638
Frankreich	493	(2.1)	95	(1.5)	331	364	425	564	613	639
Vereinigtes Königreich	492	(2.5)	93	(1.4)	337	371	430	556	610	641
Tschechische Republik	492	(2.4)	91	(1.7)	340	373	431	555	608	639
Portugal	492	(2.5)	96	(1.3)	332	365	424	561	614	644
Italien	490	(2.8)	94	(1.7)	334	368	426	555	610	640
Island	488	(2.0)	93	(1.3)	333	367	424	553	608	640
Spanien	486	(2.2)	85	(1.3)	342	374	428	546	593	621
Luxemburg	486	(1.3)	94	(1.2)	334	363	417	553	607	638
Lettland	482	(1.9)	78	(1.2)	353	382	430	536	582	608
Ungarn	477	(2.5)	94	(1.7)	321	351	411	543	598	627
Slowakische Republik	475	(2.7)	95	(1.6)	312	349	412	543	596	625
Israel	470	(3.6)	103	(2.2)	296	332	396	545	601	634
Vereinigte Staaten	470	(3.2)	88	(1.5)	323	355	408	532	585	613
Griechenland	454	(3.8)	89	(1.8)	306	336	391	517	570	598
Chile	423	(2.5)	85	(1.4)	284	313	363	483	534	563
Türkei	420	(4.1)	82	(2.3)	291	317	363	477	529	559
Mexiko	408	(2.2)	75	(1.3)	284	312	357	459	505	533
OECD-Durchschnitt	490	(0.4)	89	(0.3)	340	373	428	553	605	634
OECD-Partnerstaaten										
Singapur	564	(1.5)	95	(0.8)	399	436	500	632	682	711
Hongkong (China)	548	(3.0)	90	(1.5)	389	426	490	611	659	687
Macau (China)	544	(1.1)	80	(1.1)	408	439	491	599	643	669
Chinesisch Taipeh	542	(3.0)	103	(1.9)	364	404	474	616	670	701
B-S-J-G (China)*	531	(4.9)	106	(2.5)	351	388	458	609	664	695
Vietnam	495	(4.5)	84	(2.7)	361	388	436	551	604	636
Russische Föderation	494	(3.1)	83	(1.3)	357	387	437	552	601	629
Malta	479	(1.7)	110	(1.4)	289	331	405	558	616	648
Litauen	478	(2.3)	86	(1.4)	337	365	419	539	590	620
Kroatien	464	(2.8)	88	(1.6)	322	351	402	525	580	612
Rumänien	444	(3.8)	86	(2.1)	305	334	384	502	557	590
Bulgarien	441	(4.0)	97	(2.4)	284	315	371	509	568	601
Vereinigte Arabische Emirate	427	(2.4)	97	(1.3)	275	306	360	493	557	593
Republik Moldau	420	(2.5)	90	(1.5)	271	303	358	482	536	568
Uruguay	418	(2.5)	87	(1.7)	281	309	357	477	532	565
Republik Montenegro	418	(1.5)	87	(1.4)	279	308	358	477	531	563
Republik Trinidad und Tobago	417	(1.4)	96	(1.2)	265	294	348	484	545	578
Thailand	415	(3.0)	82	(1.9)	286	313	360	468	521	555
Albanien	413	(3.4)	86	(1.6)	272	303	354	472	525	556
Georgien	404	(2.8)	94	(2.2)	250	285	341	467	525	559
Katar	402	(1.3)	99	(1.0)	248	278	331	470	536	573
Costa Rica	400	(2.5)	68	(1.4)	292	315	353	445	489	517
Libanesische Republik	396	(3.7)	101	(2.0)	236	268	324	464	531	568
Kolumbien	390	(2.3)	77	(1.3)	269	293	335	441	492	522
Peru	387	(2.7)	83	(1.4)	254	283	329	442	495	526
Indonesien	386	(3.1)	80	(2.0)	264	289	331	436	492	528
Jordanien	380	(2.7)	86	(2.1)	238	271	324	439	489	519
Brasilien	377	(2.9)	89	(1.7)	240	267	315	434	496	533
Republik Mazedonien	371	(1.3)	96	(1.6)	217	251	306	434	496	533
Tunesien	367	(3.0)	84	(2.3)	235	263	310	421	476	510
Kosovo	362	(1.6)	75	(1.4)	238	265	310	413	460	487
Algerien	360	(3.0)	71	(1.5)	247	271	312	405	452	481
Dominikanische Republik	328	(2.7)	69	(2.0)	220	243	281	373	418	446

*B-S-J-G (China) bezieht sich auf vier Provinzen in China, die an der PISA-Studie teilgenommen haben: Peking, Shanghai, Jiangsu und Guangdong.

signifikant über dem OECD-Durchschnitt

nicht signifikant verschieden vom OECD-Durchschnitt

signifikant unter dem OECD-Durchschnitt

© Waxmann Verlag GmbH

Tabelle A6: Mittelwerte, Streuungen und Perzentile der Lesekompetenz

OECD-Staaten					Perzentile					
	M	*(SE)*	*SD*	*(SE)*	5%	10%	25%	75%	90%	95%
Kanada	527	(2.3)	93	(1.3)	366	404	466	591	642	671
Finnland	526	(2.5)	94	(1.5)	359	401	469	592	640	668
Irland	521	(2.5)	86	(1.5)	373	406	463	582	629	657
Estland	519	(2.2)	87	(1.2)	369	404	460	581	630	659
Korea	517	(3.5)	97	(1.7)	345	386	455	586	637	666
Japan	516	(3.2)	92	(1.8)	352	391	457	581	629	656
Norwegen	513	(2.5)	99	(1.7)	342	381	449	583	636	666
Neuseeland	509	(2.4)	105	(1.7)	327	368	439	584	643	674
Deutschland	509	(3.0)	100	(1.6)	334	375	442	581	634	664
Polen	506	(2.5)	90	(1.3)	349	386	446	570	617	644
Slowenien	505	(1.5)	92	(1.3)	346	382	444	570	621	648
Niederlande	503	(2.4)	101	(1.6)	330	368	434	577	630	658
Australien	503	(1.7)	103	(1.1)	324	365	435	576	631	662
Schweden	500	(3.5)	102	(1.5)	321	364	433	573	625	655
Dänemark	500	(2.5)	87	(1.2)	347	383	443	561	608	635
Frankreich	499	(2.5)	112	(2.0)	299	344	423	583	637	666
Belgien	499	(2.4)	100	(1.5)	323	360	429	573	623	650
Portugal	498	(2.7)	92	(1.1)	339	374	436	564	614	641
Vereinigtes Königreich	498	(2.8)	97	(1.1)	336	372	432	565	621	653
Vereinigte Staaten	497	(3.4)	100	(1.6)	326	364	430	568	624	655
Spanien	496	(2.4)	87	(1.4)	343	379	438	558	603	629
Schweiz	492	(3.0)	98	(1.7)	322	360	426	563	614	643
Lettland	488	(1.8)	85	(1.5)	341	374	431	548	595	621
Tschechische Republik	487	(2.6)	100	(1.7)	315	352	418	559	614	645
Österreich	485	(2.8)	101	(1.5)	308	347	417	559	611	641
Italien	485	(2.7)	94	(1.6)	323	359	421	552	602	631
Island	482	(2.0)	99	(1.7)	310	350	417	552	607	638
Luxemburg	481	(1.4)	107	(1.0)	299	336	405	561	616	647
Israel	479	(3.8)	113	(2.0)	284	326	401	562	621	655
Ungarn	470	(2.7)	97	(1.7)	306	338	399	541	593	620
Griechenland	467	(4.3)	98	(2.4)	296	334	400	539	590	618
Chile	459	(2.6)	88	(1.7)	310	342	398	521	572	599
Slowakische Republik	453	(2.8)	104	(1.8)	269	312	382	528	583	613
Türkei	428	(4.0)	82	(2.0)	291	322	372	487	535	561
Mexiko	423	(2.6)	78	(1.5)	292	321	370	478	523	549
OECD-Durchschnitt	493	(0.5)	96	(0.3)	326	364	428	561	613	642
OECD-Partnerstaaten										
Singapur	535	(1.6)	99	(1.1)	362	400	470	607	657	686
Hongkong (China)	527	(2.7)	86	(1.5)	372	412	473	587	632	656
Macau (China)	509	(1.3)	82	(1.1)	365	399	456	566	610	635
Chinesisch Taipeh	497	(2.5)	93	(1.7)	331	371	437	563	611	638
Russische Föderation	495	(3.1)	87	(1.4)	350	381	434	556	608	637
B-S-J-G (China)*	494	(5.1)	109	(2.9)	304	346	420	573	630	661
Kroatien	487	(2.7)	91	(1.6)	334	367	424	553	603	632
Vietnam	487	(3.7)	73	(2.0)	367	393	438	537	580	605
Litauen	472	(2.7)	94	(1.5)	312	347	407	541	593	622
Malta	447	(1.8)	121	(1.5)	236	284	366	533	595	631
Uruguay	437	(2.5)	97	(1.6)	280	311	368	504	563	597
Rumänien	434	(4.1)	95	(2.1)	276	310	370	499	555	588
Vereinigte Arabische Emirate	434	(2.9)	106	(1.4)	258	295	359	509	572	605
Bulgarien	432	(5.0)	115	(2.6)	241	277	347	517	578	611
Costa Rica	427	(2.6)	79	(1.6)	298	326	374	480	530	560
Republik Trinidad und Tobago	427	(1.5)	104	(1.3)	256	291	353	502	561	596
Republik Montenegro	427	(1.6)	94	(1.2)	271	304	361	493	549	581
Kolumbien	425	(2.9)	90	(1.5)	278	308	361	489	542	572
Republik Moldau	416	(2.5)	98	(1.5)	253	289	349	485	541	574
Thailand	409	(3.3)	80	(1.7)	281	308	354	463	514	543
Jordanien	408	(2.9)	94	(1.8)	241	281	348	475	522	549
Brasilien	407	(2.8)	100	(1.5)	247	279	336	477	539	576
Albanien	405	(4.1)	97	(1.8)	244	279	340	472	528	561
Katar	402	(1.0)	111	(1.0)	221	256	321	483	547	581
Georgien	401	(3.0)	104	(1.8)	226	266	332	474	533	568
Peru	398	(2.9)	89	(1.6)	253	281	333	462	514	543
Indonesien	397	(2.9)	76	(1.8)	272	300	346	448	495	522
Tunesien	361	(3.1)	82	(1.9)	228	257	305	416	467	496
Dominikanische Republik	358	(3.1)	85	(1.9)	226	250	297	416	471	503
Republik Mazedonien	352	(1.4)	99	(1.2)	187	222	284	421	480	513
Algerien	350	(3.0)	73	(1.6)	232	258	301	397	443	472
Kosovo	347	(1.6)	78	(1.1)	215	243	294	403	447	471
Libanesische Republik	347	(4.4)	115	(2.6)	167	203	265	426	503	546

*B-S-J-G (China) bezieht sich auf vier Provinzen in China, die an der PISA-Studie teilgenommen haben: Peking, Shanghai, Jiangsu und Guangdong.

signifikant über dem OECD-Durchschnitt

nicht signifikant verschieden vom OECD-Durchschnitt

signifikant unter dem OECD-Durchschnitt

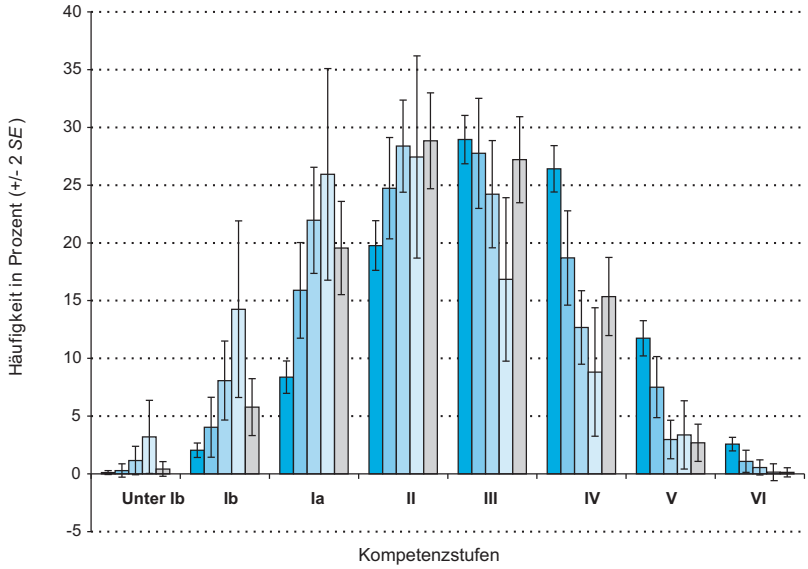
© Waxmann Verlag GmbH

Tabelle A7: Mittlere Kompetenz von Jugendlichen mit Zuwanderungshintergrund für Naturwissenschaften, Mathematik und Lesen

	Ohne Zuwanderungs-hintergrund		**Mit Zuwanderungshintergrund**							
			Insgesamt Zuwanderung		Ein Elternteil im Ausland geboren		Zweite Generation		Erste Generation	
	%	(*SE*)	*%*	(*SE*)	*%*	(*SE*)	*%*	(*SE*)	*%*	(*SE*)
Naturwissenschaften	532	(2.7)	**471**	(5.3)	**497**	(5.1)	**461**	(5.9)	**433**	(10.7)
Mathematik	524	(3.1)	**476**	(4.8)	**493**	(4.8)	**471**	(5.7)	**445**	(10.6)
Lesen	530	(3.2)	**481**	(6.2)	**502**	(5.8)	**478**	(7.0)	**429**	(12.9)

fett: signifikante Unterschiede zu Schülerinnen und Schülern ohne Zuwanderungshintergrund

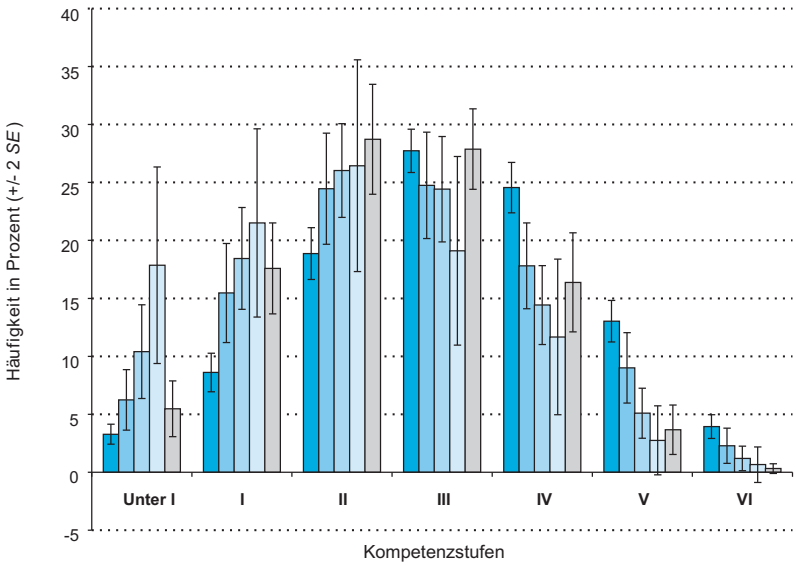
© Waxmann Verlag GmbH



	Unter Ib	Ib	Ia	II	III	IV	V	VI
	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)
Ohne Zuwanderungshintergrund	0.1 (0.1)	2.0 (0.3)	8.4 (0.7)	19.8 (1.1)	28.9 (1.0)	26.4 (1.0)	11.7 (0.8)	2.6 (0.3)
Generationsstatus								
Ein Elternteil im Ausland geboren	0.3 (0.1)	4.0 (1.3)	15.9 (2.1)	24.7 (2.2)	27.8 (2.4)	18.7 (2.0)	7.5 (1.3)	1.1 (0.5)
Zweite Generation	1.2 (0.3)	8.1 (1.7)	22.0 (2.3)	28.4 (2.0)	24.2 (2.3)	12.7 (1.6)	3.0 (0.8)	0.6 (0.3)
Erste Generation	3.2 (1.6)	14.3 (3.8)	25.9 (4.6)	27.4 (4.4)	16.8 (3.5)	8.8 (2.8)	3.4 (1.5)	0.1 (0.4)
Nicht zuzuordnen	0.4 (0.3)	5.8 (1.2)	19.6 (2.0)	28.8 (2.1)	27.2 (1.9)	15.4 (1.7)	2.7 (0.8)	0.1 (0.2)

Abbildung A1: Kompetenzstufen von fünfzehnjährigen Schülerinnen und Schülern mit Zuwanderungshintergrund für Naturwissenschaften

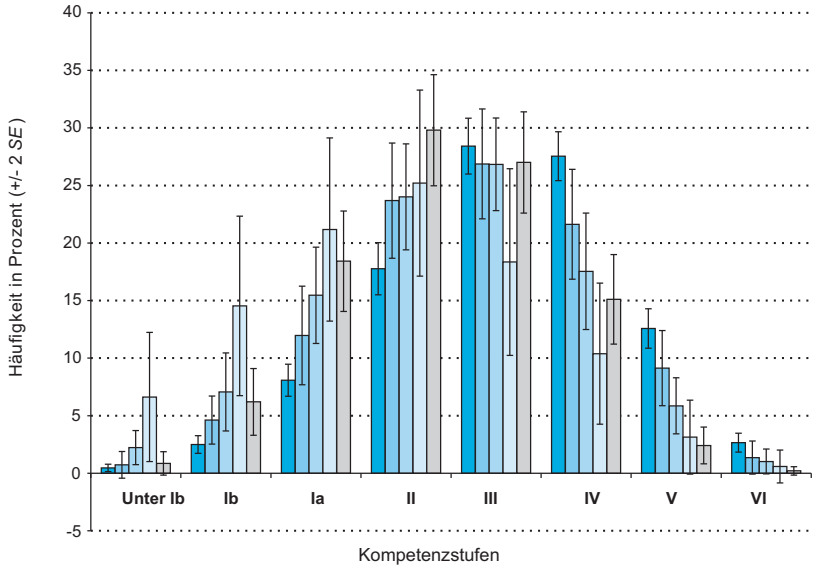
© Waxmann Verlag GmbH



	Unter I	I	II	III	IV	V	VI
	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)
Ohne Zuwanderungshintergrund	3.3 (0.4)	8.6 (0.8)	18.9 (1.1)	27.7 (0.9)	24.6 (1.1)	13.0 (0.9)	3.9 (0.5)
Generationsstatus							
Ein Elternteil im Ausland geboren	6.2 (1.3)	15.5 (2.1)	24.5 (2.4)	24.8 (2.3)	17.8 (1.9)	9.0 (1.5)	2.3 (0.8)
Zweite Generation	10.4 (2.0)	18.4 (2.2)	26.0 (2.0)	24.4 (2.3)	14.4 (1.7)	5.1 (1.1)	1.2 (0.5)
Erste Generation	17.9 (4.2)	21.5 (4.1)	26.4 (4.6)	19.1 (4.1)	11.7 (3.4)	2.8 (1.5)	0.7 (0.8)
Nicht zuzuordnen	5.5 (1.2)	17.6 (2.0)	28.7 (2.4)	27.9 (1.7)	16.4 (2.1)	3.7 (1.1)	0.3 (0.2)

Abbildung A2: Kompetenzstufen von fünfzehnjährigen Schülerinnen und Schülern mit Zuwanderungshintergrund für Mathematik

© Waxmann Verlag GmbH



	Unter Ib	Ib	Ia	II	III	IV	V	VI
	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)	% (*SE*)
Ohne Zuwanderungshintergrund	0.5 (0.2)	2.5 (0.4)	8.1 (0.7)	17.8 (1.1)	28.4 (1.2)	27.5 (1.1)	12.6 (0.9)	2.7 (0.4)
Generationsstatus								
Ein Elternteil im Ausland geboren	0.7 (0.6)	4.6 (1.0)	12.0 (2.1)	23.7 (2.5)	26.9 (2.4)	21.6 (2.4)	9.1 (1.6)	1.4 (0.7)
Zweite Generation	2.2 (0.7)	7.1 (1.7)	15.5 (2.1)	24.0 (2.3)	26.8 (2.0)	17.5 (2.5)	5.9 (1.2)	1.0 (0.5)
Erste Generation	6.6 (2.8)	14.5 (3.9)	21.2 (4.0)	25.2 (4.0)	18.3 (4.1)	10.4 (3.1)	3.1 (1.6)	0.6 (0.7)
Nicht zuzuordnen	0.8 (0.5)	6.2 (1.5)	18.4 (2.2)	29.8 (2.4)	27.0 (2.2)	15.1 (1.9)	2.4 (0.8)	0.2 (0.2)

Abbildung A3: Kompetenzstufen von fünfzehnjährigen Schülerinnen und Schülern mit Zuwanderungshintergrund für Lesen

© Waxmann Verlag GmbH

Anhang B (zu Kapitel 12) – Übersicht zu Testformen, Item-Clustern und Booklet-Design

Die Zusammenstellung der unterschiedlichen Testformen (Booklets) aus den Item-Clustern und deren Zuweisung zu den einzelnen Schülerinnen und Schülern im Rahmen der computerbasierten Testung erfolgte bei PISA 2015 in zwei Schritten.

Wie Tabelle B1 zeigt, existieren insgesamt 66 verschiedenen Basis-Testformen (Booklets), welche jeweils insgesamt vier Item-Cluster umfassen. Zwei dieser Item-Cluster enthalten dabei Items zu *naturwissenschaftlicher Kompetenz* und die anderen zwei Item-Cluster enthalten entweder Items zu *Lesekompetenz*, *Mathematikkompetenz* oder aber zum Kompetenzbereich *Problemlösen im Team*. Die in Tabelle B1 dargestellten sechs Basis-Testformgruppen werden in unterschiedlichen Anteilen den Schülerinnen und Schüler der gesamten Stichprobe zugeordnet. Diese Darstellung in Tabelle B1 indiziert für die Item-Cluster des Kompetenzbereiches Naturwissenschaften zunächst lediglich deren (relative) Position („**S**“), ohne dabei zu spezifizieren, welche konkreten Item-Cluster für Naturwissenschaften die beiden Positionen besetzen.

© Waxmann Verlag GmbH

Tabelle B1: Basis-Testformen nach Clustern

Anteil in Stichprobe	Testform (Basis)	Cluster 1	Cluster 2	Cluster 3	Cluster 4
33%	31	S	S	R01	R02
	32	S	S	R02	R03
	33	S	S	R03	R04
	34	S	S	R04	R05
	35	S	S	R05	R06ab
	36	S	S	R06ab	R01
	37	R01	R03	S	S
	38	R02	R04	S	S
	39	R03	R05	S	S
	40	R04	R06ab	S	S
	41	R05	R01	S	S
	42	R06ab	R02	S	S
33%	43	S	S	M01	M02
	44	S	S	M02	M03
	45	S	S	M03	M04
	46	S	S	M04	M05
	47	S	S	M05	M06ab
	48	S	S	M06ab	M01
	49	M01	M03	S	S
	50	M02	M04	S	S
	51	M03	M05	S	S
	52	M04	M06ab	S	S
	53	M05	M01	S	S
	54	M06ab	M02	S	S
4%	55	S	S	M01	R01
	56	S	S	R02	M02
	57	S	S	M03	R03
	58	S	S	R04	M04
	59	S	S	M05	R05
	60	S	S	R06ab	M06ab
	61	R01	M01	S	S
	62	M02	R02	S	S
	63	R03	M03	S	S
	64	M04	R04	S	S
	65	R05	M05	S	S
	66	M06ab	R06ab	S	S
4%	67	S	S	C01	M01
	68	S	S	M02	C02
	69	S	S	C03	M03
	70	S	S	M04	C03
	71	S	S	C02	M05
	72	S	S	M06ab	C01
	73	M01	C02	S	S
	74	C03	M02	S	S
	75	M03	C01	S	S
	76	C01	M04	S	S
	77	M05	C03	S	S
	78	C02	M06ab	S	S
4%	79	S	S	R01	C01
	80	S	S	C02	R02
	81	S	S	R03	C03
	82	S	S	C03	R04
	83	S	S	R05	C02
	84	S	S	C01	R06ab
	85	C02	R01	S	S
	86	R02	C03	S	S
	87	C01	R03	S	S
	88	R04	C01	S	S
	89	C03	R05	S	S
	90	R06ab	C02	S	S
22%	91	S	S	C01	C02
	92	S	S	C02	C03
	93	S	S	C03	C01
	94	C02	C01	S	S
	95	C03	C02	S	S
	96	C01	C03	S	S

Anmerkungen: Mathematik Trend-Aufgaben (M01–M06); Lesen Trend-Aufgaben (R01–R06); Naturwissenschaften Trend-Aufgaben und neue Aufgaben (S); Problemlösen im Team, neue Aufgaben (C01–C06); Standardaufgaben (a); einfache Aufgaben (b).

© Waxmann Verlag GmbH

Die Zuweisung der 66 Basis-Testformen aus Tabelle B1 zu den Schülerinnen und Schülern erfolgt nun im ersten Schritt durch die zufällige Ziehung einer zweistelligen Zahl (im Bereich 31 bis 96) für jede einzelne Schülerin, jeden einzelnen Schüler. Als Restriktion dieser Zufallsziehung werden dabei die vorgegebenen Anteile (im Bereich von 33% bis 4%) der Basis-Testformgruppen in der gesamten Stichprobe berücksichtigt.

Im zweiten Schritt können nun die in Tabelle B1 durch den Buchstaben „**S**“ indizierten zwei Positionen der Item-Cluster jeweils mit 36 unterschiedlichen Clusterkombinationen für den Bereich Naturwissenschaften besetzt werden (vgl. Tabelle B2). Die in Tabelle B2 dargestellten Clusterkombinationen für den Kompetenzbereich Naturwissenschaften bestehen dabei entweder aus (1) neuen Aufgaben und Trend-Aufgaben, (2) nur neuen Aufgaben oder (3) nur Trend-Aufgaben.

Für die Besetzung dieser in Tabelle B1 mit „**S**“ indizierten Positionen wird nun für jede Schülerin, jeden Schüler eine (ganzzahlige gleichverteilte) Zufallszahl (K) im Wertebereich von $K = 1$ bis $K = 6$ gezogen.

Tabelle B2: 36 Clusterkombinationen für Naturwissenschaften

	Clusterkombination			Clusterkombination	
N	**S**	**S**	***N***	**S**	**S**
1	S01	S07	19	S07	S08
2	S01	S10	20	S07	S09
3	S02	S08	21	S07	S11
4	S03	S09	22	S08	S10
5	S03	S12	23	S08	S12
6	S04	S07	24	S09	S08
7	S04	S10	25	S09	S11
8	S05	S11	26	S10	S07
9	S06	S12	27	S10	S09
10	S07	S06	28	S10	S12
11	S08	S01	29	S11	S08
12	S08	S05	30	S11	S10
13	S09	S02	31	S12	S07
14	S09	S06	32	S12	S09
15	S10	S03	33	S12	S11
16	S11	S02	34	S02	S04
17	S11	S04	35	S05	S01
18	S12	S05	36	S06	S03

Anmerkungen: Trend-Aufgaben: S01–S06; neue Aufgaben S07–S12; Kombinationen 1–18: Trend-Aufgaben und neue Aufgaben; Kombinationen 19–33: nur neue Aufgaben; Kombinationen 34-36: nur Trend-Aufgaben.

© Waxmann Verlag GmbH

In Verbindung mit der im ersten Schritt gezogenen zweistelligen Zufallszahl kann nun aus der Tabelle B3 die entsprechende Clusterkombination (aus Tabelle B2) für die in Tabelle B1 mit dem Buchstaben „**S**“ indizierten Clusterpositionen herausgesucht werden.

Zur Verdeutlichung dieses zweistufigen Prozesses der Zuweisung von Basis-Testformen und Item-Clustern für Naturwissenschaften zu den einzelnen Schülerinnen und Schülern sei das folgende Beispiel dargestellt:

> Dem Schüler Peter wird im ersten Schritt die Basis-Testform „**37**“ zugewiesen. Sein PISA-Test enthält daher an den ersten beiden Clusterpositionen die beiden Lesecluster **R01** (Position 1) und **R03** (Position 2). Im zweiten Schritt wird Peter die Zufallszahl $\boldsymbol{K} = \mathbf{4}$ zugewiesen. Aus Tabelle B3 ist nun in Zeile „**37**“ und Spalte „**4**“ abzulesen, dass Peter für die beiden verbleibenden Clusterpositionen seiner Basis-Testform die Item-Clusterkombination „**19**“ für Naturwissenschaften erhält. Wie in Tabelle B2 zu erkennen ist, besteht diese Clusterkombination „**19**“ aus den Clustern „**S07**“ und „**S08**“. Da Peter die Basis-Testform „**37**“ erhalten hat, erscheint für ihn der Cluster „**S07**“ an Position 3 und der Cluster „**S08**“ an Position 4 seiner so individualisierten Testform.

© Waxmann Verlag GmbH

Tabelle B3: Blockdesign zur Zuordnung der Aufgaben aus dem Bereich Naturwissenschaften zu den Basis-Testformen

Testform (Basis)	Zufallszahl (*K*)						Testform (Basis)	Zufallszahl (*K*)					
	1	2	3	4	5	6		1	2	3	4	5	6
31	1	13	6	9	22	25	64	1	13	6	9	22	25
32	2	16	12	10	31	32	65	2	16	12	10	31	32
33	11	5	17	14	26	29	66	11	5	17	14	26	29
34	35	4	7	19	23	30	67	1	13	6	9	22	25
35	34	15	8	20	24	28	68	2	16	12	10	31	32
36	3	36	18	21	27	33	69	11	5	17	14	26	29
37	35	4	7	19	23	30	70	35	4	7	19	23	30
38	34	15	8	20	24	28	71	34	15	8	20	24	28
39	3	36	18	21	27	33	72	3	36	18	21	27	33
40	1	13	6	9	22	25	73	35	4	7	19	23	30
41	2	16	12	10	31	32	74	34	15	8	20	24	28
42	11	5	17	14	26	29	75	3	36	18	21	27	33
43	1	13	6	9	22	25	76	1	13	6	9	22	25
44	2	16	12	10	31	32	77	2	16	12	10	31	32
45	11	5	17	14	26	29	78	11	5	17	14	26	29
46	35	4	7	19	23	30	79	1	13	6	9	22	25
47	34	15	8	20	24	28	80	2	16	12	10	31	32
48	3	36	18	21	27	33	81	11	5	17	14	26	29
49	35	4	7	19	23	30	82	35	4	7	19	23	30
50	34	15	8	20	24	28	83	34	15	8	20	24	28
51	3	36	18	21	27	33	84	3	36	18	21	27	33
52	1	13	6	9	22	25	85	35	4	7	19	23	30
53	2	16	12	10	31	32	86	34	15	8	20	24	28
54	11	5	17	14	26	29	87	3	36	18	21	27	33
55	1	13	6	9	22	25	88	1	13	6	9	22	25
56	2	16	12	10	31	32	89	2	16	12	10	31	32
57	11	5	17	14	26	29	90	11	5	17	14	26	29
58	35	4	7	19	23	30	91	1	13	6	9	22	25
59	34	15	8	20	24	28	92	2	16	12	10	31	32
60	3	36	18	21	27	33	93	11	5	17	14	26	29
61	35	4	7	19	23	30	94	35	4	7	19	23	30
62	34	15	8	20	24	28	95	34	15	8	20	24	28
63	3	36	18	21	27	33	96	3	36	18	21	27	33

© Waxmann Verlag GmbH

Tabelle B4: Zuordnung der neuen Aufgaben für den Kompetenzbereich Naturwissenschaften zu Clustern und Subdomänen

Cluster	ID Item	Bezeichnung	Antwort-format (CBA)	Subdomäne
S07	S627Q01	Autoreifen	MEC	Phänomene naturwissenschaftlich erklären
S07	S627Q03	Autoreifen	MKC	Phänomene naturwissenschaftlich erklären
S07	S627Q04	Autoreifen	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S07	S635Q01	Rettet die Fische	MKC	Phänomene naturwissenschaftlich erklären
S07	S635Q02	Rettet die Fische	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S07	S635Q03	Rettet die Fische	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S07	S635Q04	Rettet die Fische	OC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S07	S635Q05	Rettet die Fische	OM	Phänomene naturwissenschaftlich erklären
S07	S603Q01	Elefanten und Akazienbäume	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S07	S603Q02	Elefanten und Akazienbäume	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S07	S603Q03	Elefanten und Akazienbäume	MEC	Phänomene naturwissenschaftlich erklären
S07	S603Q04	Elefanten und Akazienbäume	MEC	Phänomene naturwissenschaftlich erklären
S07	S603Q05	Elefanten und Akazienbäume	MEC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S07	S602Q01	Städtischer Wärmeinsel-Effekt	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S07	S602Q02	Städtischer Wärmeinsel-Effekt	MKC	Phänomene naturwissenschaftlich erklären
S07	S602Q03	Städtischer Wärmeinsel-Effekt	OM	Phänomene naturwissenschaftlich erklären
S07	S602Q04	Städtischer Wärmeinsel-Effekt	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S08	S607Q01	Vögel und Raupen	MEC	Phänomene naturwissenschaftlich erklären
S08	S607Q02	Vögel und Raupen	MKC	Phänomene naturwissenschaftlich erklären
S08	S607Q03	Vögel und Raupen	OM	Phänomene naturwissenschaftlich erklären
S08	S646Q01	Nanopartikel	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S08	S646Q02	Nanopartikel	OC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S08	S646Q03	Nanopartikel	MEC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S08	S646Q04	Nanopartikel	OM	Phänomene naturwissenschaftlich erklären
S08	S646Q05	Nanopartikel	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S08	S608Q01	Ammoniten	MKC	Phänomene naturwissenschaftlich erklären
S08	S608Q02	Ammoniten	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S08	S608Q03	Ammoniten	MEC	Phänomene naturwissenschaftlich erklären

© Waxmann Verlag GmbH

Cluster	ID Item	Bezeichnung	Antwort-format (CBA)	Subdomäne
S08	S608Q04	Ammoniten	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S08	S605Q01	Geothermische Energie	MKC	Phänomene naturwissenschaftlich erklären
S08	S605Q02	Geothermische Energie	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S08	S605Q03	Geothermische Energie	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S08	S605Q04	Geothermische Energie	OM	Phänomene naturwissenschaftlich erklären
S09	S649Q01	Wetterballon	MEC	Phänomene naturwissenschaftlich erklären
S09	S649Q02	Wetterballon	OM	Phänomene naturwissenschaftlich erklären
S09	S649Q03	Wetterballon	MKC	Phänomene naturwissenschaftlich erklären
S09	S649Q04	Wetterballon	MEC	Phänomene naturwissenschaftlich erklären
S09	S634Q01	Impfung und Ausbreitung von Krankheiten	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S09	S634Q02	Impfung und Ausbreitung von Krankheiten	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S09	S634Q03	Impfung und Ausbreitung von Krankheiten	OM	Phänomene naturwissenschaftlich erklären
S09	S634Q05	Impfung und Ausbreitung von Krankheiten	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S09	S634Q04	Impfung und Ausbreitung von Krankheiten	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S09	S620Q01	Tornados	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S09	S620Q02	Tornados	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S09	S620Q04	Tornados	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S09	S638Q01	Ölteppiche	MKC	Phänomene naturwissenschaftlich erklären
S09	S638Q02	Ölteppiche	MKC	Phänomene naturwissenschaftlich erklären
S09	S638Q04	Ölteppiche	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S09	S638Q05	Ölteppiche	OM	Phänomene naturwissenschaftlich erklären
S10	S625Q01	Flächenbrände und das Verbrennungsdreieck	OM	Phänomene naturwissenschaftlich erklären
S10	S625Q02	Flächenbrände und das Verbrennungsdreieck	MEC	Phänomene naturwissenschaftlich erklären
S10	S625Q03	Flächenbrände und das Verbrennungsdreieck	MKC	Phänomene naturwissenschaftlich erklären
S10	S615Q07	Tsunamis verstehen	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S10	S615Q01	Tsunamis verstehen	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S10	S615Q02	Tsunamis verstehen	OC	Daten und Evidenz naturwissenschaftlich interpretieren
S10	S615Q05	Tsunamis verstehen	MKC	Phänomene naturwissenschaftlich erklären
S10	S604Q02	Wasser aus Nebel	MKC	Phänomene naturwissenschaftlich erklären

© Waxmann Verlag GmbH

Cluster	ID Item	Bezeichnung	Antwort-format (CBA)	Subdomäne
S10	S604Q04	Wasser aus Nebel	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S10	S645Q01	Kohlendioxid in der Erdatmosphäre	MKC	Phänomene naturwissenschaftlich erklären
S10	S645Q03	Kohlendioxid in der Erdatmosphäre	MKC	Phänomene naturwissenschaftlich erklären
S10	S645Q04	Kohlendioxid in der Erdatmosphäre	OM	Phänomene naturwissenschaftlich erklären
S10	S645Q05	Kohlendioxid in der Erdatmosphäre	OM	Phänomene naturwissenschaftlich erklären
S10	S657Q01	Invasive Arten	MKC	Phänomene naturwissenschaftlich erklären
S10	S657Q02	Invasive Arten	MEC	Phänomene naturwissenschaftlich erklären
S10	S657Q03	Invasive Arten	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S10	S657Q04	Invasive Arten	OM	Phänomene naturwissenschaftlich erklären
S11	S656Q01	Vogelzug	MEC	Phänomene naturwissenschaftlich erklären
S11	S656Q02	Vogelzug	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S11	S656Q04	Vogelzug	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S11	S643Q03	Lampen im Vergleich	OM	Phänomene naturwissenschaftlich erklären
S11	S643Q01	Lampen im Vergleich	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S11	S643Q02	Lampen im Vergleich	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S11	S643Q04	Lampen im Vergleich	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S11	S643Q05	Lampen im Vergleich	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S11	S629Q01	Solarkocher	OM	Phänomene naturwissenschaftlich erklären
S11	S629Q02	Solarkocher	MKC	Phänomene naturwissenschaftlich erklären
S11	S629Q03	Solarkocher	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S11	S629Q04	Solarkocher	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S11	S648Q01	Habitable Zone	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S11	S648Q02	Habitable Zone	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S11	S648Q03	Habitable Zone	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S11	S648Q05	Habitable Zone	OM	Phänomene naturwissenschaftlich erklären
S12	S641Q01	Meteoriten und Krater	MEC	Phänomene naturwissenschaftlich erklären
S12	S641Q02	Meteoriten und Krater	MKC	Phänomene naturwissenschaftlich erklären
S12	S641Q03	Meteoriten und Krater	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S12	S641Q04	Meteoriten und Krater	MKC	Daten und Evidenz naturwissenschaftlich interpretieren

© Waxmann Verlag GmbH

Cluster	ID Item	Bezeichnung	Antwort-format (CBA)	Subdomäne
S12	S637Q01	Untersuchung von Hangflächen	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S12	S637Q02	Untersuchung von Hangflächen	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S12	S637Q05	Untersuchung von Hangflächen	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S12	S601Q01	Nachhaltige Fischzucht	MKC	Phänomene naturwissenschaftlich erklären
S12	S601Q02	Nachhaltige Fischzucht	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S12	S601Q04	Nachhaltige Fischzucht	MEC	Phänomene naturwissenschaftlich erklären
S12	S610Q01	Gedankengesteuerte Roboter	OM	Phänomene naturwissenschaftlich erklären
S12	S610Q02	Gedankengesteuerte Roboter	MEC	Phänomene naturwissenschaftlich erklären
S12	S610Q04	Gedankengesteuerte Roboter	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S12	S626Q01	Schall im Lebensraum Meer	MEC	Phänomene naturwissenschaftlich erklären
S12	S626Q02	Schall im Lebensraum Meer	MEC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S12	S626Q03	Schall im Lebensraum Meer	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S12	S626Q04	Schall im Lebensraum Meer	OM	Daten und Evidenz naturwissenschaftlich interpretieren

Anmerkungen: **MEC**: Mehrfachwahlaufgabe (einfach) computerkodiert; **OM**: Offene Antwort (menschlicher Kodierer); **MKC**: Mehrfachwahlaufgabe (komplex) computerkodiert; **OC**: Offene Antwort (Computercodiert).

© Waxmann Verlag GmbH

Tabelle B5: Zuordnung der Trend-Aufgaben für den Kompetenzbereich Naturwissenschaften zu Clustern und Subdomänen

Cluster	ID Item	Bezeichnung	Antwort-format (CBA)	Subdomäne
S01	S269Q01	Temperatur der Erde	OM	Phänomene naturwissenschaftlich erklären
S01	S269Q03	Temperatur der Erde	OM	Phänomene naturwissenschaftlich erklären
S01	S269Q04	Temperatur der Erde	MKC	Phänomene naturwissenschaftlich erklären
S01	S408Q01	Flughafer	MEC	Phänomene naturwissenschaftlich erklären
S01	S408Q03	Flughafer	OM	Phänomene naturwissenschaftlich erklären
S01	S408Q04	Flughafer	MKC	Phänomene naturwissenschaftlich erklären
S01	S408Q05	Flughafer	MEC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S01	S521Q02	Kochen im Freien	MEC	Phänomene naturwissenschaftlich erklären
S01	S521Q06	Kochen im Freien	MEC	Phänomene naturwissenschaftlich erklären
S01	S519Q01	Airbags	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S01	S519Q02	Airbags	MKC	Phänomene naturwissenschaftlich erklären
S01	S519Q03	Airbags	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S01	S527Q01	Das Aussterben der Dinosaurier	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S01	S527Q03	Das Aussterben der Dinosaurier	MKC	Phänomene naturwissenschaftlich erklären
S01	S527Q04	Das Aussterben der Dinosaurier	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S01	S466Q01	Waldbrände	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S01	S466Q07	Waldbrände	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S01	S466Q05	Waldbrände	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S02	S326Q01	Milch	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S02	S326Q02	Milch	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S02	S326Q03	Milch	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S02	S326Q04	Milch	MKC	Phänomene naturwissenschaftlich erklären
S02	S256Q01	Löffel	MEC	Phänomene naturwissenschaftlich erklären
S02	S478Q01	Antibiotika	MEC	Phänomene naturwissenschaftlich erklären
S02	S478Q02	Antibiotika	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S02	S478Q03	Antibiotika	MKC	Phänomene naturwissenschaftlich erklären
S02	S413Q06	Kunststoffzeitalter	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S02	S413Q04	Kunststoffzeitalter	MKC	Daten und Evidenz naturwissenschaftlich interpretieren

© Waxmann Verlag GmbH

Cluster	ID Item	Bezeichnung	Antwort-format (CBA)	Subdomäne
S02	S413Q05	Kunststoffzeitalter	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S02	S498Q02	Versuch zur Verdauung	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S02	S498Q03	Versuch zur Verdauung	MEC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S02	S498Q04	Versuch zur Verdauung	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S02	S425Q03	Die Pinguin-Insel	OM	Phänomene naturwissenschaftlich erklären
S02	S425Q05	Die Pinguin-Insel	MEC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S02	S425Q02	Die Pinguin-Insel	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S02	S425Q04	Die Pinguin-Insel	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S03	S465Q01	Klima-Unterschiede	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S03	S465Q02	Klima-Unterschiede	MEC	Phänomene naturwissenschaftlich erklären
S03	S465Q04	Klima-Unterschiede	MEC	Phänomene naturwissenschaftlich erklären
S03	S131Q02	Schwingungen	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S03	S131Q04	Schwingungen	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S03	S428Q01	Bakterien in der Milch	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S03	S428Q03	Bakterien in der Milch	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S03	S428Q05	Bakterien in der Milch	OM	Phänomene naturwissenschaftlich erklären
S03	S514Q02	Bauprojekte und Naturkatastrophen	OM	Phänomene naturwissenschaftlich erklären
S03	S514Q03	Bauprojekte und Naturkatastrophen	OM	Phänomene naturwissenschaftlich erklären
S03	S514Q04	Bauprojekte und Naturkatastrophen	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S03	S438Q01	Grüne Parks	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S03	S438Q02	Grüne Parks	MEC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S03	S438Q03	Grüne Parks	OM	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S03	S415Q07	Solarzellen	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S03	S415Q02	Solarzellen	MEC	Phänomene naturwissenschaftlich erklären
S03	S415Q08	Solarzellen	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S04	S476Q01	Herzoperation	MEC	Phänomene naturwissenschaftlich erklären
S04	S476Q02	Herzoperation	MEC	Phänomene naturwissenschaftlich erklären

© Waxmann Verlag GmbH

Cluster	ID Item	Bezeichnung	Antwort-format (CBA)	Subdomäne
S04	S476Q03	Herzoperation	MEC	Phänomene naturwissenschaftlich erklären
S04	S495Q04	Strahlentherapie	MKC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S04	S495Q01	Strahlentherapie	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S04	S495Q02	Strahlentherapie	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S04	S495Q03	Strahlentherapie	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S04	S268Q01	Algen	MEC	Naturwissenschaftliche Forschung bewerten & naturwissenschaftliche Untersuchungen planen
S04	S268Q02	Algen	OM	Phänomene naturwissenschaftlich erklären
S04	S268Q06	Algen	MEC	Phänomene naturwissenschaftlich erklären
S05	S524Q06	Herstellung von Penicillin	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S05	S524Q07	Herstellung von Penicillin	OM	Phänomene naturwissenschaftlich erklären
S05	S510Q01	Magnetschwebebahnen	MKC	Phänomene naturwissenschaftlich erklären
S05	S510Q04	Magnetschwebebahnen	OM	Phänomene naturwissenschaftlich erklären
S05	S437Q01	Feuer löschen	MEC	Phänomene naturwissenschaftlich erklären
S05	S437Q03	Feuer löschen	MEC	Phänomene naturwissenschaftlich erklären
S05	S437Q04	Feuer löschen	MEC	Phänomene naturwissenschaftlich erklären
S05	S437Q06	Feuer löschen	OM	Phänomene naturwissenschaftlich erklären
S05	S304Q01	Wasser	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S05	S304Q02	Wasser	MEC	Phänomene naturwissenschaftlich erklären
S05	S304Q03a	Wasser	OM	Phänomene naturwissenschaftlich erklären
S05	S304Q03b	Wasser	OM	Phänomene naturwissenschaftlich erklären
S05	S416Q01	Der Mond	OM	Daten und Evidenz naturwissenschaftlich interpretieren
S06	S458Q01	Die Eismumie	OM	Phänomene naturwissenschaftlich erklären
S06	S458Q02	Die Eismumie	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S06	S421Q01	Groß und Klein	OC	Phänomene naturwissenschaftlich erklären
S06	S421Q02	Groß und Klein	OC	Phänomene naturwissenschaftlich erklären
S06	S421Q03	Groß und Klein	OC	Phänomene naturwissenschaftlich erklären
S06	S252Q01	Südrainea	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S06	S252Q02	Südrainea	MEC	Daten und Evidenz naturwissenschaftlich interpretieren
S06	S252Q03	Südrainea	MKC	Daten und Evidenz naturwissenschaftlich interpretieren
S06	S327Q01	Gezeitenkraft	MKC	Phänomene naturwissenschaftlich erklären

Anmerkungen: **MEC**: Mehrfachwahlaufgabe (einfach) computerkodiert; **OM**: Offene Antwort (menschlicher Kodierer); **MKC**: Mehrfachwahlaufgabe (komplex) computerkodiert; **OC**: Offene Antwort (computerkodiert).

© Waxmann Verlag GmbH

Tabelle B6: Zuordnung der Aufgaben für den Kompetenzbereich Mathematik zu Clustern und Subdomänen

Cluster	ID Item	Bezeichnung	Antwortformat (CBA)	Subdomäne (eindimensional skaliert)
M01	M033Q01	Ansicht eines Zimmers	MEC	Raum und Form
M01	M474Q01	Laufzeit	MEC	Quantität
M01	M155Q02	Bevölkerungspyramiden	OM	Veränderungen und Beziehungen
M01	M155Q01	Bevölkerungspyramiden	MKC	Veränderungen und Beziehungen
M01	M155Q03	Bevölkerungspyramiden	OM	Veränderungen und Beziehungen
M01	M155Q04	Bevölkerungspyramiden	MKC	Veränderungen und Beziehungen
M01	M411Q01	Turmspringen	OC	Quantität
M01	M411Q02	Turmspringen	MEC	Unsicherheit und Daten
M01	M803Q01	Etiketten	OC	Unsicherheit und Daten
M01	M442Q02	Blindenschrift	MKC	Quantität
M01	M462Q01	Die dritte Seite	OM	Raum und Form
M01	M034Q01	Ziegelsteine	OC	Raum und Form
M02	M305Q01	Straßenkarte	MEC	Raum und Form
M02	M496Q01	Geld abheben	MKC	Quantität
M02	M496Q02	Geld abheben	OC	Quantität
M02	M423Q01	Münzwurf	MEC	Unsicherheit und Daten
M02	*M192Q01*	*Behälter*	*Nicht administriert*	*Veränderungen und Beziehungen*
M02	M406Q01	Sportplatz	OM	Raum und Form
M02	M406Q02	Sportplatz	OM	Raum und Form
M02	M603Q01	Zahlenüberprüfung	MKC	Quantität
M02	M571Q01	Stopp das Auto	MEC	Veränderungen und Beziehungen
M02	M564Q01	Sessellift	MEC	Quantität
M02	M564Q02	Sessellift	MEC	Unsicherheit und Daten
M03	M447Q01	Fliesenanordnung	MEC	Raum und Form
M03	M273Q01	Rohrleitungen	MKC	Raum und Form
M03	M408Q01	Lotterien	MKC	Unsicherheit und Daten
M03	M420Q01	Transportmittel	MKC	Unsicherheit und Daten
M03	M446Q01	Die Thermometer-Grille	OC	Veränderungen und Beziehungen
M03	M446Q02	Die Thermometer-Grille	OM	Veränderungen und Beziehungen
M03	M559Q01	Telefongebühren	MEC	Quantität
M03	*M828Q01*	*Kohlenstoffdioxid*	*Nicht administriert*	*Veränderungen und Beziehungen*
M03	M828Q02	Kohlenstoffdioxid	OM	Unsicherheit und Daten
M03	M828Q03	Kohlenstoffdioxid	OC	Quantität
M03	M464Q01	Der Zaun	OC	Raum und Form
M03	M800Q01	Computerspiel	MEC	Quantität

© Waxmann Verlag GmbH

Cluster	ID Item	Bezeichnung	Antwortformat (CBA)	Subdomäne (eindimensional skaliert)
M04	M982Q01	Beschäftigungszahlen	OC	Unsicherheit und Daten
M04	M982Q02	Beschäftigungszahlen	OC	Unsicherheit und Daten
M04	M982Q03	Beschäftigungszahlen	MKC	Unsicherheit und Daten
M04	M982Q04	Beschäftigungszahlen	MEC	Unsicherheit und Daten
M04	M992Q01	Abstandhalter	OC	Raum und Form
M04	M992Q02	Abstandhalter	OC	Raum und Form
M04	M992Q03	Abstandhalter	OM	Veränderungen und Beziehungen
M04	M915Q01	CO2-Steuer	MEC	Unsicherheit und Daten
M04	M915Q02	CO2-Steuer	OC	Veränderungen und Beziehungen
M04	M906Q01	Gelbe Spinnerameisen	MEC	Quantität
M04	M906Q02	Gelbe Spinnerameisen	OM	Quantität
M04	M00KQ02	Rollstuhlbasketball	OM	Raum und Form
M05	M909Q01	Strafe für zu schnelles Fahren	OC	Quantität
M05	M909Q02	Strafe für zu schnelles Fahren	MEC	Quantität
M05	M909Q03	Strafe für zu schnelles Fahren	OC	Veränderungen und Beziehungen
M05	M949Q01	Dachstuhlkonstruktion	MKC	Raum und Form
M05	M949Q02	Dachstuhlkonstruktion	MKC	Raum und Form
M05	M949Q03	Dachstuhlkonstruktion	OM	Raum und Form
M05	M00GQ01	Litfaßsäule	OC	Raum und Form
M05	M955Q01	Migration	OM	Unsicherheit und Daten
M05	M955Q02	Migration	OM	Unsicherheit und Daten
M05	M955Q03	Migration	OC	Unsicherheit und Daten
M05	M998Q02	Fahrradverleih	OM	Veränderungen und Beziehungen
M05	M998Q04	Fahrradverleih	MKC	Veränderungen und Beziehungen
M06A	M905Q01	Tennisbälle	MKC	Quantität
M06A	M905Q02	Tennisbälle	OM	Quantität
M06A	M919Q01	Fanartikel	OC	Quantität
M06A	M919Q02	Fanartikel	OC	Quantität
M06A	M954Q01	Dosierung von Medikamenten	OC	Veränderungen und Beziehungen
M06A	M954Q02	Dosierung von Medikamenten	OM	Veränderungen und Beziehungen
M06A	M954Q04	Dosierung von Medikamenten	OC	Veränderungen und Beziehungen
M06A	M943Q01	Bögen	MEC	Veränderungen und Beziehungen
M06A	M943Q02	Bögen	OC	Raum und Form
M06A	M953Q02	Grippetest	OM	Unsicherheit und Daten
M06A	M953Q03	Grippetest	OC	Unsicherheit und Daten
M06A	M953Q04	Grippetest	OM	Unsicherheit und Daten
M06B*	M948Q01	Teilzeitarbeit [Part Time Work]	MEC	Quantität
M06B*	M948Q02	Teilzeitarbeit [Part Time Work]	OC	Quantität

© Waxmann Verlag GmbH

Cluster	ID Item	Bezeichnung	Antwortformat (CBA)	Subdomäne (eindimensional skaliert)
M06B*	M948Q03	Teilzeitarbeit [Part Time Work]	OC	Quantität
M06B*	M936Q01	Sitze in einem Theater [Seats in a Theatre]	OC	Veränderungen und Beziehungen
M06B*	M936Q02	Sitze in einem Theater [Seats in a Theatre]	OM	Veränderungen und Beziehungen
M06B*	M961Q02	Schokolade [Chocolate]	OM	Veränderungen und Beziehungen
M06B*	M961Q03	Schokolade [Chocolate]	MEC	Veränderungen und Beziehungen
M06B*	M961Q05	Schokolade [Chocolate]	OM	Unsicherheit und Daten
M06B*	M939Q01	Das Rennen [Racing]	MEC	Unsicherheit und Daten
M06B*	M939Q02	Das Rennen [Racing]	MEC	Unsicherheit und Daten
M06B*	M967Q01	Holzeisenbahn [Wooden Train Set]	OC	Raum und Form
M06B*	M967Q03	Holzeisenbahn [Wooden Train Set]	MKC	Raum und Form

Anmerkungen: **MEC**: Mehrfachwahlaufgabe (einfach) computerkodiert; **OM**: Offene Antwort (menschlicher Kodierer); **MKC**: Mehrfachwahlaufgabe (komplex) computerkodiert; **OC**: Offene Antwort (computerkodiert); * in Deutschland wurde das Item-Cluster M06A eingesetzt.

© Waxmann Verlag GmbH

Tabelle B7: Zuordnung der Aufgaben für den Kompetenzbereich Lesen zu Clustern und Subdomänen

Cluster	ID Item	Beschreibung	Antwort-format (CBA)	Subdomäne (eindimensional skaliert)
R01	R219Q01	Bewerbung	OM	Informationen suchen und extrahieren
R01	R219Q01EC	Bewerbung	OM	Kombinieren und Interpretieren
R01	R219Q02	Bewerbung	OM	Reflektieren und Bewerten
R01	R067Q01	Aesop	MEC	Kombinieren und Interpretieren
R01	R067Q04	Aesop	OM	Reflektieren und Bewerten
R01	R067Q05	Aesop	OM	Reflektieren und Bewerten
R01	R102Q04	Hemden	OM	Kombinieren und Interpretieren
R01	R102Q05	Hemden	OM	Kombinieren und Interpretieren
R01	R102Q07	Hemden	MEC	Kombinieren und Interpretieren
R01	R220Q01	Südpol	OC	Informationen suchen und extrahieren
R01	R220Q02	Südpol	MEC	Kombinieren und Interpretieren
R01	R220Q04	Südpol	MEC	Kombinieren und Interpretieren
R01	R220Q05	Südpol	MEC	Kombinieren und Interpretieren
R01	R220Q06	Südpol	MEC	Kombinieren und Interpretieren
R02	R227Q01	Optiker	MEC	Kombinieren und Interpretieren
R02	R227Q02	Optiker	MKC	Informationen suchen und extrahieren
R02	R227Q03	Optiker	OM	Reflektieren und Bewerten
R02	R227Q06	Optiker	OM	Informationen suchen und extrahieren
R02	R111Q01	Schüleraustausch	MEC	Kombinieren und Interpretieren
R02	R111Q02B	Schüleraustausch	OM	Reflektieren und Bewerten
R02	R111Q06B	Schüleraustausch[1]	OM	Reflektieren und Bewerten
R02	R055Q01	Spinnen unter Drogen	MEC	Kombinieren und Interpretieren
R02	R055Q02	Spinnen unter Drogen	OM	Reflektieren und Bewerten
R02	R055Q03	Spinnen unter Drogen	OM	Kombinieren und Interpretieren
R02	R055Q05	Spinnen unter Drogen	OM	Kombinieren und Interpretieren
R02	R104Q01	Telefoninformationen	OC	Informationen suchen und extrahieren
R02	R104Q02	Telefoninformationen	OC	Informationen suchen und extrahieren
R02	R104Q05	Telefoninformationen	OC	Informationen suchen und extrahieren
R03	R420Q02	Die Zukunft der Kinder	OM	Informationen suchen und extrahieren
R03	R420Q10	Die Zukunft der Kinder	OM	Kombinieren und Interpretieren
R03	R420Q06	Die Zukunft der Kinder	OM	Reflektieren und Bewerten
R03	R420Q09	Die Zukunft der Kinder	OM	Informationen suchen und extrahieren
R03	R453Q01	Wie man einen Ferienjob findet	MEC	Kombinieren und Interpretieren
R03	R453Q04	Wie man einen Ferienjob findet	OM	Reflektieren und Bewerten
R03	R453Q05	Wie man einen Ferienjob findet	MKC	Informationen suchen und extrahieren
R03	R453Q06	Wie man einen Ferienjob findet	OM	Reflektieren und Bewerten
R03	R412Q01	Die Sprachen der Welt	MEC	Informationen suchen und extrahieren
R03	R412Q05	Die Sprachen der Welt	MEC	Kombinieren und Interpretieren
R03	R412Q08	Die Sprachen der Welt	OM	Kombinieren und Interpretieren

© Waxmann Verlag GmbH

Cluster	ID Item	Beschreibung	Antwort-format (CBA)	Subdomäne (eindimensional skaliert)
R03	R412Q06	Die Sprachen der Welt	MKC	Kombinieren und Interpretieren
R03	R437Q01	Narziss	MEC	Kombinieren und Interpretieren
R03	R437Q07	Narziss	OM	Kombinieren und Interpretieren
R03	R437Q06	Narziss	MEC	Kombinieren und Interpretieren
R04	R456Q01	Kekse	MEC	Informationen suchen und extrahieren
R04	R456Q02	Kekse	OM	Kombinieren und Interpretieren
R04	R456Q06	Kekse	OM	Kombinieren und Interpretieren
R04	R466Q02	Richtig Arbeiten	OM	Informationen suchen und extrahieren
R04	R466Q03	Richtig Arbeiten	MKC	Kombinieren und Interpretieren
R04	R466Q06	Richtig Arbeiten	OC	Informationen suchen und extrahieren
R04	R446Q03	Stellenangebot	MKC	Informationen suchen und extrahieren
R04	R446Q06	Stellenangebot	OM	Reflektieren und Bewerten
R04	R432Q01	Über ein Buch	OM	Kombinieren und Interpretieren
R04	R432Q05	Über ein Buch	OM	Reflektieren und Bewerten
R04	R432Q06	Über ein Buch	MKC	Kombinieren und Interpretieren
R04	R460Q01	Der Golf von Mexiko	OM	Informationen suchen und extrahieren
R04	R460Q05	Der Golf von Mexiko	MEC	Informationen suchen und extrahieren
R04	R460Q06	Der Golf von Mexiko	MEC	Kombinieren und Interpretieren
R05	R424Q02	Fairer Handel	MKC	Kombinieren und Interpretieren
R05	R424Q03	Fairer Handel	MEC	Reflektieren und Bewerten
R05	R424Q07	Fairer Handel	MEC	Reflektieren und Bewerten
R05	R404Q03	Schlaf	MEC	Kombinieren und Interpretieren
R05	R404Q06	Schlaf	MEC	Kombinieren und Interpretieren
R05	R404Q07	Schlaf	MKC	Kombinieren und Interpretieren
R05	R404Q10A	Schlaf	OM	Reflektieren und Bewerten
R05	R404Q10B	Schlaf	OM	Reflektieren und Bewerten
R05	R406Q01	Kokeshi-Puppen	OM	Kombinieren und Interpretieren
R05	R406Q05	Kokeshi-Puppen	OM	Kombinieren und Interpretieren
R05	R406Q02	Kokeshi-Puppen	OM	Kombinieren und Interpretieren
R05	R455Q02	Schokolade und Gesundheit	OM	Reflektieren und Bewerten
R05	R455Q03	Schokolade und Gesundheit	OM	Informationen suchen und extrahieren
R05	R455Q04	Schokolade und Gesundheit	MEC	Kombinieren und Interpretieren
R05	R455Q05	Schokolade und Gesundheit	MKC	Kombinieren und Interpretieren
R06A	R083Q01	Aufgaben im Haushalt	MEC	Kombinieren und Interpretieren
R06A	R083Q02	Aufgaben im Haushalt	OC	Informationen suchen und extrahieren
R06A	R083Q03	Aufgaben im Haushalt	MEC	Informationen suchen und extrahieren
R06A	R083Q04	Aufgaben im Haushalt	MEC	Kombinieren und Interpretieren
R06A	R442Q02	Galileo	OM	Informationen suchen und extrahieren
R06A	R442Q03	Galileo	OM	Kombinieren und Interpretieren
R06A	R442Q05	Galileo	OM	Reflektieren und Bewerten
R06A	R442Q06	Galileo	OM	Reflektieren und Bewerten

© Waxmann Verlag GmbH

Cluster	ID Item	Beschreibung	Antwort-format (CBA)	Subdomäne (eindimensional skaliert)
R06A	R442Q07	Galileo	MEC	Kombinieren und Interpretieren
R06A	R245Q01	Filmbesprechungen	MKC	Informationen suchen und extrahieren
R06A	R245Q02	Filmbesprechungen	MKC	Kombinieren und Interpretieren
R06A	R101Q01	Nashörner	MEC	Kombinieren und Interpretieren
R06A	R101Q02	Nashörner	MEC	Kombinieren und Interpretieren
R06A	R101Q03	Nashörner	MEC	Reflektieren und Bewerten
R06A	R101Q04	Nashörner	MEC	Kombinieren und Interpretieren
R06A	R101Q05	Nashörner	MEC	Kombinieren und Interpretieren
R06B*	R445Q03	Straße [Road]	MEC	Kombinieren und Interpretieren
R06B*	R445Q01	Straße [Road]	OM	Kombinieren und Interpretieren
R06B*	R445Q04	Straße [Road]	MEC	Kombinieren und Interpretieren
R06B*	R445Q06	Straße [Road]	MEC	Kombinieren und Interpretieren
R06B*	R462Q02	Paketpost [Parcel Post]	OM	Informationen suchen und extrahieren
R06B*	R462Q05	Paketpost [Parcel Post]	OM	Kombinieren und Interpretieren
R06B*	R462Q04	Paketpost [Parcel Post]	MEC	Informationen suchen und extrahieren
R06B*	R435Q02	Staubmilben [Dust Mites]	OC	Informationen suchen und extrahieren
R06B*	R435Q01	Staubmilben [Dust Mites]	MEC	Kombinieren und Interpretieren
R06B*	R435Q05	Staubmilben [Dust Mites]	OM	Reflektieren und Bewerten
R06B*	R435Q08	Staubmilben [Dust Mites]	MKC	Reflektieren und Bewerten
R06B*	R465Q01	Wie man bei der Arbeit überlebt [How to survive at work]	OM	Informationen suchen und extrahieren
R06B*	R465Q02	Wie man bei der Arbeit überlebt [How to survive at work]	OM	Kombinieren und Interpretieren
R06B*	R465Q06	Wie man bei der Arbeit überlebt [How to survive at work]	OM	Reflektieren und Bewerten
R06B*	R465Q05	Wie man bei der Arbeit überlebt [How to survive at work]	OM	Reflektieren und Bewerten

Anmerkungen: **MEC**: Mehrfachwahlaufgabe (einfach) computerkodiert; **OM**: Offene Antwort (menschlicher Kodierer); **MKC**: Mehrfachwahlaufgabe (komplex) computerkodiert; **OC**: Offene Antwort (computerkodiert); * in Deutschland wurde das Item-Cluster R06A eingesetzt und die Titel der Aufgaben nicht übersetzt; [1] Die Mehrfachwahlaufgabe R111Q06A wurde nicht unabhängig bewertet und skaliert und ist daher hier nicht aufgeführt. Die Antworten auf diese Frage wurden bei Bewertung und Skalierung des Items R111Q06B berücksichtigt.

© Waxmann Verlag GmbH

Tabelle B8: Übersicht zu den Ergebnissen der Item-Kalibrierung für den Kompetenzbereich Naturwissenschaften

Staatenspezifische Parameter	ID der Einheit	α	σ	δ_1	δ_2
Unique Item Parameters	CS466Q07S	0.3424	-2.5146		
Unique Item Parameters	CS466Q05S	0.9269	-0.6088		
Unique Item Parameters	CS415Q07S	1.0328	-1.0296		
Unique Item Parameters	DS408Q03C	1.0401	1.1206		
Unique Item Parameters	DS326Q01C	1.1170	0.3759		
Unique Item Parameters	DS438Q03C	1.3219	0.6724		
Unique Item Parameters	CS610Q04S	1.3526	-0.2374		
Unique Item Parameters	DS268Q02C	1.8277	0.7153		
	CS421Q02S	0.4017	1.5239		
	CS252Q02S	0.4583	-0.8238		
	CS615Q05S	0.4783	1.9530		
	CS637Q02S	0.4865	1.9552	0.5116	-0.5116
	CS519Q02S	0.4918	-0.2002		
	CS498Q03S	0.4975	0.4869		
	CS648Q03S	0.5107	-0.4725		
	CS649Q01S	0.5159	1.2546		
	CS465Q04S	0.5615	0.5609		
	CS608Q01S	0.5770	0.7940		
	CS510Q01S	0.5868	0.0849		
	CS521Q02S	0.5995	-0.1214		
	CS649Q04S	0.6296	0.2225		
	CS527Q03S	0.6335	-0.0672		
	CS641Q01S	0.6487	-0.4900		
	CS408Q04S	0.6529	-0.0980		
	CS608Q03S	0.6705	0.2650		
	CS627Q01S	0.6715	0.3764		
	CS498Q02S	0.6777	0.2254		
	CS635Q01S	0.6778	-0.1660	0.1033	-0.1033
	CS605Q03S	0.6833	-0.1822		
	CS437Q03S	0.6839	0.0938		
	CS657Q02S	0.6865	0.4234		
	DS629Q01C	0.7005	-0.1978	-0.3920	0.3920
	CS252Q03S	0.7079	-0.1964		
	CS478Q01S	0.7144	0.2286		
	CS603Q04S	0.7207	-0.2012		
	CS645Q01S	0.7420	-0.1270	-0.3290	0.3290

© Waxmann Verlag GmbH

Staatenspezifische Parameter	ID der Einheit	α	σ	δ_1	δ_2
	DS519Q01C	0.7424	0.3834	-0.2462	0.2462
	CS657Q01S	0.7555	-0.8143		
	CS269Q04S	0.7963	0.6375		
	CS626Q01S	0.7965	-0.3339		
	CS602Q02S	0.8104	0.6132		
	DS607Q03C	0.8137	0.1978	-0.3083	0.3083
	CS625Q02S	0.8297	-0.4315		
	DS519Q03C	0.8380	0.8660		
	CS527Q04S	0.8460	-0.0537		
	CS495Q02S	0.8565	-0.2514		
	CS629Q04S	0.8575	-0.1372		
	CS252Q01S	0.8584	0.0324		
	CS476Q02S	0.8782	-0.7293		
	CS495Q04S	0.8903	0.4236		
	CS495Q01S	0.9031	0.4087		
	CS627Q04S	0.9068	-0.4001		
	CS413Q05S	0.9080	-0.4989		
	DS425Q03C	0.9152	0.2055		
	CS268Q06S	0.9163	0.0884		
	CS629Q02S	0.9176	0.2478		
	CS465Q02S	0.9218	-0.2564		
	CS627Q03S	0.9318	-0.7981		
	CS626Q02S	0.9334	-0.0560		
	CS408Q05S	0.9412	0.2270		
	CS648Q02S	0.9461	0.2769		
	DS498Q04C	0.9518	-0.1754	-1.0223	1.0223
	DS657Q04C	0.9589	0.6401	0.1498	-0.1498
	CS413Q04S	0.9629	0.2495		
	CS408Q01S	0.9647	-0.3385		
	CS478Q03S	0.9744	-0.4318		
	CS524Q06S	0.9766	-0.4214		
	CS604Q02S	0.9789	0.1715		
	DS465Q01C	0.9852	0.1779	-0.0768	0.0768
	CS327Q01S	0.9909	-0.0657		
	CS458Q02S	0.9937	0.0132		
	CS620Q02S	0.9940	0.4138		
	CS256Q01S	1.0000	-1.4119		
	CS476Q01S	1.0000	-0.6170		

© Waxmann Verlag GmbH

Staatenspezifische Parameter	ID der Einheit	α	σ	δ_1	δ_2
	CS268Q01S	1.0000	-0.5800		
	CS466Q01S	1.0000	-0.5779		
	CS425Q05S	1.0000	-0.4525		
	CS304Q02S	1.0000	-0.3232		
	CS437Q04S	1.0000	-0.0474		
	CS425Q02S	1.0000	0.0936		
	CS326Q04S	1.0000	0.8513		
	CS527Q01S	1.0000	1.3042		
	CS656Q01S	1.0041	-0.2278		
	DS458Q01C	1.0323	1.2835		
	CS601Q04S	1.0329	0.2832		
	CS641Q02S	1.0358	-0.5196		
	CS605Q01S	1.0401	0.2309		
	DS656Q02C	1.0463	0.5522		
	CS620Q01S	1.0583	-0.9675		
	DS649Q02C	1.0609	1.2793		
	CS635Q04S	1.0610	0.1107	0.2714	-0.2714
	DS602Q03C	1.0614	0.7139		
	DS416Q01C	1.0677	0.2173		
	CS421Q01S	1.0678	0.1812		
	CS638Q04S	1.0685	0.5916		
	CS625Q03S	1.0686	-0.1856		
	DS643Q05C	1.0859	0.7942		
	CS645Q03S	1.0923	-0.2008		
	CS657Q03S	1.1069	0.0308		
	CS415Q08S	1.1131	-0.2111		
	CS437Q01S	1.1132	-0.5680		
	CS603Q03S	1.1136	-0.5285		
	DS425Q04C	1.1139	0.6470		
	CS641Q03S	1.1329	-1.3954		
	CS438Q01S	1.1379	-0.7641		
	CS603Q05S	1.1387	-0.1725		
	CS626Q03S	1.1424	-0.4660		
	CS649Q03S	1.1508	0.5411		
	CS634Q02S	1.1515	0.4306	0.3014	-0.3014
	DS605Q04C	1.1571	-0.0960	0.0271	-0.0271
	CS656Q04S	1.1748	0.2475		
	DS510Q04C	1.1757	0.3211		

© Waxmann Verlag GmbH

Staatenspezifische Parameter	ID der Einheit	α	σ	δ_1	δ_2
	CS438Q02S	1.1808	-0.2913		
	CS476Q03S	1.1834	-0.3024		
	DS514Q03C	1.1964	0.2624		
	CS478Q02S	1.2002	-0.0538		
	DS131Q04C	1.2111	0.5512		
	DS604Q04C	1.2315	0.5904		
	CS646Q02S	1.2433	-0.0453		
	DS620Q04C	1.2452	0.3970		
	CS610Q02S	1.2656	-1.0388		
	DS610Q01C	1.2695	0.5771		
	CS602Q04S	1.2772	-0.5581		
	DS645Q05C	1.2869	0.8469		
	CS607Q01S	1.2900	-0.9305		
	CS641Q04S	1.3161	-0.5302		
	DS637Q05C	1.3337	0.3702		
	CS428Q01S	1.3339	-0.1800		
	CS646Q03S	1.3352	-0.5106		
	DS637Q01C	1.3565	-0.0540		
	CS421Q03S	1.3621	-0.2830		
	CS635Q02S	1.3648	-0.5045		
	CS601Q02S	1.3751	-0.4128		
	CS608Q02S	1.3812	-0.2699		
	DS524Q07C	1.3815	0.4370		
	DS304Q01C	1.3868	0.3689		
	DS629Q03C	1.3963	-0.1112		
	CS638Q02S	1.4054	-0.6675		
	DS625Q01C	1.4065	0.1400		
	CS415Q02S	1.4093	-0.6720		
	CS326Q03S	1.4212	-0.1436		
	DS131Q02C	1.4253	0.0748		
	DS643Q03C	1.4267	0.4761		
	CS602Q01S	1.4396	-0.8079		
	DS648Q05C	1.4453	0.3690		
	CS615Q01S	1.4609	-0.8049		
	DS635Q05C	1.4869	0.7525	-0.2122	0.2122
	CS615Q07S	1.4969	0.4505		
	CS634Q04S	1.4972	0.0242		
	DS646Q04C	1.5131	0.5232		

© Waxmann Verlag GmbH

Staatenspezifische Parameter	ID der Einheit	α	σ	δ_1	δ_2
	DS603Q02C	1.5177	0.3278		
	CS638Q01S	1.5299	-0.0506		
	DS269Q01C	1.5406	0.0538		
	CS521Q06S	1.5446	-1.0997		
	DS634Q05C	1.5470	1.0499		
	DS646Q05C	1.5486	0.9972		
	DS638Q05C	1.5707	0.0475		
	CS605Q02S	1.5783	0.3884		
	CS634Q01S	1.5811	0.7932		
	DS304Q03AC	1.5881	0.2581		
	CS615Q02S	1.6041	0.1461		
	CS413Q06S	1.6231	0.2393		
	DS269Q03C	1.6247	0.2297		
	DS495Q03C	1.6285	0.2353		
	DS634Q03C	1.6303	0.7617		
	CS607Q02S	1.6463	0.0452		
	CS601Q01S	1.6516	1.3058		
	DS514Q02C	1.6823	-0.6716		
	CS428Q03S	1.6835	-0.4981		
	DS645Q04C	1.7222	-0.0836		
	DS428Q05C	1.7555	0.1085		
	CS603Q01S	1.7585	-0.5400		
	DS437Q06C	1.8094	-0.4582		
	CS643Q04S	1.8094	0.4647		
	DS326Q02C	1.8238	-0.2059		
	CS646Q01S	1.8920	-0.6916		
	DS648Q01C	1.8979	0.2467		
	DS608Q04C	1.9008	0.0466		
	DS514Q04C	1.9135	-0.1157		
	CS643Q02S	1.9927	-0.1109		
	DS626Q04C	2.0533	-0.0665		
	CS643Q01S	2.1770	-0.3915		
	DS635Q03C	2.1875	0.1659		
	DS304Q03BC	2.4747	0.0085		

Anmerkungen: Trennschärfe (α); Schwierigkeit (σ – Logit-Metrik); relative Kategorieschwierigkeit (δ – bei mehrstufigem Antwortformat)

© Waxmann Verlag GmbH

Tabelle B9: Übersicht zu den Ergebnissen der Item-Kalibrierung für den Kompetenzbereich Mathematik

Staatenspezifische Parameter	ID der Einheit	α	σ	δ_1	δ_2
Unique Item Parameters	CM949Q02S	1.0158	0.7498		
Unique Item Parameters	CM915Q01S	1.0199	0.4609		
Unique Item Parameters	CM408Q01S	1.3234	-0.0119		
	CM998Q04S	0.2196	1.1316		
	CM423Q01S	0.5864	-1.5480		
	CM305Q01S	0.6210	0.1488		
	CM564Q01S	0.6307	-0.0416		
	DM949Q03C	0.6313	0.3128	-2.3684	2.3684
	CM411Q02S	0.6858	0.0553		
	CM273Q01S	0.7369	0.0610		
	CM943Q01S	0.7373	-0.2017		
	CM603Q01S	0.7649	0.3262		
	CM919Q02S	0.8299	0.1001		
	CM982Q02S	0.8304	0.5106		
	CM420Q01S	0.8401	-0.0436		
	DM955Q01C	0.8567	-0.8505		
	DM462Q01C	0.8832	1.1417	-0.3741	0.3741
	DM998Q02C	0.9955	-0.7339		
	CM800Q01S	1.0000	-1.8045		
	CM982Q01S	1.0000	-1.4754		
	CM909Q01S	1.0000	-1.4554		
	CM919Q01S	1.0000	-1.2462		
	CM033Q01S	1.0000	-0.9560		
	CM905Q01S	1.0000	-0.9367		
	CM992Q01S	1.0000	-0.9358		
	CM155Q01S	1.0000	-0.6719		
	CM474Q01S	1.0000	-0.6699		
	CM982Q03S	1.0000	-0.6289		
	CM559Q01S	1.0000	-0.6164		
	CM496Q02S	1.0000	-0.6099		
	CM447Q01S	1.0000	-0.5452		
	CM906Q01S	1.0000	-0.4989		
	CM909Q02S	1.0000	-0.4423		
	DM155Q02C	1.0000	-0.3573	-0.4244	0.4244
	CM155Q04S	1.0000	-0.2756		
	CM564Q02S	1.0000	-0.0938		

© Waxmann Verlag GmbH

Staatenspezifische Parameter	ID der Einheit	α	σ	δ_1	δ_2
	CM496Q01S	1.0000	-0.0233		
	CM571Q01S	1.0000	0.0313		
	DM953Q02C	1.0000	0.0842		
	CM034Q01S	1.0000	0.2508		
	DM953Q04C	1.0000	0.7302	-0.4271	0.4271
	CM955Q03S	1.0000	0.9898	-0.7730	0.7730
	DM00KQ02C	1.0000	1.1157		
	CM00GQ01S	1.0000	1.6223		
	DM446Q02C	1.0000	1.7054		
	DM828Q02C	1.0056	-0.5344		
	DM906Q02C	1.0113	-0.0091	-0.5411	0.5411
	CM828Q03S	1.0705	0.4459		
	CM982Q04S	1.0865	-0.0551		
	DM155Q03C	1.0868	0.7350	-0.2012	0.2012
	CM915Q02S	1.2324	-0.7371		
	DM955Q02C	1.2740	0.3083		
	CM949Q01S	1.2817	-0.6425		
	DM954Q02C	1.2861	0.3615		
	CM992Q02S	1.3206	0.8826		
	CM446Q01S	1.4022	-0.5607		
	CM411Q01S	1.4262	-0.0746		
	CM442Q02S	1.4827	0.4380		
	CM954Q01S	1.5051	-0.6111		
	CM954Q04S	1.5910	0.4998		
	CM464Q01S	1.6434	0.5235		
	CM953Q03S	1.6616	-0.1894		
	CM803Q01S	1.6952	0.5137		
	DM406Q01C	1.7802	0.5120		
	CM943Q02S	1.8044	1.3310		
	CM909Q03S	1.8214	0.2583		
	DM905Q02C	1.8699	0.1215		
	DM992Q03C	2.0971	1.0177		
	DM406Q02C	2.3029	0.7832		

Anmerkungen: Trennschärfe (α); Schwierigkeit (σ – Logit-Metrik); relative Kategorieschwierigkeit (δ – bei mehrstufigem Antwortformat)

© Waxmann Verlag GmbH

Tabelle B10: Übersicht zu den Ergebnissen der Item-Kalibrierung für den Kompetenzbereich Lesen

Staatenspezifische Parameter	ID der Einheit	α	σ	δ_1	δ_2
Unique Item Parameters	CR104Q02S	0.4387	2.3297		
Unique Item Parameters	DR406Q02C	0.7466	1.5641		
Unique Item Parameters	DR219Q01EC	0.8313	-0.7099		
Unique Item Parameters	DR111Q02BC	0.9881	0.8211	0.7900	-0.7900
Unique Item Parameters	CR455Q05S	0.9890	1.8062		
Unique Item Parameters	DR456Q02C	1.0491	-1.0128		
Unique Item Parameters	CR220Q06S	1.0738	-0.3953		
Unique Item Parameters	DR055Q03C	1.1216	0.9277		
Unique Item Parameters	DR219Q01C	1.1482	-0.4546		
Unique Item Parameters	CR404Q07S	1.2485	0.4749		
Unique Item Parameters	DR102Q04C	1.2640	0.8230		
Unique Item Parameters	DR412Q08C	1.2846	1.0963		
Unique Item Parameters	DR442Q02C	1.3127	0.1651		
Unique Item Parameters	DR406Q01C	1.3519	-0.3107		
Unique Item Parameters	DR442Q03C	2.2000	0.2974		
	CR412Q06S	0.1468	2.6559		
	CR424Q02S	0.5334	0.8369		
	DR067Q04C	0.5398	0.0542	0.1542	-0.1542
	CR437Q01S	0.5634	0.5854		
	CR424Q03S	0.5722	0.2658		
	DR067Q05C	0.5918	-0.1867	-0.9643	0.9643
	CR227Q01S	0.5986	0.0801		
	CR101Q05S	0.6325	0.7670		
	CR437Q06S	0.6846	0.2200		
	CR404Q06S	0.7009	0.4907		
	CR083Q04S	0.7127	-0.1795		
	CR412Q05S	0.7187	0.1189		
	DR111Q06C	0.7237	0.6631	-0.5637	0.5637
	DR420Q06C	0.7248	0.6596		
	DR455Q02C	0.7623	1.0386		
	DR437Q07C	0.7740	1.7755		
	CR227Q02S	0.7786	0.5386	0.6766	-0.6766
	CR101Q01S	0.8231	0.4518		
	CR460Q06S	0.8257	0.0273		
	CR455Q04S	0.8443	0.0283		
	CR220Q04S	0.8574	0.0420		

© Waxmann Verlag GmbH

Staatenspezifische Parameter	ID der Einheit	α	σ	δ_1	δ_2
	DR455Q03C	0.9168	-0.6651		
	CR104Q05S	0.9410	2.0684	0.6279	-0.6279
	CR245Q01S	0.9787	0.0037		
	DR453Q04C	0.9945	-0.0592		
	CR456Q01S	1.0000	-1.9121		
	CR446Q03S	1.0000	-1.2923		
	CR067Q01S	1.0000	-1.2226		
	CR412Q01S	1.0000	-0.9548		
	CR101Q02S	1.0000	-0.7987		
	CR220Q05S	1.0000	-0.7412		
	DR456Q06C	1.0000	-0.6546		
	CR460Q05S	1.0000	-0.6351		
	CR055Q01S	1.0000	-0.6288		
	CR083Q02S	1.0000	-0.5453		
	CR424Q07S	1.0000	-0.5249		
	CR404Q03S	1.0000	-0.4548		
	DR420Q09C	1.0000	-0.4284		
	CR083Q03S	1.0000	-0.3408		
	DR406Q05C	1.0000	-0.3215		
	DR420Q10C	1.0000	0.0758	-1.3732	1.3732
	CR220Q02S	1.0000	0.3113		
	CR466Q03S	1.0000	1.6209		
	CR432Q06S	1.0000	2.1040		
	DR055Q02C	1.0110	0.4615		
	CR111Q01S	1.0245	-0.1123		
	DR102Q05C	1.0518	0.8618		
	CR245Q02S	1.0847	-0.0060		
	CR101Q04S	1.0939	-0.5283		
	DR446Q06C	1.1234	-0.2495		
	CR101Q03S	1.1315	0.0845		
	DR420Q02C	1.1366	-0.4888		
	CR102Q07S	1.1383	-0.7854		
	DR460Q01C	1.1448	-0.1149		
	CR083Q01S	1.1503	0.1660		
	DR227Q03C	1.1513	0.2452		
	CR453Q05S	1.1722	0.2288		
	CR453Q01S	1.1724	-0.5529		
	DR219Q02C	1.1751	-0.4885		

© Waxmann Verlag GmbH

Staatenspezifische Parameter	ID der Einheit	α	σ	δ_1	δ_2
	CR104Q01S	1.1815	0.2186		
	CR220Q01S	1.2371	1.2731		
	DR466Q02C	1.2926	0.5359		
	DR227Q06C	1.3554	-0.1749		
	DR432Q05C	1.3989	-0.1215		
	DR055Q05C	1.4166	-0.1350		
	DR453Q06C	1.4380	-0.2017		
	CR442Q07S	1.4537	0.7570		
	DR404Q10BC	1.4538	0.7202		
	CR466Q06S	1.4656	-0.4292		
	DR432Q01C	1.4815	-0.6482		
	DR442Q06C	1.5455	1.2272		
	DR404Q10AC	1.5609	0.5482		
	DR442Q05C	1.6691	0.8431		

Anmerkungen: Trennschärfe (α); Schwierigkeit (σ – Logit-Metrik); relative Kategorieschwierigkeit (δ – bei mehrstufigem Antwortformat)

© Waxmann Verlag GmbH

Abbildungsverzeichnis

Abbildung 2.1: Die PISA-2015-Rahmenkonzeption naturwissenschaftlicher Grundbildung 49
Abbildung 2.2: Beispielaufgabe *Völkerkollaps bei Bienen*, Item 1 60
Abbildung 2.3: Beispielaufgabe *Völkerkollaps bei Bienen*, Item 2 61
Abbildung 2.4: Beispielaufgabe *Völkerkollaps bei Bienen*, Item 3 62
Abbildung 2.5: Beispielaufgabe *Energieeffiziente Häuser,* Einleitung 64
Abbildung 2.6: Beispielaufgabe *Energieeffiziente Häuser*, Erläuterung der Simulation 64
Abbildung 2.7: Beispielaufgabe *Energieeffiziente Häuser,* Item 1 65
Abbildung 2.8: Beispielaufgabe *Energieeffiziente Häuser*, Item 2 66
Abbildung 2.9: Beispielaufgabe *Energieeffiziente Häuser*, Item 3 67
Abbildung 2.10: Mittelwerte, Streuungen und Perzentilbänder der Gesamtskala naturwissenschaftliche Kompetenz der OECD-Staaten 73
Abbildung 2.11: Prozentuale Anteile von Schülerinnen und Schülern auf Kompetenzstufe I oder darunter beziehungsweise auf Kompetenzstufe V oder VI für die Gesamtskala Naturwissenschaften 75
Abbildung 2.12: Mittelwerte naturwissenschaftlicher Kompetenz von Jungen und Mädchen in den OECD-Staaten 78
Abbildung 2.13: Mittelwerte und Streuungen der Teilskala *Phänomene naturwissenschaftlich erklären* in den OECD-Staaten 80
Abbildung 2.14: Mittelwerte und Streuungen der Teilskala *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* in den OECD-Staaten 81
Abbildung 2.15: Mittelwerte und Streuungen der Teilskala *Daten und Evidenz naturwissenschaftlich interpretieren* in den OECD-Staaten 82
Abbildung 2.16: Perzentilbänder naturwissenschaftlicher Kompetenz in Deutschland nach Schulart 86
Abbildung 2.17: Prozentuale Anteile Fünfzehnjähriger auf den Stufen der naturwissenschaftlichen Kompetenz in Deutschland nach Schulart 87
Abbildung 2.18: Mittelwerte naturwissenschaftlicher Kompetenz in den OECD-Staaten in PISA 2006 und PISA 2015 90
Abbildung 3.1: Freude und Interesse an Naturwissenschaften im internationalen Vergleich 110
Abbildung 3.2: Naturwissenschaftsbezogene instrumentelle Motivation und Selbstwirksamkeitserwartungen im internationalen Vergleich 114
Abbildung 3.3: Motivationale Orientierungen und Selbstwirksamkeitserwartung differenziert nach Schulart 116
Abbildung 3.4: Veränderungen der motivationalen Orientierungen und Selbstwirksamkeitserwartungen in Deutschland zwischen PISA 2006 und PISA 2015 117
Abbildung 3.5: Naturwissenschaftsbezogene Berufserwartungen im internationalen Vergleich 119
Abbildung 3.6: Verteilung der Arten naturwissenschaftsbezogener Berufe in Deutschland 120
Abbildung 3.7: Verteilung der Arten naturwissenschaftsbezogener Berufe in Deutschland differenziert nach Geschlecht 121
Abbildung 3.9: Die Veränderung der naturwissenschaftsbezogenen Berufserwartungen zwischen PISA 2006 und PISA 2015 123
Abbildung 3.8: Verteilung der Arten naturwissenschaftsbezogener Berufe in Deutschland differenziert nach Schulart 122
Abbildung 4.1: Disziplin im Klassenzimmer und wahrgenommene Unterstützung im internationalen Vergleich 143
Abbildung 4.2: Rückmeldung und Differenzierung im internationalen Vergleich 144
Abbildung 4.3: Forschend-entdeckender Unterricht im internationalen Vergleich 146
Abbildung 4.4: Domänenübergreifende Unterrichtsmerkmale am Gymnasium und an nicht gymnasialen Schularten 152

© Waxmann Verlag GmbH

Abbildung 4.5: Forschend-entdeckender Unterricht am Gymnasium und an nicht gymnasialen Schularten ... 153
Abbildung 4.6: Veränderung des forschend-entdeckenden Unterrichts zwischen PISA 2006 und PISA 2015 im internationalen Vergleich ... 154
Abbildung 4.7: Profillinien der latenten Klassen für die Unterrichtswahrnehmung der Fünfzehnjährigen in Deutschland ... 158
Abbildung 4.8: Vergleich der Unterrichtsmuster hinsichtlich domänenübergreifender Merkmale der Unterrichtsqualität ... 162
Abbildung 4.9: Vergleich der Unterrichtsmuster hinsichtlich naturwissenschaftlicher Kompetenz und Interesse ... 166
Abbildung 5.1: Varianzanteile der Schülerkompetenz (Naturwissenschaften) für alle Schulen sowie für Gymnasien und nicht gymnasiale Schularten (Bildungsbeteiligung in Klammern) ... 188
Abbildung 5.2: Prozentuale Anteile der Fünfzehnjährigen in Klassenstufe 9 und 10 (PISA 2000 und PISA 2015), inkl. berufliche Schulen und Förderschulen ... 191
Abbildung 6.1: Beispielaufgabe „Bergsteigen am Mount Fuji“ ... 228
Abbildung 6.2: Perzentilbänder mathematischer Kompetenz in den OECD-Staaten ... 231
Abbildung 6.3: Prozentuale Anteile der Schülerinnen und Schüler auf Kompetenzstufe I oder darunter beziehungsweise auf Kompetenzstufe V oder VI ... 233
Abbildung 6.4: Mittelwerte mathematischer Kompetenz von Mädchen und Jungen in den OECD-Staaten ... 235
Abbildung 6.5: Perzentilbänder mathematischer Kompetenz in Deutschland nach Schulart und für die Gesamtstichprobe ... 237
Abbildung 6.6: Prozentuale Anteile von Schülerinnen und Schülern auf den Stufen mathematischer Kompetenz in Deutschland nach Schulart ... 239
Abbildung 6.7: Prozentuale Anteile der Mädchen und Jungen auf den Stufen der mathematischen Kompetenz in Deutschland ... 240
Abbildung 6.8: Mathematische Kompetenz in Deutschland und an Gymnasien von PISA 2003 bis PISA 2015 ... 242
Abbildung 7.1: Der Text „Geizhals“ zu den in Abbildung 7.3 dargestellten Items 5 und 7 ... 259
Abbildung 7.2: Der Text „Heißluftballon“ zu den in Abbildung 7.3 dargestellten Items 6 und 8 ... 260
Abbildung 7.3: Itembeispiele zur Veranschaulichung der Zuordnung von Items zu den Kompetenzstufen der Lesekompetenz ... 261
Abbildung 7.4: Perzentilbänder der Lesekompetenz in den OECD-Staaten ... 264
Abbildung 7.5: Prozentuale Anteile von Schülerinnen und Schülern auf Kompetenzstufe Ia, Kompetenzstufe Ib oder darunter sowie auf Kompetenzstufe V und Kompetenzstufe VI ... 267
Abbildung 7.6: Mittelwerte der Lesekompetenz von Mädchen und Jungen in den OECD-Staaten ... 269
Abbildung 7.7: Perzentilbänder für die Lesekompetenz in Deutschland für nicht gymnasiale Schularten, Gymnasium und Gesamtstichprobe ... 272
Abbildung 7.8: Prozentuale Anteile der Schülerinnen und Schüler in Deutschland auf den Stufen der Lesekompetenz an nicht gymnasialen Schularten und am Gymnasium ... 274
Abbildung 7.9: Prozentuale Anteile der Mädchen und Jungen in Deutschland auf den Stufen der Lesekompetenz ... 275
Abbildung 7.10: Prozentuale Anteile von Mädchen und Jungen in Deutschland bei PISA 2015 und 2009 auf Kompetenzstufe Ia, Ib und darunter sowie auf Kompetenz-stufe V und VI der Gesamtskala Lesen ... 276
Abbildung 7.11: Veränderung der mittleren Lesekompetenz in Deutschland von 2000 bis 2015 ... 277
Abbildung 7.12: Veränderung der mittleren Lesekompetenz in Deutschland von 2000 bis 2015 für Mädchen und Jungen ... 278

© Waxmann Verlag GmbH

Abbildung 8.1: Verteilung des sozioökonomischen Status (HISEI) in den OECD-Staaten ... 293
Abbildung 8.2: Naturwissenschaftliche Kompetenz und Varianzaufklärung (R^2) durch den ökonomischen, sozialen und kulturellen Status (ESCS) ... 298
Abbildung 8.3: Verteilung der Kompetenzstufen in Lesen differenziert nach EGP-Klassen ... 308
Abbildung 10.1: Allgemeine und schulbezogene Aktivitäten und Interaktionen in der Familie ... 356
Abbildung 10.2: Emotionale Unterstützung der Schülerinnen und Schüler durch ihre Eltern ... 357
Abbildung 10.3: Naturwissenschaftliche Förderung in der Familie während der Grundschulzeit ... 358
Abbildung 10.4 Einrichtungen zur frühkindlichen Betreuung und Erziehung ... 359
Abbildung 10.5: Eintrittsalter in Kindergarten und Schule ... 360
Abbildung 10.6: Gründe für die Wahl von frühkindlichen Betreuungseinrichtungen ... 361
Abbildung 10.7: Häufigkeit der Gespräche zwischen Eltern und Lehrkräften ... 362
Abbildung 10.8: Kriterien der Eltern für die Schulwahl ... 363
Abbildung 10.9: Merkmale der Schulqualität aus der Perspektive der Eltern ... 364
Abbildung 10.10: Theoretisches Pfadmodell der Zusammenhänge zwischen Struktur- und Prozessmerkmalen der Unterstützung in der Familie und der naturwissenschaftlichen Kompetenz der Schülerinnen und Schüler ... 367
Abbildung 12.1 Übersicht der inhaltlichen Schwerpunktsetzung in den PISA-Runden seit 2000 ... 384
Abbildung 12.2: Vier Schritte der PISA-Stichprobenziehung ... 388
Abbildung 12.3: Stichprobendesign PISA 2015 Deutschland ... 393
Abbildung 12.4: Zuordnung der PISA-Domänen zu Basis-Testformen und deren prozentuale Verteilung in der Stichprobe ... 400
Abbildung 12.5: Item Characteristic Curves (ICC) für drei Items mit zweistufigen Antwortformat (richtig / falsch) mit unterschiedlichen Trennschärfen α ... 404
Abbildung 12.6: *Item Characteristic Curves* (ICC) für zwei Items *i* und *j* mit zweistufigen Antwortformat (richtig / falsch) mit unterschiedlichen Trennschärfen α und Schwierigkeiten σ ... 405
Abbildung 12.7: Beispielaufgabe „Geizhals“ im bisherigen Testheft-Layout (oben) und im neuen CBA-Layout (unten) ... 409
Abbildung 13.1: Die modulare Struktur der PISA-2015-Kontextfragebögen. In Anlehnung an Abb. 6.2 in OECD (2016), Übersetzung und Überarbeitung durch die Autoren ... 435
Abbildung 13.2: Internationales Erhebungsdesign der Fragebögen in PISA 2015 ... 439
Abbildung 14.1: Perzentilband und Normalverteilungskurve ... 451

Abbildung A1: Kompetenzstufen von fünfzehnjährigen Schülerinnen und Schülern mit Zuwanderungshintergrund für Naturwissenschaften ... 467
Abbildung A2: Kompetenzstufen von fünfzehnjährigen Schülerinnen und Schülern mit Zuwanderungshintergrund für Mathematik ... 468
Abbildung A3: Kompetenzstufen von fünfzehnjährigen Schülerinnen und Schülern mit Zuwanderungshintergrund für Lesen ... 469

© Waxmann Verlag GmbH

Tabellenverzeichnis

Tabelle 1.1: An PISA 2015 teilnehmende Staaten ... 22
Tabelle 1.2: Organisation der theoretischen Rahmenkonzeptionen der drei kognitiven Domänen in PISA 2015 ... 23
Tabelle 2.1: Naturwissenschaftliche Kontexte bei PISA 2015 ... 50
Tabelle 2.2: Teilkompetenzen der naturwissenschaftlichen Grundbildung ... 51
Tabelle 2.3: Drei Wissenssysteme ... 53
Tabelle 2.4: Prozedurales naturwissenschaftliches Wissen ... 54
Tabelle 2.5: Epistemisches naturwissenschaftliches Wissen ... 55
Tabelle 2.6: Verteilung der Items des Naturwissenschaftstests auf die Kontexte ... 57
Tabelle 2.7: Verteilung der Items des Naturwissenschaftstests auf die drei Teilkompetenzen, Wissensbereiche und Wissenssysteme ... 58
Tabelle 2.8: Kognitive Anforderungsniveaus in PISA 2015 ... 59
Tabelle 2.9: Einordnung Beispielaufgabe *Völkerkollaps bei Bienen* ... 62
Tabelle 2.10: Einordnung Beispielaufgabe *Energieeffiziente Häuser* ... 67
Tabelle 2.11: Stufen naturwissenschaftlicher Kompetenz in PISA 2015 ... 69
Tabelle 2.12: Mittelwerte der Teilskalen in Deutschland nach Schulart ... 89
Tabelle 2.13: Naturwissenschaftliche Kompetenz in Deutschland bei PISA 2006 und PISA 2015 getrennt nach Geschlecht und Schulart ... 91
Tabelle 3.1: *Interesse an naturwissenschaftlichen Themen* im internationalen Vergleich ... 112
Tabelle 3.2: Zusammenhang zwischen Schülermerkmalen und der naturwissenschaftsbezogenen Berufserwartung mit 30 Jahren ... 125
Tabelle 4.1: *Soziale Lernktivitäten* im internationalen Vergleich ... 147
Tabelle 4.2: *Prozedurale Lernktivitäten* im internationalen Vergleich ... 148
Tabelle 4.3: *Epistemische Lernaktivitäten*, *naturwissenschaftliches Modellieren* und *Lebensweltbezug* im internationalen Vergleich ... 149
Tabelle 4.4: Relativer Modellvergleich für LCA-Modelle mit unterschiedlicher Anzahl latenter Klassen ... 158
Tabelle 4.5: Prozentuale Häufigkeiten der Unterrichtsmuster insgesamt und getrennt für Gymnasien und andere Schularten ... 160
Tabelle 5.1: Varianz der naturwissenschaftlichen Kompetenz im internationalen Vergleich ... 185
Tabelle 5.2: Klassenstufe der PISA-Kohorte in den OECD-Staaten ... 190
Tabelle 5.3: Prozentualer Anteil der Schülerinnen und Schüler, die mindestens eine Klasse wiederholt haben ... 194
Tabelle 5.4: Anteile der Fünfzehnjährigen in Prozent, Verteilung auf Klassenstufen und Klassenwiederholungen ... 196
Tabelle 5.7: Durchschnittliche Klassengröße in den OECD-Staaten (Sekundarstufe I) ... 204
Tabelle 5.8: Instrumente der Qualitätssicherung und -entwicklung in Schulen ... 205
Tabelle 5.9: Beeinträchtigungen des Lernens durch Absenzen von Schülerinnen und Schülern und Lehrkräften ... 212
Tabelle 6.1 Verteilung der Items des Mathematiktests auf die Inhalte, Prozesse und Kontexte ... 225
Tabelle 6.2: Stufen mathematischer Kompetenz ... 226
Tabelle 6.3: Mittelwerte und Standardabweichungen nach Schulform und für die Gesamtstichprobe ... 237
Tabelle 7.1: Überblick über die typischen Anforderungen der sieben Kompetenzstufen (Ib – VI) der Lesekompetenz (vgl. Naumann, Artelt, Schneider & Stanat, 2010, S. 27–28) ... 258
Tabelle 7.2: Kodieranweisungen zu den in Abbildung 7.3 dargestellten Items 5 und 7 zu dem Text „Geizhals“ ... 262
Tabelle 7.3: Mittelwerte und Standardabweichungen der Lesekompetenz in Deutschland für nicht gymnasiale Schularten, Gymnasium und Gesamtstichprobe ... 271

© Waxmann Verlag GmbH

Tabelle 8.1: Zusammenhang zwischen Naturwissenschaftskompetenz und ökonomischem Status im internationalen Vergleich (Prädiktorvariable: HISEI) ... 294
Tabelle 8.2: Zusammenhang zwischen Naturwissenschaftskompetenz und ökonomischem, kulturellem und sozialem Status im internationalen Vergleich (Prädiktorvariable: ESCS) ... 296
Tabelle 8.3: Veränderungen in den sozialen Disparitäten des Kompetenzerwerbs über die Zeit im internationalen Vergleich des Trend-ESCS-Index und seines Zusammenhangs mit der naturwissenschaftlichen Kompetenz zwischen PISA 2006 und PISA 2015 ... 299
Tabelle 8.4: Regressionsmodelle zur Schätzung von sozialen Disparitäten in der naturwissenschaftlichen Kompetenz ... 302
Tabelle 8.5: Beispielhafte Beschreibung der EGP-Klassifikation (Baumert & Schümer, 2001, S. 339) ... 303
Tabelle 8.6: Merkmale der sozialen Lage und der Bildungswege differenziert nach EGP-Klassen ... 305
Tabelle 8.7: Mittelwerte und Standardabweichungen der Lesekompetenz differenziert nach EGP-Klassen (Bezugsperson) im Vergleich zwischen PISA 2000 und PISA 2015 in Deutschland ... 307
Tabelle 8.8: Prozentuale Anteile der Schülerinnen und Schüler in den Schularten, differenziert nach EGP-Klassen im Vergleich zwischen PISA 2000 und PISA 2015 in Deutschland ... 307
Tabelle 8.9: Merkmale der sozialen Lage differenziert nach EGP-Klassen in der Gruppe der im Lesen leistungstarken (mindestens auf Kompetenzstufe V) und leistungsschwachen (unter Kompetenzstufe II) Schülerinnen und Schüler ... 309
Tabelle 9.1: Prozentuale Anteile von fünfzehnjährigen Schülerinnen und Schülern mit Zuwanderungshintergrund in ausgewählten OECD-Staaten ... 326
Tabelle 9.2: Prozentuale Anteile von fünfzehnjährigen Schülerinnen und Schülern mit Zuwanderungshintergrund, die zu Hause die Sprache des Einwanderungslandes sprechen ... 328
Tabelle 9.3: Disparitäten im sozioökonomischen Status zwischen Familien ohne Zuwanderungshintergrund und Familien mit Zuwanderungshintergrund in ausgewählten OECD-Staaten ... 329
Tabelle 9.4: Mittlere naturwissenschaftliche Kompetenz von Jugendlichen mit Zuwanderungshintergrund in ausgewählten OECD-Staaten ... 331
Tabelle 9.5: Prozentuale Anteile fünfzehnjähriger Schülerinnen und Schüler mit Zuwanderungshintergrund in Deutschland ... 333
Tabelle 9.6: Unterschiede in der gesprochenen Sprache sowie in der sozialen Herkunft zwischen Jugendlichen mit und Jugendlichen ohne Zuwanderungshintergrund ... 335
Tabelle 9.7: Prozentuale Anteile fünfzehnjähriger Schülerinnen und Schüler mit Zuwanderungshintergrund an Gymnasien und nicht gymnasialen Schularten (Hauptschule, Integrierte Gesamtschule, Realschule und Schule mit mehreren Bildungsgängen) ... 336
Tabelle 9.8: Regressionsmodelle zur Schätzung von Disparitäten in der naturwissenschaftlichen Kompetenz zwischen Jugendlichen ohne Zuwanderungshintergrund und Jugendlichen mit Zuwanderungshintergrund ... 338
Tabelle 10.1 Regressionsmodelle zur Vorhersage der naturwissenschaftlichen Kompetenz unter Kontrolle der häuslichen Lernumgebung ... 368
Tabelle 12.1: Bruttostichprobe der Schulen nach Bundesland und Schulart ... 392
Tabelle 12.2: Überblick über die Cluster und Kompetenzbereiche mit Trend-Aufgaben und neu entwickelten Aufgaben ... 400
Tabelle 12.3: Kodierdesign und mittlere prozentuale Übereinstimmung nach Kodiererin und Kodierer für den Bereich Naturwissenschaften ... 413
Tabelle 12.4: Cut-off-Werte der Kompetenzstufen in PISA 2015 ... 419
Tabelle 12.5: Robuster Link-Fehler für Trendbetrachtungen zwischen PISA 2015 und früheren Runden ... 422

© Waxmann Verlag GmbH

Tabelle A1: Mittelwerte, Streuungen und Perzentile naturwissenschaftlicher Kompetenz 460
Tabelle A2: Mittelwerte, Streuungen und Perzentile der Teilskala *Phänomene naturwissenschaftlich erklären* 461
Tabelle A3: Mittelwerte, Streuungen und Perzentile der Teilskala *naturwissenschaftliche Forschung bewerten und Untersuchungen planen* 462
Tabelle A4: Mittelwerte, Streuungen und Perzentile der Teilskala *Daten und Evidenz naturwissenschaftlich interpretieren* 463
Tabelle A5: Mittelwerte, Streuungen und Perzentile mathematischer Kompetenz 464
Tabelle A6: Mittelwerte, Streuungen und Perzentile der Lesekompetenz 465
Tabelle A7: Mittlere Kompetenz von Jugendlichen mit Zuwanderungshintergrund für Naturwissenschaften, Mathematik und Lesen 466
Tabelle B1: Basis-Testformen nach Clustern. 471
Tabelle B2: 36 Clusterkombinationen für Naturwissenschaften 472
Tabelle B3: Blockdesign zur Zuordnung der Aufgaben aus dem Bereich Naturwissenschaften zu den Basis-Testformen 474
Tabelle B4: Zuordnung der neuen Aufgaben für den Kompetenzbereich Naturwissenschaften zu Clustern und Subdomänen 475
Tabelle B5: Zuordnung der Trend-Aufgaben für den Kompetenzbereich Naturwissenschaften zu Clustern und Subdomänen 479
Tabelle B6: Zuordnung der Aufgaben für den Kompetenzbereich Mathematik zu Clustern und Subdomänen 482
Tabelle B7: Zuordnung der Aufgaben für den Kompetenzbereich Lesen zu Clustern und Subdomänen 485
Tabelle B8: Übersicht zu den Ergebnissen der Item-Kalibrierung für den Kompetenzbereich Naturwissenschaften 488
Tabelle B9: Übersicht zu den Ergebnissen der Item-Kalibrierung für den Kompetenzbereich Mathematik 493
Tabelle B10: Übersicht zu den Ergebnissen der Item-Kalibrierung für den Kompetenzbereich Lesen 495

© Waxmann Verlag GmbH

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. Timo Ehmke (Leuphana Universität Lüneburg)
Prof. Dr. Frank Goldhammer (ZIB/DIPF Frankfurt)
Prof. Dr. Marcus Hammann (Universität Münster)
Prof. Dr. Hendrik Härtig (Universität Duisburg-Essen)
Prof. Dr. Aiso Heinze (IPN Kiel)
Prof. Dr. Silke Hertel (Universität Heidelberg)
Prof. Dr. Oliver Lüdtke (ZIB/IPN Kiel)
Prof. Dr. Knut Neumann (IPN Kiel)
Prof. Dr. Ilka Parchmann (IPN Kiel)
Prof. Dr. Manfred Prenzel (TU München)
Prof. Dr. Dominique Rauch (DIPF Frankfurt)
PD Dr. Maximilian Pfost (Universität Bamberg)
PD Dr. Christine Sälzer (ZIB/TU München)
Dr. Lars Borchert (IEA DPC Hamburg)
Dr. Nicole Haag (IQB Berlin)
Dr. Sabine Hammer (TU München)
Dr. Nina Jude (ZIB/DIPF Frankfurt)
Dr. Ulf Kröhne (ZIB/DIPF Frankfurt)
Dr. Susanne Kuger (ZIB/DIPF Frankfurt)
Dr. Katharina Müller (Leuphana Universität Lüneburg)
Dr. Silke Rönnebeck (IPN Kiel)
Dr. Anja Schiepe-Tiska (ZIB/TU München)
Dr. Katrin Schöps (IPN Kiel)
Dr. Mirjam Weis (ZIB/TU München)
Dr. Fabian Zehner (ZIB/TU München/DIPF Frankfurt)
Jens Gomolka (IEA DPC Hamburg)
Jörg-Henrik Heine (ZIB/TU München)
Matthias Lehner (TU München)
Julia Mang (ZIB/TU München)
Stefanie Schmidtner (ZIB/TU München)
Inga Simm (ZIB/TU München)
Anselm Strohmaier (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 (Mathema-

© Waxmann Verlag GmbH

tik, Informatik, Naturwissenschaften, Technik) sowie der Aus- und Weiterbildung von Lehrkräften an beruflichen Schulen im gewerblich-technischen Bereich. Das nationale Projektmanagement für PISA 2012, 2015 und 2018 ist an der *TUM School of Education* angesiedelt.

Das *DIPF (Deutsches Institut für Internationale Pädagogische Forschung)* 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. Das nationale Projektmanagement für PISA 2009 oblag dem DIPF.

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.

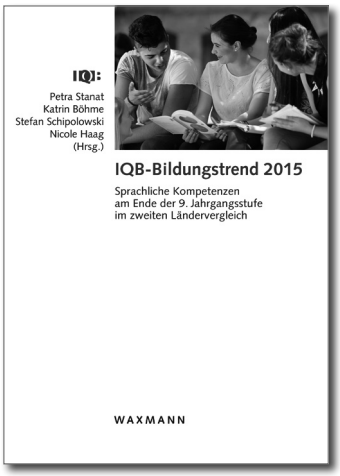
Das *ZIB (Zentrum für Internationale Bildungsvergleichsstudien, www.zib.education)* ist ein An-Institut der Technischen Universität München und umfasst als gemeinnütziger Verein die drei Institutionen *TUM School of Education, IPN Kiel* und *DIPF Frankfurt*. Der Zusammenschluss der drei Institutionen 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.

Das *IEA DPC Hamburg (Data Processing and Research Center der International Association for the Evaluation of Educational Achievement)* in Hamburg führt seit 1958 international vergleichende Schulleistungsstudien durch und ist in Deutschland seit der ersten PISA-Studie Kooperationspartner der wissenschaftlichen Studienleitung.

© Waxmann Verlag GmbH

UNSERE BUCHEMPFEHLUNG



Petra Stanat, Katrin Böhme,
Stefan Schipolowski, Nicole Haag
(Hrsg.)

IQB-Bildungstrend 2015

Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im zweiten Ländervergleich

2016, 544 Seiten, br., 44,90 €,
ISBN 978-3-8309-3535-3

Im IQB-Bildungstrend 2015 wird über die Ergebnisse des zweiten Ländervergleichs des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) in den sprachlichen Fächern berichtet. Untersucht werden Kompetenzen von Schülerinnen und Schülern der 9. Jahrgangsstufe im Jahr 2015 in den Fächern Deutsch, Englisch und Französisch. Ein Fokus liegt dabei auf Trendanalysen, die zeigen, inwieweit sich das von Neuntklässlerinnen und Neuntklässlern erreichte Kompetenzniveau in den sprachlichen Fächern seit dem IQB-Ländervergleich 2009 verändert hat. Die Referenzgröße bilden die länderübergreifenden Bildungsstandards der Kultusministerkonferenz, die fächerspezifisch festlegen, welche Kompetenzen Schülerinnen und Schüler bis zu einem bestimmten Punkt in ihrer Schullaufbahn entwickelt haben sollen.

Neben der Untersuchung der sprachlichen Kompetenzen in den Bereichen *Lesen*, *Zuhören* und *Orthografie* im Fach Deutsch sowie im *Lese-* und *Hörverstehen* in den fremdsprachlichen Fächern werden in diesem Bericht auch geschlechtsbezogene, soziale und zuwanderungsbezogene Disparitäten analysiert und ebenfalls überprüft, inwieweit hierfür seit dem Jahr 2009 Veränderungen festzustellen sind. Ergänzend werden Befunde zur Aus- und Fortbildung von Deutsch- und Englischlehrkräften berichtet.

www.waxmann.com

© Waxmann Verlag GmbH

UNSERE BUCHEMPFEHLUNG

Heike Wendt, Wilfried Bos,
Christoph Selter, Olaf Köller,
Knut Schwippert, Daniel Kasper
(Hrsg.)

TIMSS 2015

Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich

2016, 408 Seiten, br., 39,90 €
ISBN 978-3-8309-3566-7

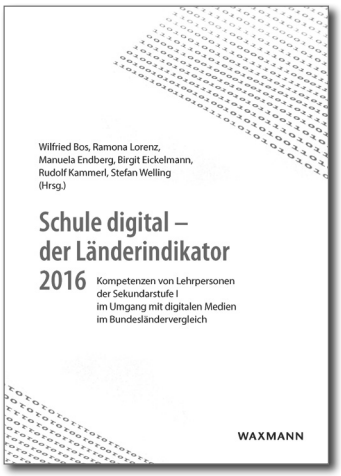
Im Jahr 2015 beteiligte sich Deutschland zum dritten Mal an der Grundschuluntersuchung *Trends in International Mathematics and Science Study* (TIMSS). Mit TIMSS werden alle vier Jahre die Fachleistungen von Schülerinnen und Schülern der vierten Jahrgangsstufe in den Bereichen Mathematik und Naturwissenschaften im internationalen Vergleich erhoben.

In diesem Band werden die Ergebnisse von TIMSS 2015 für die Bildungsdiskussion in Deutschland erschlossen. Die Ergebnisse des internationalen Vergleichs werden vor dem Hintergrund von Themen dargestellt, die das Lehren und Lernen an Grundschulen in Deutschland verändert und den Bildungsdiskurs der letzten Jahre besonders geprägt haben. Im Fokus stehen neben der Betrachtung von Schülerleistungen im internationalen Vergleich Leistungsdisparitäten zwischen Jungen und Mädchen sowie Kindern unterschiedlicher sozialer und kultureller Herkunft. Zudem werden zentrale Lehr- und Lernbedingungen in den Blick genommen, wobei Gestaltungsmerkmale des Unterrichts und die Aus- und Fortbildung von Lehrkräften, Lernbedingungen an Ganztagsschulen sowie die Inanspruchnahme von Nachhilfe differenziert betrachtet werden. Darüber hinaus werden der Übergang von der Grundschule in die Sekundarstufe I und soziale Kompetenzen von Grundschulkindern untersucht.

www.waxmann.com

© Waxmann Verlag GmbH

UNSERE BUCHEMPFEHLUNG



Wilfried Bos, Ramona Lorenz,
Manuela Endberg, Birgit Eickelmann,
Rudolf Kammerl, Stefan Welling
(Hrsg.)

Schule digital – der Länderindikator 2016

Kompetenzen von Lehrpersonen der Sekundarstufe I im Umgang mit digitalen Medien im Bundesländervergleich

2016, 292 Seiten, br., 34,90 €,
ISBN 978-3-8309-3540-7

Der *Länderindikator 2016* stellt zum zweiten Mal für Deutschland repräsentative Befunde zur schulischen Medienbildung vor. Mit diesem Bildungsmonitoring kann der Stand der Implementation digitaler Medien in Lehr- und Lernprozessen auf Bundeslandebene erfasst werden, sodass auch spezifische Entwicklungsimpulse gegeben werden können. Entwicklungen seit dem *Länderindikator 2015* können dabei für die Bereiche der schulischen Ausstattung mit digitalen Medien, deren Nutzung im Unterricht sowie für die Förderung der medienbezogenen Kompetenzen von Schülerinnen und Schülern abgebildet werden.

Darüber hinaus wird mit dem *Länderindikator 2016* der thematische Schwerpunkt auf die medienbezogenen Kompetenzen von Lehrkräften gelegt. Dazu wird die Relevanz der Lehrerausbildung für die Vermittlung entsprechender Kompetenzen, die Selbsteinschätzung der Lehrkräfte ihrer medienbezogenen Kompetenzen, Fortbildungsaktivitäten sowie die medienbezogene Kooperation unter Lehrkräften vertiefend im Bundesländervergleich in den Blick genommen. Zudem wird mit einem Länderportrait zur schulischen Medienbildung in Dänemark ein Land hervorgehoben, das im Vergleich zu Deutschland einen vertieften Implementationsstand digitaler Medien im Schulsystem aufweist.

www.waxmann.com

© Waxmann Verlag GmbH