peDOCS
Open Access Erziehungswissenschaften

DIPF

Schwippert, Knut [Hrsg.]; Kasper, Daniel [Hrsg.]; Eickelmann, Birgit [Hrsg.]; Goldhammer, Frank [Hrsg.]; Köller, Olaf [Hrsg.]; Selter, Christoph [Hrsg.]; Steffensky, Mirjam [Hrsg.]

TIMSS 2023. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich

Münster ; New York : Waxmann 2024, 365 S.

Quellenangabe/ Reference:
Schwippert, Knut [Hrsg.]; Kasper, Daniel [Hrsg.]; Eickelmann, Birgit [Hrsg.]; Goldhammer, Frank [Hrsg.]; Köller, Olaf [Hrsg.]; Selter, Christoph [Hrsg.]; Steffensky, Mirjam [Hrsg.]: TIMSS 2023. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Münster ; New York : Waxmann 2024, 365 S. - URN: urn:nbn:de:0111-pedocs-332656 - DOI: 10.25656/01:33265; 10.31244/9783830999591

https://nbn-resolving.org/urn:nbn:de:0111-pedocs-332656
https://doi.org/10.25656/01:33265

in Kooperation mit / in cooperation with:

WAXMANN
www.waxmann.com

http://www.waxmann.com

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: http://creativecommons.org/licenses/by-nc-sa/4.0/deed.de - Sie dürfen das Werk bzw. den Inhalt unter folgenden Bedingungen vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen: Sie müssen den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen. Dieses Werk bzw. der Inhalt darf nicht für kommerzielle Zwecke verwendet werden. Die neu entstandenen Werke bzw. Inhalte dürfen nur unter Verwendung von Lizenzbedingungen weitergegeben werden, die mit denen dieses Lizenzvertrages identisch oder vergleichbar sind.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: http://creativecommons.org/licenses/by-nc-sa/4.0/deed.en - You may copy, distribute and transmit, adapt or exhibit the work in the public and alter, transform or change this work as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work. If you alter, transform, or change this work in any way, you may distribute the resulting work only under this or a comparable license.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Mitglied der Leibniz-Gemeinschaft



Knut Schwippert
Daniel Kasper
Birgit Eickelmann
Frank Goldhammer
Olaf Köller
Christoph Selter
Mirjam Steffensky
(Hrsg.)

TIMSS 2023

Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich

WAXMANN

Knut Schwippert, Daniel Kasper, Birgit Eickelmann,
Frank Goldhammer, Olaf Köller, Christoph Selter,
Mirjam Steffensky (Hrsg.)

TIMSS 2023

Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich



Waxmann 2024
Münster · New York

Die Beteiligung Deutschlands an TIMSS 2023 wurde gefördert durch das Bundesministerium für Bildung und Forschung (BMBF) und die Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (KMK)

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

KULTUSMINISTER
KONFERENZ

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

Print-ISBN 978-3-8309-4959-6
E-Book-ISBN 978-3-8309-9959-1
https://doi.org/10.31244/9783830999591

2024, Waxmann Verlag GmbH,
Steinfurter Straße 555, 48159 Münster

www.waxmann.com
info@waxmann.com

Umschlaggestaltung: Anne Breitenbach, Münster
Titelfoto: © skynesher – Istockphoto.com
Satz: Roger Stoddart, Münster

Dieses E-Book steht open access unter der Creative-Commons-Lizenz CC BY-NC-SA 4.0 zur Verfügung.

Diese Lizenz gilt nur für das Originalmaterial. Alle gekennzeichneten Fremdinhalte (z.B. Abbildungen, Fotos, Zitate etc.) sind von der CC-Lizenz ausgenommen und für deren Wiederverwendung ist es ggf. erforderlich, weitere Nutzungsgenehmigungen beim jeweiligen Rechteinhaber einzuholen.

Inhalt

Vorwort **9**

Kapitel 1

TIMSS 2023: Wichtige Ergebnisse im Überblick **11**

Knut Schwippert, Daniel Kasper, Birgit Eickelmann, Frank Goldhammer, Olaf Köller, Christoph Selter und Mirjam Steffensky

1.1 Einleitung 11
1.2 Zentrale Ergebnisse 13
1.3 Bildungspolitische und didaktische Folgerungen 20
Literatur 23

Kapitel 2

Ziele, Anlage und Durchführung der *Trends in International Mathematics and Science Study* (TIMSS) 2023 ..... **25**

Christin Beese, Luise A. Scholz, Donieta Jusufi, Daniel Kasper und Knut Schwippert

2.1 Einleitung 25
2.2 Forschungsnetzwerk in TIMSS 2023 26
2.3 Ziele der TIMSS-Untersuchung 27
2.4 Anlage der Studie 32
2.5 Vorbereitung und Durchführung von TIMSS 2023 44
2.6 Technische und empirische Voraussetzungen für die Auswertung 47
2.7 Hinweise für die Ergebnisdarstellung 47
Literatur 50

Kapitel 3

Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse **53**

Christoph Selter, Daniel Walter, Aiso Heinze, Johanna Brandt und Daniel Kasper

3.1 Einleitung 53
3.2 Mathematische Grundbildung in Deutschland 55
3.3 TIMSS-Rahmenkonzeption zur Erfassung mathematischer Kompetenzen 58
3.4 Der TIMSS-Test zur Erfassung der mathematischen Kompetenzen 62
3.5 Ergebnisse 73
3.6 Zusammenfassung 97
Literatur 102

Kapitel 4

Naturwissenschaftliche Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse 105

Mirjam Steffensky, Luise A. Scholz und Olaf Köller

4.1 Einleitung 105
4.2 Naturwissenschaftliche Grundbildung in der Primarstufe in Deutschland 107
4.3 TIMSS-Rahmenkonzeption zur Erfassung der naturwissenschaftlichen Kompetenzen 110
4.4 TIMSS-Test 116
4.5 Kompetenzen im internationalen Vergleich 127
4.6 Zusammenfassung 151
Literatur 154

Kapitel 5

Lehren und Lernen mit digitalen Medien in der Grundschule 159

Donieta Jusufi, Knut Schwippert, Birgit Eickelmann, Julia Gerick und Ulf Kröhne

5.1 Einleitung 159
5.2 Relevante Merkmale im Zusammenhang mit digitalen Medien 161
5.3 Verfügbarkeit schulischer und häuslicher digitaler Ressourcen im internationalen Vergleich 164
5.4 Digitalisierungsbezogene Lernangebote im Fachunterricht im internationalen Vergleich 168
5.5 Selbsteingeschätzte Kompetenzen der Schüler:innen im Umgang mit digitalen Medien im internationalen Vergleich 175
5.6 Nutzung digitaler Medien in der Schule und im Unterricht in Deutschland 180
5.7 Bedarf und Teilnahme von Lehrkräften an Fortbildungen zu digitalen Medien im internationalen Vergleich und im Trend 186
5.8 Zusammenfassung 189
Literatur 192

Kapitel 6

Geschlechterunterschiede in mathematischen und naturwissenschaftlichen Kompetenzen 197

Sonja Nonte, Eva Grommé und Luise A. Scholz

6.1 Einleitung 197
6.2 Theoretische Erklärungsansätze zu Geschlechterunterschieden in mathematischen und naturwissenschaftlichen Leistungen 198
6.3 Befunde zu Geschlechterunterschieden in mathematischen und naturwissenschaftlichen Leistungen in Schulleistungsstudien 199
6.4 Ergebnisse zu geschlechterspezifischen Leistungsunterschieden in Mathematik 201
6.5 Ergebnisse zu geschlechterspezifischen Leistungsunterschieden in den Naturwissenschaften 211
6.6 Ergebnisse zu geschlechterspezifischen Unterschieden in Einstellungen und Selbstkonzepten 220
6.7 Zusammenfassung 225
Literatur 229

Kapitel 7

Soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässler:innen 233

Tobias C. Stubbe, Leonhard Schulz und Christin Beese

7.1 Einleitung.....233
7.2 Soziale Herkunft und Bildungserfolg.....234
7.3 Indikatoren der sozialen Herkunft.....235
7.4 Soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen im internationalen Vergleich 237
7.5 Soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen in Deutschland.....246
7.6 Zusammenfassung 254
Literatur.....256

Kapitel 8

Mathematische und naturwissenschaftliche Kompetenzen von Schüler:innen mit und ohne Migrationshintergrund.....259

Knut Schwippert, Donieta Jusufi, Martje Lott und Tobias C. Stubbe

8.1 Einleitung.....259
8.2 Weitere relevante Heterogenitätsmerkmale der Schüler:innen und ihrer Familien 261
8.3 Erfassung des Migrationshintergrundes in TIMSS 2023.....262
8.4 Mathematische und naturwissenschaftliche Kompetenzen der Schüler:innen nach Migrationshintergrund im internationalen Vergleich 263
8.5 Lebensbedingungen von Schüler:innen mit und ohne Migrationshintergrund in Deutschland.....267
8.6 Kompetenzunterschiede von Schüler:innen mit und ohne Migrationshintergrund in Deutschland im Trend.....268
8.7 Fachbezogene Selbstkonzepte, Einstellung und soziale Eingebundenheit von Schüler:innen mit und ohne Migrationshintergrund.....271
8.8 Zusammenfassung 272
Literatur.....274

Kapitel 9

Trends in der Schülerzusammensetzung und den mathematischen und naturwissenschaftlichen Kompetenzen am Ende der Grundschulzeit in Deutschland 277

Daniel Kasper, Jakob Sierau, Knut Schwippert und Olaf Köller

9.1 Einleitung.....277
9.2 Charakteristika der Viertklässler:innen in TIMSS 2007, 2019 und 2023 280
9.3 Entwicklungen während der COVID-19-Pandemie.....281
9.4 Trends in den Kompetenzen der Schüler:innen.....287
9.5 Veränderungen in den Schülerzusammensetzungen.....295
9.6 Zusammenfassung 300
Literatur.....302

Kapitel 10

Unterrichten in der Grundschule: Unterrichtsqualität und Lehrkräftefortbildung 305
Luise A. Scholz, Christin Beese, Armin Jentsch, Jakob Sierau und Mirjam Steffensky

10.1 Einleitung 305
10.2 Unterrichtsqualität 306
10.3 Lehrkräftefortbildung 314
10.4 Zusammenfassung 322
Literatur 324

Kapitel 11

Privater Nachhilfeunterricht und Förderungen zum Ausgleich pandemiebedingter Lernrückstände am Ende der Grundschulzeit 327
Lisa Benckwitz, Karin Guill und Christin Beese

11.1 Einleitung 327
11.2 Verbreitung von Nachhilfeunterricht 329
11.3 Teilnahme an Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände 334
11.4 Zusammenfassung 337
Literatur 338

Kapitel 12

Schullaufbahnpräferenzen am Übergang in die Sekundarstufe und der Zusammenhang mit leistungsrelevanten und sozialen Merkmalen 341
Tobias C. Stubbe, Annika Beier und Martje Lott

12.1 Einleitung 341
12.2 Theorie und Forschungsstand zu Schullaufbahnpräferenzen am Übergang in die Sekundarstufe 342
12.3 Schullaufbahnpräferenzen von Lehrkräften und Erziehungsberechtigten: Verteilung und Ausmaß der Übereinstimmung 345
12.4 Schullaufbahnpräferenzen von Lehrkräften: Leistungsrelevante Merkmale der Schüler:innen 347
12.5 Schullaufbahnpräferenzen von Lehrkräften und Erziehungsberechtigten: Soziale Merkmale der Schüler:innen 350
12.6 Zusammenfassung 357
Literatur 358

Anhang 361

Vorwort

Das Zitat „Ohne Daten ist man nur eine Person mit einer Meinung“ wird dem amerikanischen Physiker, Statistiker und Pionier für Qualitätsmanagement William Edwards Deming zugeschrieben.[1] Gerade in Zeiten, in denen anhand digitaler Medien Informationen, ohne diese zu reflektieren oder zu prüfen, schnell verbreitet werden, sowie die Zugangsmöglichkeiten und Häufigkeiten der Nutzung eben solcher ‚Informationsquellen‘ dramatisch steigt, erscheint es umso wichtiger solide Daten zur Entscheidungsfindung zur Verfügung zu haben. Genau diesen Anspruch erhebt die *Trends in International Mathematics and Science Study* (TIMSS), die international seit 1995 und in Deutschland seit 2007 regelmäßig alle vier Jahre über die Kompetenzstände von Viertklässler:innen informiert und dabei auch eine Vielzahl von relevanten Rahmenbedingungen des Lehrens und Lernens mit betrachtet. Die TIMSS-Untersuchung zählt zu den groß angelegten Schulleistungsstudien, deren spezifischer Fokus auf das Bildungssystem gerichtet ist und nicht auf die Bewertung einzelner Schüler:innen. Die besondere Stärke von TIMSS ist, dass international vergleichend Schüler:innen am Ende ihrer Grundschulzeit untersucht werden und damit ein *natürliches Experiment* (Husén, 1967)[2] zur Verfügung steht, um international vergleichend Gemeinsamkeiten und Unterschiede auf Basis einheitlich erhobener Daten zu dokumentieren und anhand von wissenschaftlichen Theorien zu reflektieren und für Veränderung im Bildungssystem zu nutzen. Hierbei sind die Adressat:innen der Studien diejenigen, die sich mit der Struktur und den Rahmenbedingungen von Bildungssystemen auseinandersetzen. Von den durchführenden Wissenschaftler:innen werden die gewonnenen Informationen für die Bildungspolitik und -administration, aber auch für Lehrkräfte, Schulleitungen und eine weitere interessierte Öffentlichkeit aufbereitet und präsentiert.

Im Rahmen von TIMSS werden über die Zeit systematisch Informationen ermittelt, die zum einen den Vergleich zwischen den an der Studie beteiligten Staaten und Regionen erlauben und zum anderen die alle vier Jahre wiederholten Bestandsaufnahmen von relevanten Indikatoren zur Beschreibung von Bildungssystemen im Trend ermöglichen. Gleichzeitig wird in jedem Zyklus auch immer wieder der Blick auf aktuelle Veränderungen oder Gegebenheiten gerichtet, um hierüber ebenso relevante wie auch verlässliche Daten zu erheben. Dabei wird kontinuierlich vor dem Hintergrund des aktuellen Forschungsstandes abgewogen, welche neu entwickelten spezifischen Fragen berücksichtigt werden und welche Fragen nicht mehr relevant sind und somit aus den Fragebögen entfernt werden, um die befragten Schüler:innen, deren Eltern bzw. Erziehungsberechtigte, die sie unterrichtenden Lehrkräfte und verantwortlichen Schulleitungen nicht übermäßig in Anspruch zu nehmen. Hierbei hat sich die konsequente Verzahnung von bereits vorliegenden Forschungsbefunden, empirisch fundierten Theorien und der fokussierten Weiterentwicklung der Fragebögen bewährt. Diese Herangehensweise sichert eine solide und belastbare Datenbasis für zukünftige Reformen im Bildungssystem. In TIMSS 2023 wird die, insbesondere auch für die Schule und den Unterricht, besonders prägende COVID-19-Pandemie berücksichtigt, weshalb

1 https://www.quality.de/lexikon/william-edwards-deming/ [o. S.; abgerufen am 27.09.2024].
2 Husén, T. (1967). *International study of achievement in mathematics: A comparison of twelve countries*. Wiley/Almqvist & Wiksell.

vertieft Fragen zur Nutzung und Nutzbarkeit digitaler Medien gestellt wurden, die im Rahmen des Fern- und Wechselunterrichts genutzt wurden.

Deutschland ist nicht nur Teilnehmer an verschiedenen internationalen Schulleistungsstudien (wie z. B. die *Internationale Grundschul-Lese-Untersuchung*/*Progress in International Reading Literacy Study*, IGLU/PIRLS; *Programme for International Student Assessment*, PISA; *International Computer and Information Literacy Study*, ICILS). In Deutschland werden zudem sowohl übergreifend über die Länder der Bundesrepublik Deutschland hinweg (wie z. B. IQB-Bildungstrend) als auch spezifisch in einzelnen dieser Länder entsprechende Untersuchungen (wie z. B. *Kompetenzen ermitteln*, Kermit in Hamburg) – die auch als Bildungsmonitoringstudien bezeichnet werden – realisiert. Zwar sind alle diese Studien für sich genommen auf die Untersuchung bestimmter Aspekte spezialisiert, dennoch sind sie konsequent miteinander wissenschaftlich vernetzt. Die Besonderheiten der diesem Bericht zugrundeliegenden TIMSS-Untersuchung lassen sich folgendermaßen zusammenfassen: TIMSS ist in Deutschland die einzige Bildungsmonitoringstudie, die die mathematischen und naturwissenschaftlichen Kompetenzen von Schüler:innen am Ende ihrer Grundschulzeit repräsentativ für Deutschland in den Fokus nimmt und dabei die Befunde international vergleichend mit einer großen Zahl an Bildungssystemen aus der ganzen Welt gegenüber stellen kann. Gerade diese Perspektive bietet für Deutschland die Möglichkeit, die eigene Gestaltung des Bildungssystems vor dem Hintergrund international erhobener und wissenschaftlich fundierter Informationen zu reflektieren und ggf. Vorgehensweisen anderer Staaten zur Orientierung für eigene Reformen zum Vorbild zu nehmen, die in Deutschland (noch) nicht realisiert wurden.

Kapitel 1
TIMSS 2023: Wichtige Ergebnisse im Überblick

Knut Schwippert, Daniel Kasper, Birgit Eickelmann, Frank Goldhammer, Olaf Köller, Christoph Selter und Mirjam Steffensky

1.1 Einleitung

Die *Trends in International Mathematics and Science Study* (TIMSS) 2023 hat das Ziel, mathematische und naturwissenschaftliche Kompetenzen von Viertklässler:innen sowie die Rahmenbedingungen und Voraussetzungen des Lernens zu untersuchen. Die TIMS-Studie legt den Fokus auf die Dokumentation und Interpretation langfristiger Entwicklungen von Bildungssystemen im internationalen Vergleich. Durch die gleichzeitige Betrachtung von mathematischen und naturwissenschaftlichen Kompetenzen, curricularen Vorgaben sowie schulischen und außerschulischen Lehr- und Lernbedingungen wird es möglich, zu Erkenntnissen über die Zusammenhänge dieser Rahmenbedingungen und den Lernerfolgen der Grundschüler:innen zu gelangen und Entscheidungshilfen für bildungspolitische Fragen zur Weiterentwicklung der Bildungssysteme zu geben. Deutschland beteiligt sich seit 2007 an der alle vier Jahre stattfindenden Grundschuluntersuchung, sodass für den Mathematik- und den naturwissenschaftsbezogenen Sachunterricht mittlerweile Entwicklungen im deutschen Bildungssystem über einen Zeitraum von 16 Jahren abgebildet werden können. Im Mittelpunkt des vorliegenden Berichtsbandes von TIMSS 2023 stehen die Zyklen von 2007 (als erster Zyklus mit deutscher Beteiligung), 2019 (als unmittelbar vorangegangener Zyklus) und 2023 (als aktueller Zyklus). Die zwischen 2007 und 2019 liegenden Zyklen von 2011 und 2015 sind ausführlich im Berichtsband von TIMSS 2019 dokumentiert (Schwippert et al., 2020). Der Vergleich der Zyklen 2019 und 2023 ist wegen der pandemiebedingten Schulschließungen in den Schuljahren 2019/20 und 2020/21 von großem bildungspolitischen Interesse, um möglichen (negativen) Effekten der Schulschließungen auf den Kompetenzerwerb in der Grundschule nachzugehen.

Die groß angelegte und weltweit durchgeführte Untersuchung basiert auf der Zusammenführung unterschiedlicher Expertisen aus dem Bereich der empirischen Bildungsforschung. So wirken in dieser Untersuchung internationale und nationale Expert:innen unterschiedlicher Fachrichtungen, einerseits spezialisiert für die teilnehmenden Staaten und Regionen, andererseits koordiniert über alle Teilnehmer hinweg, eng zusammen.

Initiator ist die *International Association for the Evaluation of Educational Achievement* (IEA). Die IEA hat seit 1958 weitreichende Erfahrungen in der Konzeption, Durchführung und Auswertung von internationalen Schulleistungsstudien. Der Ausgangspunkt für die alle vier Jahre durchgeführte Trendstudie TIMSS wurde 1995 gesetzt.

Die operative Planung und Implementierung der TIMSS-Untersuchung hat die IEA auch für das Jahr 2023 wieder an das *TIMSS & PIRLS International Study Center* (ISC) am Boston College vergeben. Damit obliegt es dem ISC, das Design der Studie, die internationale Koordination der Entwicklung der Test- und Befragungsinstrumente und der Erhebungsprozeduren sowie die Überprüfung der Qualität der Datenerhebung zu initiieren und mit den beteiligten Staaten und Regionen abzustimmen. Das ISC übernimmt dabei auch hauptverantwortlich die technische Dokumentation der Untersuchung und die Erstellung des internationalen Ergebnisberichts. In Deutschland liegt die inhaltliche Verantwortung für die TIMSS-Untersuchung beim interdisziplinär besetzten nationalen TIMSS-2023-Konsortium.

Zur Untersuchung der mathematischen und naturwissenschaftlichen Kompetenzen der Schüler:innen wurde eine Rahmenkonzeption entwickelt, die möglichst große Überschneidungen mit den Curricula aller teilnehmenden Staaten und Regionen aufweisen soll. Aus nationaler Perspektive zeigt sich für Mathematik, dass die Rahmenkonzeption und entsprechend die Kompetenztests weitgehend die aktuellen von der *Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland* (KMK) verabschiedeten Bildungsstandards widerspiegeln (KMK, 2004). Diese legen fest, welche mathematischen Kompetenzen Schüler:innen am Ende ihrer Grundschulzeit erworben haben sollten. Für die Naturwissenschaften gibt es keine länderübergreifenden Bildungsstandards für die Grundschule. Die Bildungspläne ausgewählter Länder sowie der Perspektivrahmen der *Gesellschaft für Didaktik des Sachunterrichts* (GDSU) zeigen aber eine hohe Passung zu der TIMSS-Rahmenkonzeption und den Kompetenztests.

In der TIMSS-Rahmenkonzeption werden sowohl für Mathematik als auch für die Naturwissenschaften drei Inhalts- und Anforderungsbereiche formuliert. Die Inhalte für Mathematik umfassen die Bereiche *Arithmetik*, *Messen und Geometrie* sowie *Daten*, die Inhalte für die Naturwissenschaften *Biologie*, *Physik/Chemie* und *Geografie*. Sowohl für den Kompetenztest in Mathematik als auch in den Naturwissenschaften werden drei kognitive Anforderungsbereiche unterschieden:

- *Reproduzieren* von Wissen, Fertigkeiten und Grundvorstellungen,
- *Anwenden* von Wissen, Fertigkeiten und Grundvorstellungen beim Bearbeiten von Standardaufgaben sowie
- *Problemlösen* bei komplexeren, anspruchsvolleren Aufgabenstellungen.

Neben der Auseinandersetzung mit den Inhalts- und Anforderungsbereichen der Aufgaben spielt auch das Testformat eine wichtige Rolle. In TIMSS 2023 wurde die im Jahr 2019 begonnene Umstellung der Erhebungen auf ein digitales Format weitergeführt, sodass in TIMSS 2023 neben den Kompetenztests auch drei von

vier Hintergrundfragebögen digital erhoben wurden. Lediglich den Eltern bzw. Erziehungsberechtigten wurden die Fragen in einem papierbasierten Format vorgelegt, da hierdurch eine höhere Rücklaufquote der Fragebögen erwartet wurde.

Der Bedeutsamkeit des digitalen Wandels wurde durch ein vertiefendes Teilprojekt zur Untersuchung des Lehrens und Lernens mit digitalen Medien in der Grundschule Rechnung getragen. Hierbei wurde der Fokus auf die vorhandenen Ressourcen und deren Nutzung im Grundschulunterricht gerichtet. Eine weitere Besonderheit der TIMSS-2023-Untersuchung ist der Blick auf die potenziellen Auswirkungen der COVID-19-Pandemie. Aufgrund des Testzeitpunktes im Frühjahr 2023 nahmen Viertklässler:innen teil, die die ersten beiden Schuljahre ihrer Grundschulzeit während der COVID-19-Pandemie erlebt haben. Die Kinder waren sowohl im Schuljahr 2019/20 als auch im Schuljahr 2020/21 von bundesweiten und regionalen Schulschließungen betroffen. Dieser Umstand muss einerseits bei der Interpretation der Ergebnisse berücksichtigt werden, bietet andererseits aber auch die Möglichkeit, den unterschiedlichen Umgang mit dem Lehren und Lernen während der COVID-19-Pandemie und mögliche Auswirkungen im direkten Vergleich zu TIMSS 2019 zu beleuchten.

1.2 Zentrale Ergebnisse

Der vorliegende Berichtsband bietet eine nationale Vertiefung und Ausdifferenzierung der internationalen Berichtslegung des ISC (von Davier et al., 2024). Wesentliche Befunde für die 58 Teilnehmerstaaten sowie die fünf Benchmark-Teilnehmer finden sich in den internationalen Berichten. Dagegen werden in diesem Band Ergebnisse mit besonderem Fokus auf und wichtigen Implikationen für das deutsche Bildungssystem herausgearbeitet und um vertiefende Analysen ergänzt. Dabei werden nationale Hintergrundinformationen berücksichtigt, die über die internationale Berichtslegung hinausgehen.

1.2.1 Mathematische Kompetenzen im internationalen Vergleich

Mit TIMSS 2023 liegen Befunde über einen Zeitraum von 16 Jahren vor, die es ermöglichen, mathematische Kompetenzen, Einstellungen und Selbstkonzepte von Viertklässler:innen in Deutschland im internationalen Vergleich und im Trend in den Blick zu nehmen. Die von Selter, Walter, Heinze, Brandt und Kasper im dritten Kapitel vorgelegten Analysen erlauben es, differenzierte Aussagen über spezifische Stärken und Schwächen der Schüler:innen in den Inhaltsbereichen und den kognitiven Anforderungsbereichen des Mathematikunterrichts am Ende der Grundschulzeit zu treffen.

Der Leistungsmittelwert Deutschlands liegt mit 524 Punkten oberhalb des internationalen Mittelwerts und im oberen Mittelfeld der internationalen Rangreihe. Im Vergleich zur Erhebung von TIMSS 2019 können für Deutschland keine signifikanten Veränderungen festgestellt werden. Angesichts der gestiegenen Anforderungen an die Grundschule sowie der pandemiebedingten Unterrichtsausfälle ist dieses Resultat positiv zu beurteilen. Ein ‚Abwärtstrend' in den Leistungen der Schüler:innen, wie er im vom *Institut zur Qualitätsentwicklung im Bildungswesen* (IQB) durchgeführten IQB-Bildungstrend 2021 (Stanat et al., 2022) berichtet wurde, ist nicht erkennbar. Erwähnenswert ist jedoch, dass für die Einstellung zur Mathematik und für das mathematikbezogene Selbstkonzept

signifikant gesunkene mittlere Skalenwerte berichtet werden. Zudem zeigt sich, dass Deutschland einer von fünf teilnehmenden Staaten ist, für die keine signifikanten Leistungsunterschiede zwischen 2007 und 2023 zu verzeichnen sind. Ein ‚Aufwärtstrend' ist also ebenfalls nicht erkennbar. Darüber hinaus ist festzuhalten, dass sich weiterhin etwa ein Viertel der Schüler:innen auf den Kompetenzstufen I und II befindet und nur über geringe mathematische Kompetenzen verfügt. Auf der anderen Seite des Leistungsspektrums sind es lediglich 8 Prozent der Schüler:innen in Deutschland, die die höchste Kompetenzstufe erreichen. Auch wenn es hier in Gegenüberstellung zu TIMSS 2019 zu einer signifikanten Steigerung kam, handelt es sich im internationalen Vergleich um einen immer noch niedrigen Wert.

Die Viertklässler:innen zeigen relative Stärken im Inhaltsbereich *Messen und Geometrie* sowie im Anforderungsbereich *Reproduzieren*. Für den Inhaltsbereich *Daten* sowie den Anforderungsbereich *Anwenden* zeigen sich hingegen in TIMSS 2023 relative Schwächen.

1.2.2 Naturwissenschaftliche Kompetenzen im internationalen Vergleich

Naturwissenschaftliche Kompetenzen von Viertklässler:innen stehen im Fokus des vierten Kapitels von Steffensky, Scholz und Köller. Der Mittelwert der naturwissenschaftlichen Kompetenzen der Grundschüler:innen liegt in TIMSS 2023 in Deutschland mit 515 Punkten lediglich drei Punkte unter dem Mittelwert des vorherigen Studienzyklus, gegenüber TIMSS 2007, 2011 und 2015 sind es 13 Punkte weniger. Das im Vergleich zu früheren TIMSS-Erhebungen festgestellte schwächere Abschneiden der Schüler:innen im Jahr 2023 kann nicht an einem bestimmten Inhaltsbereich (*Biologie*, *Physik/Chemie* und *Geografie*) festgemacht werden. Auch zwischen den drei Anforderungsbereichen (*Reproduzieren*, *Anwenden* und *Problemlösen*) finden sich zwar kleinere Unterschiede, die aber praktisch nicht bedeutsam sind. Neben den etwas niedrigeren Kompetenzen ist eine Zunahme der Streuung bei den naturwissenschaftlichen Kompetenzen zu beobachten. Diese ergibt sich vor allem durch die Zunahme der Streuung am unteren Ende des Leistungsspektrums. Mit Blick auf die Kompetenzstufen zeigt sich ein ähnliches Bild wie in den vorherigen Studienzyklen: So finden sich eher wenige Schüler:innen auf der höchsten Kompetenzstufe (9 %) und recht viele auf den beiden untersten Kompetenzstufen (ca. 30 %).

Systematische pandemiebedingte Kompetenzeinbußen lassen sich weder in Deutschland noch in anderen teilnehmenden Staaten und Regionen klar ausmachen. Ob das damit zusammenhängt, dass naturwissenschaftlicher Unterricht in den ersten zwei Schuljahren möglicherweise eine etwas geringere Rolle spielt und damit die Schulschließungen für die hier betrachtete Kohorte im Hinblick auf die naturwissenschaftliche Kompetenz nicht so einschneidend waren, lässt sich aus den Daten nicht eindeutig ableiten.

In TIMSS 2023 ist erstmalig ein Teil der naturwissenschaftlichen Aufgaben, die den Schwerpunkt auf Umweltthemen setzen, international gesondert skaliert worden. Der Mittelwert dieser Nachhaltigkeitskompetenzen unterscheidet sich in Deutschland nicht wesentlich von der Gesamtskala Naturwissenschaften. Die ebenfalls erhobenen nachhaltigkeitsbezogenen Einstellungen zeigen für Grundschüler:innen in Deutschland eine hohe Präferenz zum Schutz der Natur und deren Wertschätzung.

1.2.3 Lehren und Lernen mit digitalen Medien in der Grundschule

Im fünften Kapitel zeichnen Jusufi, Schwippert, Eickelmann, Gerick und Kröhne nach, dass in Deutschland, auch im internationalen Vergleich, eine gute Ausstattung mit digitalen Medien im häuslichen und schulischen Kontext vorliegt. Nur ein sehr geringer Anteil von Kindern kann zu Hause nicht auf digitale Geräte wie Computer, Laptop, Tablet oder E-Reader zurückgreifen, um diese für schulische Zwecke, wie beispielsweise das Erledigen von Hausaufgaben, zu nutzen. In der vertieften Betrachtung für Deutschland zeigt sich weiter, dass mehr als ein Drittel der Viertklässler:innen von Lehrkräften unterrichtet wird, die angeben, für ihren Mathematikunterricht keine digitalen Medien zur Verfügung zu haben. Für den naturwissenschaftlichen Sachunterricht ist dieser Anteil nur ungefähr halb so groß. Auch wenn die generelle Verfügbarkeit digitaler Medien in den Grundschulen in Deutschland gegeben ist, finden sich Grenzen in der tatsächlichen Nutzung im Fachunterricht. Diese für 2023 dokumentierten Befunde überraschen, da auch in Grundschulen in Deutschland die COVID-19-Pandemie zu einem ‚Digitalisierungsschub' geführt haben sollte. Zwar kann hinsichtlich der IT-Ausstattung im schulischen Kontext positiv hervorgehoben werden, dass etwas mehr als 57 Prozent der Schüler:innen von Lehrkräften unterrichtet werden, denen es gelang, den Unterricht an die herausfordernden Bedingungen während der COVID-19-Pandemie in kurzer Zeit anzupassen. Dennoch ist die Nutzung von digitalen Medien im Fachunterricht im Jahr 2023 nicht weit verbreitet. Im Gegenteil: Bei der national vertieften Betrachtung des Einsatzes digitaler Medien für unterschiedliche Tätigkeiten zeigt sich, dass der Einsatz digitaler Medien – trotz der Verfügbarkeit digitaler Geräte – sowohl für den Mathematikunterricht als auch für den naturwissenschaftlichen Sachunterricht in Deutschland deutlich ausbaufähig ist.

Mit Blick auf die von den Schüler:innen erreichten Kompetenzstände ist festzuhalten, dass – neben dem erwarteten Zusammenhang mit der generellen Ausstattung der Haushalte mit bildungsrelevanten Gütern und den in den Haushalten verfügbaren digitalen Medien – schul- bzw. unterrichtsbezogene Merkmale kaum Erklärungskraft besitzen. Ein Fehlschluss wäre sicherlich, wenn auf eine generelle Unwirksamkeit von digitalen Medien im Unterricht geschlossen würde. Möglicherweise fehlt es bislang an passgenauen digitalen Angeboten für den Mathematikunterricht und den naturwissenschaftsbezogenen Sachunterricht, die das Potenzial haben, Schüler:innen kognitiv zu aktivieren und zu unterstützen. Anzunehmen ist zudem, dass die Implementierung solcher Angebote auch durch Fortbildungen der Lehrkräfte begleitet werden muss.

In diesem Zusammenhang lässt sich hinsichtlich digitalisierungsbezogener Fortbildungen von Lehrkräften ein positiver Anstieg der Teilnahme verzeichnen. Im Vergleich zum Jahr 2019 sind in TIMSS 2023 die Anteile der Mathematiklehrkräfte, die an Fortbildungen zur *Integration von Informationstechnologien* teilgenommen haben, signifikant höher ausgefallen. Für die Naturwissenschaften sind die Teilnahmequoten vergleichbar mit denen von 2019. Darüber hinaus bekunden die Lehrkräfte in 2023 auch ein Interesse an solchen Fortbildungen.

1.2.4 Geschlechterunterschiede in mathematischen und naturwissenschaftlichen Kompetenzen

Geschlechterunterschiede in mathematischen und naturwissenschaftlichen Kompetenzen stehen im Fokus des sechsten Kapitels von Nonte, Grommé und Scholz. Dabei wurde als Reaktion auf gesellschaftliche Entwicklungen bei der Erfassung des Geschlechts eine dritte Antwortoption angeboten. Aufgrund sehr geringer Fallzahlen und einer eingeschränkten internationalen Vergleichbarkeit der Antwortoption können jedoch keine empirisch generalisierbaren Aussagen für diese Gruppe vorgenommen werden. Deshalb muss die Darstellung der Befunde auf Schüler:innen, die sich den Kategorien *Mädchen* oder *Junge* zugeordnet haben, beschränkt bleiben.

In TIMSS 2023 lagen die Kompetenzwerte der Jungen im Fach Mathematik im Durchschnitt 13 Punkte über denen der Mädchen. Die Trendanalysen für die Jahre 2007, 2019 und 2023 in diesem Fach zeigen für Deutschland keine signifikanten Veränderungen in den Kompetenzunterschieden zwischen den beiden Geschlechtern. Für Naturwissenschaften zeichnet sich hingegen eine Abnahme in den Differenzen zwischen Mädchen und Jungen ab, sodass in TIMSS 2023 keine signifikanten Unterschiede in der Testleistung zwischen Mädchen und Jungen mehr vorliegen. Dieser Rückgang in den Kompetenzunterschieden, die in TIMSS 2007 noch zugunsten der Jungen verzeichnet werden konnten, ist allerdings nicht auf einen Anstieg der Leistungen der Mädchen, sondern auf eine Abnahme bei den Jungen zurückzuführen.

Hinsichtlich der fachbezogenen Einstellungen und Selbstkonzepte zeigt sich, dass Mädchen in Mathematik wie in den vergangenen TIMSS-Zyklen weniger positive Einstellungen und ein niedrigeres Selbstkonzept aufweisen als Jungen. Für den Sachunterricht zeigen sich keine Unterschiede. Im Trend kann allerdings für beide Fächer eine leichte Abnahme in den durchschnittlichen fachbezogenen Einstellungen und dem Selbstkonzept von Mädchen und Jungen festgestellt werden. Hierdurch sind diese nicht mehr wie noch in TIMSS 2019 als *hoch*, sondern aktuell nur noch als *mittel* zu charakterisieren.

1.2.5 Soziale Disparitäten in den Kompetenzen von Viertklässler:innen

Die Ungleichheit von Bildungschancen resultiert in vielen Staaten aus institutionellen Bedingungen des Bildungssystems, dem Lernverhalten der Schüler:innen sowie aus den in den Familien der Schüler:innen getroffenen Bildungsentscheidungen. Dabei zeigt sich für den Grundschulbereich in Deutschland nach wie vor eine deutliche Verknüpfung zwischen der sozialen Herkunft und dem Bildungserfolg der Schüler:innen am Ende der vierten Jahrgangsstufe, die im Zeitverlauf relativ stabil ausfällt. Der Untersuchung dieser Zusammenhänge im Rahmen von TIMSS 2023 widmen sich Stubbe, Schulz und Beese im siebten Kapitel dieses Bandes.

Auch in TIMSS 2023 zeigt sich eine enge Kopplung zwischen dem sozioökonomischen Status von Familien und dem Bildungserfolg ihrer Kinder in Mathematik und den Naturwissenschaften. International betrachtet sind zwischen den teilnehmenden Staaten und Regionen deutliche Unterschiede in den Zusammenhängen zwischen den sozialen Hintergrundmerkmalen und den Kompetenzen der Viertklässler:innen zu beobachten. Im Trend betrachtet ist für Deutschland festzustellen, dass keine signifikanten Veränderungen des Zusammenhangs über die Zeit vorliegen.

Ein besonderes Maß des sozialen Hintergrundes ist der Indikator der *Armutsgefährdung*. Diese Gefährdung besteht insbesondere für Kinder aus Familien mit Migrationshintergrund und für Familien mit einem niedrigen Bildungsniveau. Werden die Kompetenzunterschiede von armutsgefährdeten mit nicht armutsgefährdeten Kindern verglichen, zeigt sich in TIMSS 2023 in Mathematik erneut eine Differenz von knapp einem Lernjahr. Bei den naturwissenschaftlichen Kompetenzen beträgt diese Differenz bis zu zwei Lernjahre. Im Trend zeigen sich seit 2007 keine signifikanten Veränderungen bei den festgestellten sozialen Disparitäten.

1.2.6 Kompetenzen von Schüler:innen mit und ohne Migrationshintergrund

In internationalen Schulleistungsstudien hat es sich bewährt, den Migrationshintergrund von Kindern anhand des Geburtslandes der Eltern zu differenzieren. Daher unterscheiden Schwippert, Jusufi, Lott und Stubbe im achten Kapitel Schüler:innen, deren Eltern beide in Deutschland geboren wurden, von Schüler:innen, bei denen ein Elternteil oder beide Elternteile im Ausland geboren wurden. In der Zusammenschau der Befunde seit TIMSS 2007 zeigt sich in 2023 ein wenig verändertes Bild: Die im Durchschnitt höchsten Kompetenzen in Mathematik und den Naturwissenschaften weisen Grundschulkinder auf, deren Eltern beide in Deutschland geboren wurden. Etwas niedriger fallen die mittleren Kompetenzstände von Schüler:innen aus, bei denen ein Elternteil im Ausland geboren wurde, und noch geringer sind die Kompetenzstände von Kindern, bei denen beide Elternteile im Ausland geboren wurden. Insgesamt sind die Differenzen bei den naturwissenschaftlichen Kompetenzen deutlich stärker ausgeprägt als in Mathematik. Diese Unterschiede variieren in Deutschland seit 2007 nur wenig und auch im internationalen Vergleich zeigen sich im Gesamteindruck keine substanziellen Entwicklungen.

1.2.7 Trends in der Schülerzusammensetzung und in den Kompetenzen

Verglichen mit 2007 ist die Schülerschaft an Grundschulen in Deutschland in TIMSS 2023 von größerer Vielfalt geprägt, wie Kasper, Sierau, Schwippert und Köller im neunten Kapitel zeigen. So ist ein Anstieg des Anteils von Kindern mit sonderpädagogischen Förderbedarfen an Regelschulen festzustellen. Auch hat sich die Schülerschaft hinsichtlich ihrer soziodemografischen Zusammensetzung, vor allem in Bezug auf die migrationsbezogenen Merkmale der Familien, verändert. In den Trendanalysen der Leistungsergebnisse ist für Deutschland im Bereich Mathematik unter Berücksichtigung der veränderten Schülerschaft keine signifikante Veränderung zu beobachten. Bei den naturwissenschaftlichen Kompetenzen zeigt sich in TIMSS 2023 ein zurückgehender Kompetenzstand, der nicht durch die veränderte Schülerschaft zu erklären ist.

Hinsichtlich der heimischen Lernsituation der Kinder während der COVID-19-Pandemie gaben viele Erziehungsberechtigte an, dass sie ihre Kinder während der Schulschließung gut beim Lernen unterstützen konnten. Jedoch berichteten auch 20 Prozent der Erziehungsberechtigten, dass sie ihre Kinder während der Schulschließung nicht gut beim Lernen unterstützen konnten. Nach Angaben der Erziehungsberechtigten konnten die meisten Viertklässler:innen während der

COVID-19-Pandemie gute häusliche Lernumgebungen (etwa einen Raum zum ungestörten Lernen oder ausreichenden Internetzugang) nutzen. Jedoch stand den Kindern ein eigenes Endgerät zur Teilnahme am Distanzunterricht nur in 64 Prozent der Fälle die ganze Zeit zur Verfügung, in 22 Prozent der Fälle nur einen Teil der Zeit. In Deutschland wurden in Zeiten von Schulschließungen alternative Unterrichtsformen wie Wechselunterricht, Distanzunterricht, Distanzlernen (asynchron und eigenständig) und Notbetreuung (mit Unterrichtsinhalten) in den Schulen realisiert. Mit Blick auf die Umsetzung des Fernunterrichts gaben Lehrkräfte und Erziehungsberechtigte an, dass diese Unterrichtsformen eher gut funktioniert haben. Unterschiedliche Wahrnehmungen gab es hinsichtlich des Kontakts zwischen Schüler:innen und Lehrkräften, der Übermittlung von Arbeitsaufträgen und Lernmaterialien sowie der Rückmeldung zu bearbeiteten Arbeitsaufträgen durch die Lehrkräfte. Bei der retrospektiven Betrachtung des Effekts der durch die COVID-19-Pandemie bedingten veränderten Lernumgebung auf die Kompetenzstände lässt sich vermuten, dass ein durch die Lehrkräfte als gut eingeschätzter Distanzunterricht kompensatorisch sowohl für die mathematischen als auch naturwissenschaftlichen Kompetenzen wirken konnte.

1.2.8 Unterrichten in der Grundschule: Unterrichtsqualität und Lehrkräftefortbildung

Bei Untersuchungen zur Unterrichtsqualität werden im deutschsprachigen Raum insbesondere die drei Basisdimensionen *Klassenführung*, *kognitive Aktivierung* und *konstruktive Unterstützung* in den Blick genommen. Diese Dimensionen stellen eine wichtige Voraussetzung für das schulische Lernen dar. Gezielte Fortbildungen können Lehrkräfte darin unterstützen, Unterrichtsqualität sicherzustellen und weiterzuentwickeln. Scholz, Beese, Jentsch, Sierau und Steffensky beschreiben im zehnten Kapitel einerseits zentrale Unterrichts(qualitäts)merkmale in Deutschland und im internationalen Vergleich und werfen andererseits einen Blick auf die Teilnahme an Lehrkräftefortbildungen sowie spezifisch für Deutschland auf Fortbildungsangebote und -bedarfe.

Es zeigt sich in TIMSS 2023, dass die Schüler:innen in Deutschland die Basisdimensionen *konstruktive Unterstützung* und *kognitive Aktivierung* für den Mathematik- und Sachunterricht ähnlich positiv einschätzen. Die *Klassenführung* (hier mit dem Fokus auf unterrichtliche Störungen) wird dagegen eher unterdurchschnittlich eingeschätzt. Die Einschätzungen hängen teilweise vom Kompetenzniveau der Schüler:innen in beiden Fächern ab: Leistungsstärkere Kinder nehmen die *Klassenführung* und *konstruktive Unterstützung* im Vergleich zu leistungsschwächeren Kindern etwas höher wahr. Für die *kognitive Aktivierung* lassen sich keine eindeutigen Tendenzen ausmachen.

Hinsichtlich des Fortbildungsangebots und dessen Nutzung durch die Lehrkräfte zeigt sich, dass die Schüler:innen in Deutschland von Lehrkräften unterrichtet werden, die im internationalen Vergleich eher selten Fortbildungen besucht haben. Im Vergleich zu TIMSS 2019 ist in 2023 dabei ein Rückgang bei der Teilnahme zu verzeichnen. Dabei wurden im Vergleich zu Mathematik für den Sachunterricht sowohl niedrigere Teilnahmequoten als auch ein stärkerer Rückgang festgestellt – trotz teilweise deutlicher Bedarfe auf Seiten der Lehrkräfte. Lediglich im Bereich *Integration von Informationstechnologien* sind gleichbleibende oder sogar höhere Teilnahmequoten zu verzeichnen.

1.2.9 Privater Nachhilfeunterricht und Förderungen zum Ausgleich pandemiebedingter Lernrückstände am Ende der Grundschulzeit

Neben dem regulären Schulunterricht erhält ein großer Anteil von Schüler:innen zusätzlich privaten Nachhilfeunterricht. Diese Verbreitung von Nachhilfeunterricht kann einerseits positiv als Engagement der Erziehungsberechtigten für gute Schulleistungen gewertet werden. Mit Blick auf internationale Trends wird seit vielen Jahren andererseits auch kritisch auf eine zunehmende Verbreitung von Nachhilfeunterricht hingewiesen, da dies auch als Zeichen eines Vertrauensverlustes in das reguläre Schulsystem interpretiert werden kann. Benckwitz, Guill und Beese untersuchen im elften Kapitel die aktuelle Situation zur Nutzung von privatem Nachhilfeunterricht in Deutschland sowie das Wissen der Lehrkräfte darüber. Sie geben darüber hinaus auch einen Einblick in die Nutzung von staatlich finanzierten Förderprogrammen zum Ausgleich negativer Folgen der COVID-19-Pandemie für Schüler:innen.

In Deutschland erhielt in TIMSS 2023 knapp jedes zehnte Kind Nachhilfeunterricht in Mathematik. Mit Blick auf den Sachunterricht spielt Nachhilfe praktisch keine Rolle in Deutschland. Es zeigt sich, dass insbesondere Schüler:innen mit ausreichenden oder schlechteren Mathematiknoten Nachhilfeunterricht bekommen. Auch wenn privater Nachhilfeunterricht einen finanziellen Aufwand für die Familien bedeutet, konnte festgestellt werden, dass die Anteile der Kinder aus armutsgefährdeten Haushalten, die Nachhilfe erhalten, höher sind als die der Kinder aus nicht armutsgefährdeten Familien. Bemerkenswert erscheint, dass die Erziehungsberechtigten die Lehrkräfte nicht immer über den Nachhilfeunterricht ihrer Kinder informieren. Dies schränkt eine gegenseitig informierte und somit zusammenwirkende Förderung der Kinder durch Nachhilfe und Regelunterricht ein – eine mögliche Synergie wird in diesen Fällen somit nicht genutzt.

In Bezug auf die Inanspruchnahme von staatlichen Förderprogrammen zum Ausgleich der pandemiebedingten Lernrückstände nahm knapp jedes vierte Kind mindestens ein Angebot in Anspruch. Schüler:innen in den untersten Kompetenzstufen, aus armutsgefährdeten Familien und mit großen subjektiv wahrgenommenen Lernverlusten durch die COVID-19-Pandemie nahmen am häufigsten an Förderprogrammen teil. Auf Schulebene kann festgestellt werden, dass weder das Einkommensniveau in der Gegend der Schule noch die Länge der Schulschließungen eine zentrale Rolle für die Nutzung der Förderprogramme spielten.

1.2.10 Schullaufbahnpräferenzen am Übergang in die Sekundarstufe und der Zusammenhang mit leistungsrelevanten und sozialen Merkmalen

Länderspezifische Reformen der Sekundarstufenstruktur haben in den letzten Jahren zu einer Heterogenität der Schulformbezeichnungen in Deutschland geführt. Gegenwärtig existieren 16 verschiedene Bezeichnungen, die in den Ländern der Bundesrepublik Deutschland teils divergierende Schularten benennen, was insbesondere in bildungsfernen Haushalten bei Schulwechseln zwischen Ländern zu Verunsicherungen führen kann. Stubbe, Beier und Lott befassen sich im zwölften Kapitel mit den Übergangsempfehlungen von Lehrkräften und den Entscheidungen der Erziehungsberechtigten nach der vierten Jahrgangsstufe in Deutschland. Der Übergang in die Sekundarstufe ist für den zukünftigen Bildungserfolg von Schüler:innen in Deutschland bedeutsam. Der Übergang von der Primar- in

die Sekundarstufe ist u. a. auf Leistungsdisparitäten zurückzuführen, die mit der familiären Herkunft verbunden sind. Zahlreiche empirische Studien belegen den Zusammenhang zwischen der sozialen Herkunft der Schüler:innen, den Entscheidungen der Erziehungsberechtigten und den Übergangsempfehlungen ihrer Lehrkräfte. Die unterschiedlichen Anspruchsniveaus der verschiedenen Bildungsgänge im Sekundarschulbereich verstärken den Einfluss der sozialen Herkunft auf die Schulleistungen.

Die seit 2011 dokumentierten Trendanalysen zum Übergang von der Grundschule in die Sekundarstufe I ergeben ein insgesamt wenig verändertes Bild. Wie für die vorangegangenen Zyklen zeigt sich in TIMSS 2023, dass vor allem leistungsrelevante Merkmale der Schüler:innen für die Schullaufbahnpräferenz der Lehrkräfte ausschlaggebend sind. Allerdings besteht auch unter Kontrolle der in TIMSS getesteten mathematischen und naturwissenschaftlichen Kompetenzen weiterhin ein deutlicher Zusammenhang zwischen Merkmalen der sozialen Herkunft und den Schullaufbahnpräferenzen der Lehrkräfte und Erziehungsberechtigten.

1.3 Bildungspolitische und didaktische Folgerungen

Nationale und internationale Schulleistungsuntersuchungen sind wichtige Informationsquellen für die Weiterentwicklung von Bildungssystemen. Die einzelnen Studien nehmen dabei spezifische Perspektiven ein, zum Beispiel im Hinblick auf einen Vergleich der Länder der Bundesrepublik Deutschland (IQB-Bildungstrend), die computer- und informationsbezogenen Kompetenzen von Jugendlichen im internationalen Vergleich (*International Computer and Information Literacy Study;* ICILS), die mathematischen, naturwissenschaftlichen und Lesekompetenzen im internationalen Vergleich von Schüler:innen, die vor dem Übergang in eine Ausbildung oder in die gymnasiale Oberstufe stehen (*Programme for International Student Assessment*; PISA) oder die mathematischen und naturwissenschaftlichen Kompetenzen von Kindern vor dem Übergang in die weiterführenden Schulen (TIMSS).

Gemein ist all diesen Untersuchungen die solide wissenschaftliche Fundierung, die Orientierung an aktuellen gesellschaftlichen und politischen Herausforderungen und die regelmäßige Wiederholung zur Dokumentation von Trends. Hierbei können Effekte, die mit besonderen Herausforderungen einhergehen, oder aber langfristige Trends, die sich einer kurzfristig wiederholten oder nur einmaligen Betrachtung entziehen, dokumentiert werden. Somit kann die wiederholte Untersuchung relevanter Merkmale des Bildungssystems zum einen zukünftige Innovationen begründen, zum anderen zurückliegende evaluieren und schließlich auch als ‚Seismograf' für bisher nicht bekannte Herausforderungen im Bildungssystem dienen.

In der vorliegenden TIMSS-Untersuchung wurden Herausforderungen für das Bildungssystem mit dem Fokus auf die Grundschule untersucht. Stand bei TIMSS 2019 noch eine sich verändernde Komposition der Schülerschaft im Fokus, so bekam im Vergleich der TIMSS-Zyklen von 2019 und 2023 die COVID-19-Pandemie eine besondere Bedeutung. Befürchtet wurde für die 2023 befragten und getesteten Schüler:innen, die einen längeren Zeitraum ihrer bis dahin noch relativ kurzen Bildungslaufbahn unter den besonderen Einschränkungen durch die COVID-19-Pandemie erlebt haben, dass diese deutliche und auch sehr differenzierte Nachteile im Vergleich zu Viertklässer:innen haben könnten, die vor 2020

nicht unter diesen Einschränkungen beschult wurden und gelernt haben. Tatsächlich zeigt sich dies in TIMSS 2023 weder für die beiden untersuchten Domänen noch für spezifische Schülergruppen.

Nachfolgend sind vor dem Hintergrund der hier vorgelegten Befunde von TIMSS 2023 Herausforderungen bzw. Handlungsbedarfe summarisch und kursorisch aufgeführt. Bei der Zusammenstellung dieser Punkte ist den Autor:innen bewusst, dass damit verbundene Reformen, Innovationen oder Entscheidungen auch immer Fragen der verfügbaren und zum Teil nur eingeschränkt einsetzbaren Ressourcen beinhalten. Dies sind Ressourcen wie die räumliche Ausstattung in den Schulen, das Vorhandensein von qualifizierten Lehrkräften, zur Verfügung stehende Budgets sowie Zeit und Belastbarkeit der verschiedenen Akteur:innen im Bildungssystem, die sich praktisch einvernehmlich – wenn auch zum Teil mit unterschiedlichen Interessenlagen – um das Wohlergehen und eine gute schulische Entwicklung der Grundschulkinder bemühen, um diese auf eine Zukunft in Schule, Beruf und Gesellschaft vorzubereiten.

Die Abwägung, welche der identifizierten Herausforderungen und Handlungsbedarfe bei zukünftigen Reformen bzw. Veränderungen priorisiert werden, obliegt nicht den an der Studie beteiligten Wissenschaftler:innen. Die schwierige Abwägung, welche der hier identifizierten Herausforderungen und Handlungsbedarfe bei zukünftigen Veränderungen unter Berücksichtigung von gesellschaftlichen Erwartungen und Rahmenbedingungen priorisiert werden, ist in erster Linie Aufgabe der Bildungspolitik.

Folgende Handlungsfelder lassen sich aufgrund der Resultate der TIMSS-2023-Untersuchung exemplarisch hervorheben:

- In den mathematischen und naturwissenschaftlichen Befunden von TIMSS 2023 zeigt sich bei der Betrachtung der Kompetenzstände ein etwas unterschiedliches Profil. Während sich in Mathematik im Trend im Mittel keine Veränderungen feststellen lassen, ist in den Naturwissenschaften seit 2007 ein negativer Trend zu verzeichnen. Dabei ist in Mathematik zu beobachten, dass in TIMSS 2023 der Anteil von Schüler:innen auf der höchsten Kompetenzstufe zwar signifikant, aber auf sehr niedrigem Niveau gestiegen ist. Für beide Domänen ist der Anteil von Schüler:innen auf der höchsten Kompetenzstufe weiterhin vergleichsweise niedrig. Gleichzeitig konnte der Anteil von Lernenden auf der untersten Kompetenzstufe nicht abgebaut werden. Die Förderung von leistungsschwachen und leistungsstarken Kindern ist sowohl für den Mathematik- als auch den naturwissenschaftlichen Sachunterricht weiterhin eine zentrale Herausforderung.
- Betrachtet man die zunehmende Heterogenität der Schülerschaft unter gleichzeitiger Berücksichtigung verschiedener Hintergrundmerkmale im Trend, so zeigt sich in Mathematik kein Unterschied in den Kompetenzen von 2007, 2019 und 2023. In den naturwissenschaftlichen Kompetenzen ist ein negativ ausgeprägter Trend zu verzeichnen. Beide Trends müssen im Blick behalten werden, um sich auch zukünftig auf Herausforderungen durch eine sich wandelnde Zusammensetzung der Schülerschaft einstellen zu können.
- Bildung für nachhaltige Entwicklung ist ein aktuelles gesellschaftliches Thema – daher wurden auch in TIMSS 2023 die nachhaltigkeitsbezogenen Einstellungen der Grundschüler:innen erfasst. Dabei konnte festgestellt werden, dass generell höhere Kompetenzen in den Naturwissenschaften mit höheren Präferenzen für den Schutz der Natur zusammenhängen. Dies zeigt, wie wich-

tig naturwissenschaftliche Bildung für umweltbewusstes Handeln ist. Bildung für Nachhaltigkeit ist demnach als Auftrag für die naturwissenschaftlichen Fächer, aber auch als überfachlicher Auftrag der Schule anzusehen.

- Die Verfügbarkeit und die Nutzung digitaler Medien in der Gesellschaft nehmen zu. Dies als Ausgangspunkt nehmend hat die KMK bereits 2020 darauf hingewiesen, dass auch in den Grundschulen die Kinder an die kompetente und zukunftsorientierte Nutzung digitaler Medien herangeführt werden sollten (KMK, 2020). Im Rahmen von TIMSS 2023 wurde gerade im Nachgang zur COVID-19-Pandemie ein ‚Digitalisierungsschub' in den Schulen erwartet. Dieser schlägt sich auch tatsächlich in der Verfügbarkeit digitaler Geräte nieder, deren Nutzung im Fachunterricht wird jedoch – und dies wird vor dem Hintergrund des internationalen Vergleichs deutlich – in Deutschland zögerlich und auf einem eher niedrigen Niveau umgesetzt. Um die Grundschüler:innen auf die globalen Herausforderungen in einer digitalisierten Welt vorzubereiten, liegt hier ein umfangreicher Handlungsbedarf vor. Dieser umfasst weiter auszubauende Aus- und Fortbildungsangebote für Lehrkräfte im Umgang mit digitalen Medien im Unterricht. Er betrifft aber auch die fachdidaktische Entwicklung verstehensorientierter digitaler Angebote für den Unterricht sowie die Unterstützung von Lehrkräften bei der Nutzung solcher digitalen Angebote. Neben der Verfügbarkeit digitaler Geräte spielt auch deren technische Einsatzfähigkeit in den Schulen eine zentrale Rolle – eine Aufgabe, die sicherlich nicht zusätzlich zu den Kernaufgaben von den Lehrkräften geleistet werden kann.
- Die sich in den letzten Jahren und Jahrzehnten wandelnde Zusammensetzung der Schülerschaft stellt eine stete Herausforderung für die Lehrkräfte und die von ihnen genutzten Unterrichtskonzepte dar. Dies gilt auch hinsichtlich der sich über die Jahre verändernden Gruppen von Schüler:innen mit und ohne Migrationshintergrund. So zeigt sich seit TIMSS 2007 ein im Wesentlichen unverändertes Bild in den Kompetenzunterschieden der beobachteten Schülergruppen. Ebenso hat sich auch der Status Deutschlands im internationalen Vergleich kaum geändert. Vor dem Hintergrund des gesellschaftlichen Anspruchs des auch von der Herkunft unabhängigen Zugangs zur Bildung gestaltet sich dieser Umstand als eine dringliche Herausforderung im Bildungssystem. Trotz jahrzehntelanger Bemühungen, durch ganz unterschiedliche Konzepte Leistungsrückstände von Kindern mit Migrationshintergrund abzubauen, zeigt sich kein substanzieller bzw. nachhaltiger Effekt. Die Herausforderung, diese Disparitäten abzubauen, wird vor dem Hintergrund des gesellschaftlichen Generationswechsels und dem damit verbundenen Fachkräftemangel noch virulenter.
- Auch wenn sich in den aktuellen Daten nur für Mathematik Unterschiede zwischen Mädchen und Jungen zeigen, sollte auch der naturwissenschaftliche Sachunterricht die differenziellen Interessenlagen von Mädchen und Jungen nicht vernachlässigt werden, da der MINT-Bereich nach wie vor ein Bereich ist, in dem Mädchen bzw. Frauen unterrepräsentiert sind.
- Angesichts der Bedeutung von motivationalen Orientierungen für die kurz- und langfristige Kompetenzentwicklung, aber auch für spätere Kurswahlen und Berufsentscheidungen ist die leichte aber stetige Abnahme der motivationalen Orientierungen im Mathematik- und naturwissenschaftsbezogenen Sachunterricht problematisch. Motivations- und interessenfördernde Ansätze im Unterricht, z. B. die Einbettung von Inhalten in relevante Kontexte und Zeiträume, um verstehensorientiertes, forschendes und explorierendes Lernen

zu ermöglichen, sind entsprechend nicht nur für die weiterführenden Schulen, sondern auch für die Grundschule relevant.

- Trotz fortdauernder Bemühungen Bildungsdisparitäten im Zusammenhang mit dem sozioökonomischen Hintergrund, der Armutsgefährdung, dem Migrationshintergrund und dem Geschlecht der Grundschüler:innen abzubauen, gelingt es seit 2007 nicht, die Leistungsdiskrepanzen in beiden Domänen nachhaltig zu verringern. Unter simultaner Berücksichtigung dieser Hintergrundvariablen zeigt sich, dass es den Lehrkräften in Deutschland nicht gelungen ist, den sich durch die wandelnde Zusammensetzung der Schülerschaft ergebenden Herausforderungen erfolgreich zu begegnen. Mit Blick auf Bildungsentscheidungen im Zuge des Übergangs von der Grundschule in weiterführende Schulen wird virulent, dass möglichen verallgemeinernden, übergeneralisierenden Zuschreibungen im Zusammenhang mit bestimmten Gruppen von Schüler:innen zu begegnen ist. Dabei sollte eine alltagstaugliche Diagnostik sowie ein handhabbares und wirkungsvolles Förderrepertoire entwickelt werden, damit die Unterstützung der Grundschüler:innen so zielgerichtet erfolgen kann, wie sie erforderlich ist.

Literatur

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2004). *Bildungsstandards der Kultusministerkonferenz. Erläuterungen zur Konzeption und Entwicklung.* https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Konzeption-Entwicklung.pdf

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2020). *Ländervereinbarung über die gemeinsame Grundstruktur des Schulwesens und die gesamtstaatliche Verantwortung der Länder in zentralen bildungspolitischen Fragen*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2020/2020_10_15-Laendervereinbarung.pdf

Schwippert, K., Kasper, D., Köller, O., McElvany, N., Selter, C., Steffensky, M., & Wendt, H. (Hrsg.). (2020). *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Waxmann. https://doi.org/10.31244/9783830993193

Stanat, P., Schipolowski, S., Schneider, R., Weirich, S., Henschel, S., & Sachse, K. A. (Hrsg.). (2022). *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich*. Waxmann. https://doi.org/10.31244/9783830996064

von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U., & Yin, L. (2024). *TIMSS 2023 international results in mathematics and science*. TIMSS & PIRLS International Study Center.

Kapitel 2
Ziele, Anlage und Durchführung der *Trends in International Mathematics and Science Study* (TIMSS) 2023

Christin Beese, Luise A. Scholz, Donieta Jusufi, Daniel Kasper und Knut Schwippert

2.1 Einleitung

Die *Trends in International Mathematics and Science Study* (TIMSS) ist eine internationale Schulleistungstudie, die den Schwerpunkt auf die Untersuchung der mathematischen und naturwissenschaftlichen Kompetenzen von Schüler:innen verschiedener Staaten und Regionen unter Berücksichtigung curricularer Vorgaben und zentraler Rahmenbedingungen schulischer und familiärer Lernumgebungen legt. Gemeinsam mit dem *Programme for International Student Assessment* (PISA) und der *Internationalen Grundschul-Lese-Untersuchung/ Progress in International Reading Literacy Study* (IGLU/PIRLS) ist sie Teil der Gesamtstrategie der *Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland* (KMK) zum Bildungsmonitoring in Deutschland (KMK, 2016). Bei dieser geht es vor allem darum, langfristige Entwicklungen in Deutschland zu dokumentieren, zu interpretieren und Entscheidungshilfen für mögliche zukünftige Weiterentwicklungen des Bildungssystems zu bieten.

Die TIMSS-Untersuchung bietet alle vier Jahre die Möglichkeit, die mathematischen und naturwissenschaftlichen Kompetenzen von Schüler:innen am Ende der Grundschulzeit (i. d. R. in der 4. Klasse) und gegen Ende der Pflichtschulzeit (i. d. R. in der 8. Klasse) zu erfassen. Dies ermöglicht gleichzeitig, dass Staaten, die beide Altersgruppen untersuchen, in der achten Klassenstufe die gleiche Schülerkohorte befragen wie vier Jahre zuvor. Für Deutschland wurde mit Blick auf die parallele Teilnahme an PISA entschieden, nur mit der vierten Klassenstufe an TIMSS teilzunehmen, um eine Doppelbelastung der Bildungseinrichtungen in der Sekundarstufe zu vermeiden. Deutschland beteiligt sich seit 2007 in dieser Form an TIMSS, weshalb sich die Ausführungen in diesem Band vorrangig auf die Grundschuluntersuchung in TIMSS beziehen. Die TIMS-Studie bietet die Möglichkeit, die wiederholten Untersuchungen im Trend systematisch zu betrachten und Entwicklungen über einen längeren Zeitraum hinweg abzubilden.

Mithilfe wissenschaftlicher (Vergleichs-)Untersuchungen werden den verschiedenen Akteur:innen im Bildungssystem aktuelle und relevante Informationen bereitgestellt, auf deren Basis konkrete Pläne und Strategien zur Bewältigung von Herausforderungen im Bildungssystem entwickelt werden können.

In diesem Kapitel werden zunächst die Organisation sowie die beteiligten Institutionen in TIMSS 2023 vorgestellt, um anschließend auf die Ziele, die Anlage und die Durchführung der TIMSS-2023-Untersuchung einzugehen und wesentliche Merkmale der Studie zu dokumentieren. Zum Abschluss dieses Kapitels werden technische und empirische Voraussetzungen für die Auswertung sowie einige Hinweise zur Ergebnisdarstellung in diesem Band vorgestellt.

2.2 Forschungsnetzwerk in TIMSS 2023

Zur Realisierung groß angelegter Schulvergleichsuntersuchungen wie TIMSS ist ein komplexes internationales sowie nationales Forschungsnetzwerk nötig. Initiiert wurde die TIMSS-Untersuchung von der *International Association for the Evaluation of Educational Achievement* (IEA). Die IEA ist ein gemeinnütziger Verein, der seit 1958 Erfahrung in der Konzeption, Durchführung und Auswertung von internationalen Systemmonitoringstudien hat. Auf internationaler Ebene wird die TIMS-Studie vom *TIMSS & PIRLS International Study Center* (ISC) am Boston College verantwortet. Das ISC ist verantwortlich für die internationale Koordination und Implementierung der Studie, das Studiendesign, die Entwicklung der Test- und Befragungsinstrumente sowie für die Sicherstellung der Qualität der Datenerhebung. Darüber hinaus wird am ISC die internationale Skalierung der Tests durchgeführt und der internationale Ergebnisbericht verfasst. In Zusammenarbeit mit dem *Research Triangle Institute* (RTI International) und der Abteilung für Stichprobenziehung der internationalen Abteilung der IEA Hamburg werden zudem die Stichprobenziehung und Dokumentation sowie die Berechnung internationaler Stichprobengewichte realisiert.

Auf nationaler Ebene liegt die wissenschaftliche Gesamtleitung und Koordination des TIMSS-Konsortiums bei Prof. Dr. Knut Schwippert und Prof. Dr. Mirjam Steffensky (Universität Hamburg), unterstützt durch ein Projektleitungsteam bestehend aus Dr. Daniel Kasper im Bereich Methoden sowie Dr. Christin Beese und Luise A. Scholz im Bereich Administration und Organisation. Dabei haben die wissenschaftlichen Mitarbeiter:innen Donieta Jusufi, Celestine Kleinesper, Neele Krüger, Julia Levin, Martje Lott, Jakob Sierau und Catharina Tippe wesentlich zum Gelingen des Projekts und der nationalen Berichtslegung beigetragen. Darüber hinaus wurde das Projekt während der gesamten Laufzeit von vielen engagierten studentischen Mitarbeiter:innen unterstützt.[1]

Das nationale TIMSS-2023-Konsortium besteht aus einschlägig ausgewiesenen Wissenschaftler:innen, die ihre unterschiedlichen Expertisen aus der Erziehungswissenschaft, der Empirischen Bildungsforschung, der Pädagogischen Psychologie, der Mathematikdidaktik und der Didaktik des Sachunterrichts in die Konzeption, Durchführung und Auswertung der TIMSS-Untersuchung einbringen. Im TIMSS-2023-Konsortium sind als Verbundpartner:innen Prof. Dr. Olaf Köller (Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathe-

1 Unser Dank gilt hierbei Asha Baskaran, Sebastian Brunsmeier, Angelika Dembek, Thuy Dinh, Samuel Drößler, Isabelle Fahr, Carlotta Gemkow, Fabian Hellmold, Tobias Knickrehm, Julia Koplin, Mara Krauß, Maximilian Normann, Marisa Urban, Sümerya Uzun und Verena Zimmer.

matik – IPN Kiel), Prof. Dr. Christoph Selter (Technische Universität Dortmund) und Prof. Dr. Mirjam Steffensky (Universität Hamburg) beteiligt. Als weitere Mitantragsteller:innen wirken darüber hinaus Prof. Dr. Birgit Eickelmann (Universität Paderborn), Prof. Dr. Frank Goldhammer (DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation in Frankfurt), Prof. Dr. Aiso Heinze (IPN Kiel), Dr. Ulf Kröhne (DIPF), Prof. Dr. Tobias C. Stubbe (Georg-August-Universität Göttingen), JProf. Dr. Daniel Walter (Technische Universität Dortmund) und Prof. Dr. Heike Wendt (Karl-Franzens-Universität Graz, Österreich) mit. Darüber hinaus war die nationale Abteilung der IEA Hamburg als Sub-Kontraktor für die Testorganisation und -durchführung in den Grundschulen zuständig und hat die nationale Leitung bei der Vorbereitung der Test- und Befragungsinstrumente und den Genehmigungsverfahren in den Ländern der Bundesrepublik Deutschland unterstützt.

Ermöglicht wurde die Beteiligung Deutschlands an der internationalen TIMSS-Untersuchung mithilfe der Finanzierung durch das *Bundesministerium für Bildung und Forschung* (BMBF) und die KMK. Ebenso übernahmen auch die in Deutschland an der Durchführung der TIMSS-Untersuchung beteiligten Universitäten Anteile der Finanzierung.

2.3 Ziele der TIMSS-Untersuchung

Im Fokus der TIMSS-Untersuchung steht die Erfassung der Kompetenzen von Schüler:innen in den Domänen Mathematik und Naturwissenschaften. Unter Kompetenzen von Schüler:innen werden kognitive Leistungsdispositionen verstanden, die im Sinne einer Problemlösefähigkeit als Reaktion auf an sie gestellte Anforderungen aufgefasst werden können (Weinert, 2001). Um die Kompetenzen der Viertklässler:innen angemessen kontextualisieren zu können, werden zusätzlich zahlreiche Informationen zu den Lehr- und Lernbedingungen im schulischen und häuslichen Umfeld anhand von Hintergrundfragebögen für Schüler:innen, Lehrkräfte, Erziehungsberechtigte und Schulleitungen erhoben.

Der grundlegende Ansatz von TIMSS ist dabei, die Schüler:innen in ihren jeweiligen Klassen, Schulen und Bildungsgängen als Repräsentant:innen der gesamten Schülerschaft eines Jahrgangs zu betrachten. Es geht somit ausdrücklich nicht um eine möglichst präzise Beschreibung der einzelnen Schüler:innen im Sinne einer Individualdiagnostik.

2.3.1 Theoretische Grundlagen

Die von der IEA durchgeführten internationalen Schulleistungsstudien folgen in ihrer theoretischen Rahmung traditionell einem Input-Prozess-Output-Modell. Dieses umfasst zur Beschreibung der Curricula der teilnehmenden Staaten drei verschiedene Ebenen:

1. Input: Das intendierte Curriculum entspricht den generellen curricularen Vorgaben seitens Politik bzw. Administration.
2. Prozess: Das potenzielle Curriculum beschreibt die möglichen Lehrinhalte, die z. B. in Schulbüchern dokumentiert sind, während das implementierte Curriculum die tatsächlichen Inhalte des Unterrichts umfasst.
3. Output: Das realisierte Curriculum schließlich beschreibt, was von den Schüler:innen tatsächlich gelernt wurde.

Mit diesem Modell wird angestrebt, die Teilnehmer hinsichtlich curricularer Vorgaben, der Umsetzung dieser im Unterricht und der von den Schüler:innen erreichten Kompetenzen zu vergleichen. Durch die repräsentative Datengrundlage leisten Analysen im Rahmen von TIMSS einen wichtigen Beitrag (1) zu Erkenntnissen bezüglich Steuerungs- und Entscheidungsprozessen im Bildungssystem, (2) zur Theoriebildung unterrichtlicher Lehr-Lernprozesse und (3) zur Verbesserung pädagogischer Praxis im Bereich Mathematik und Naturwissenschaften.

2.3.2 Hauptanliegen der TIMSS-Untersuchung

Für die beteiligten Staaten bietet die TIMS-Studie im Wesentlichen drei Untersuchungsperspektiven:

1. Aufgrund der Beteiligung zahlreicher Staaten weltweit können verschiedene Bildungssysteme verglichen werden. Dabei nehmen an TIMSS einerseits Staaten teil, deren Bildungssysteme hohe Ähnlichkeiten aufweisen und die somit als Ideengeber für das eigene Bildungssystem dienen können. Andererseits beteiligen sich aber auch Staaten, deren Bildungssysteme sehr unterschiedlich konzipiert sind und so kontrastreiche Vergleiche ermöglichen.
2. Durch die regelmäßige Wiederholung der Untersuchung können für die an TIMSS teilnehmenden Staaten Aussagen über Entwicklungen in deren Bildungssystemen über einen längeren Zeitraum hinweg getroffen werden.
3. Auf nationaler Ebene ermöglicht TIMSS zudem die Untersuchung spezifischer, aktueller Herausforderungen im eigenen Bildungssystem. So können denjenigen Informationen zur Verfügung gestellt werden, die für die Gestaltung des Bildungssystems zuständig sind.

Internationale Vergleiche

In TIMSS sind Vergleiche von mathematischen und naturwissenschaftlichen Kompetenzen sowie von Hintergrundmerkmalen und Rahmenbedingungen zwischen den teilnehmenden Staaten möglich. An der Grundschuluntersuchung von TIMSS 2023 beteiligten sich insgesamt 58 Staaten sowie fünf Benchmark-Teilnehmer (siehe Abbildung 2.1).[2] Zur allgemeinen Bestandsaufnahme ist es hilfreich, viele Staaten mit in den Blick zu nehmen. Damit dies im Rahmen von internationalen Vergleichsstudien möglich ist, ist sicherzustellen, dass die herangezogenen Informationen und Erhebungsinstrumente über die verschiedenen Staaten auch tatsächlich vergleichbar sind. Sollen zudem Vergleiche mit Staaten, die einen ähnlichen historischen, kulturellen oder ökonomischen Hintergrund aufweisen, vorgenommen werden oder mit Staaten, die zu gleichen Zeitpunkten wiederholt an TIMSS teilgenommen haben, ist eine entsprechende Auswahl zu treffen.[3]

2 Zusätzlich haben an der Grundschuluntersuchung von TIMSS 2023 auch Irak und Kurdistan (Region Kurdistan-Irak) teilgenommen. Da für diese jedoch keine international vergleichbaren Daten vorliegen, werden sie im Folgenden nicht in den Ausführungen und Ergebnissen berücksichtigt.

3 Weiterführende Informationen zu den Bildungssystemen der teilnehmenden Staaten können der TIMSS-Enzyklopädie entnommen werden (Reynolds et al., 2024), für eine Übersicht über das Bildungssystem in Deutschland siehe Beese et al. (2024).

Im vorliegenden Berichtsband zu TIMSS 2023 werden dazu Listen mit Vergleichsgruppen (VG) gebildet. Es finden sich

1. Vergleiche mit allen Staaten und Regionen, die an TIMSS 2023 beteiligt waren (VG 1),
2. Vergleiche mit Staaten und Regionen, die der OECD oder der EU angehören und/oder in der jeweiligen Domäne signifikant besser als Deutschland sind bzw. sich nicht signifikant vom Mittelwert Deutschlands unterscheiden (VG 2) und
3. Vergleiche mit allen Staaten und Regionen, die seit 2007 alle vier Jahre mit der vierten Klassenstufe an TIMSS teilgenommen haben (VG 3).

Zusätzlich wird in einigen Abbildungen jeweils ein gemeinsamer Mittelwert für alle OECD- bzw. EU-Staaten als Referenzwert berichtet. In Tabelle 2.1 sind die teilnehmenden OECD- und EU-Staaten aufgeführt.

Tabelle 2.1: Vergleichsgruppe OECD und EU in TIMSS 2023

VG_{EU} (n = 22)		VG_{OCED} (n = 29)	
Belgien (Fläm. Gem.)	Schweden	Australien	Neuseeland
Belgien (Frz. Gem.)	Slowakei	Belgien (Fläm. Gem.)	Niederlande
Bulgarien	Slowenien	Belgien (Frz. Gem.)	Norwegen
Dänemark	Spanien	Chile	Polen
Deutschland	Tschechien	Dänemark	Portugal
Finnland	Ungarn	Deutschland	Republik Korea (Südkorea)
Frankreich	Zypern	England	Schweden
Irland		Finnland	Slowakei
Italien		Frankreich	Slowenien
Lettland		Irland	Spanien
Litauen		Italien	Tschechien
Niederlande		Japan	Türkei
Polen		Kanada	Ungarn
Portugal		Lettland	USA
Rumänien		Litauen	

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Trendvergleiche

Die Fortsetzung des bereits bestehenden Trends bildet die zweite Untersuchungsperspektive in TIMSS. Damit können die Kompetenzen verschiedener Schülergruppen und deren Entwicklungen im Trend langfristig zur Verfügung gestellt werden. Darüber hinaus werden deskriptive Statistiken zur Gestaltung von Schule und Unterricht und deren Zusammenhang mit den erfassten Kompetenzen der Schüler:innen dokumentiert.

Bei der Betrachtung eines komplexen Systems, wie es die Schule darstellt, ermöglicht diese Perspektive die analytische Trennung zwischen kurzfristig bedingten Unterschieden in Schule und Unterricht sowie langfristig angelegten Reformen und Förderlinien. Dabei bietet die internationale Vergleichsperspektive in TIMSS wichtige Anhaltspunkte, da eine Vielzahl von Staaten mitunter gleichermaßen von globalen oder regionalen politischen, ökonomischen und humani-

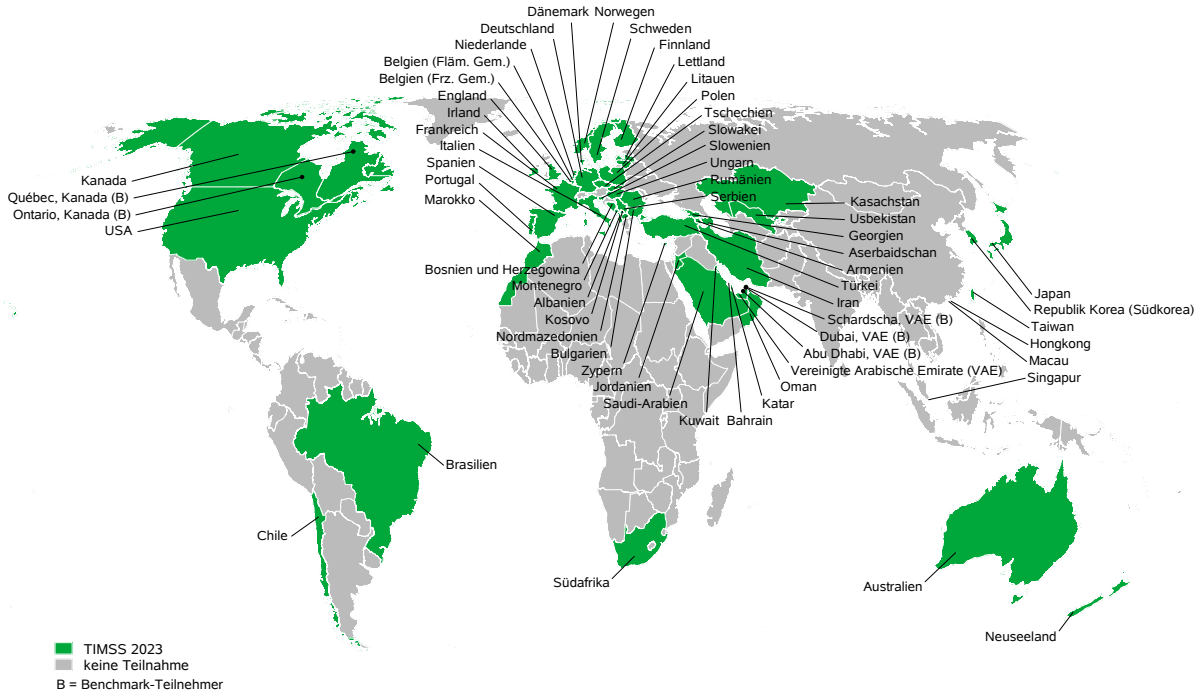
tären Veränderungen und Herausforderungen betroffen gewesen ist, diesen jedoch im Bildungssystem mit unterschiedlichen Maßnahmen begegnete.

Mit der erneuten Teilnahme Deutschlands an TIMSS bietet sich die Möglichkeit, den Trend seit 2007 in den mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässler:innen bis 2023 weiterzuführen. Dadurch können Aussagen über Unterschiede und Entwicklungen in Bezug auf Merkmale des Bildungssystems über nunmehr 16 Jahre getroffen werden. In dieser Zeit stellten beispielsweise der demografische Wandel, regionale Bevölkerungsverschiebungen (z. B. Konzentration in Ballungsgebieten) und auch der ansteigende Anteil von Schüler:innen mit Migrationshintergrund und Fluchterfahrungen ständig neue Anforderungen an das Bildungssystem in Deutschland. Aspekte dieser Entwicklungen, soweit sie im Rahmen von Studien wie TIMSS erfasst werden können, gilt es nachzuzeichnen und daraus Impulse für Weiterentwicklungen zu generieren. Individuelle Entwicklungen von Schüler:innen können und sollen – anders als bei Panelstudien, in denen Proband:innen im zeitlichen Verlauf mehrfach befragt werden – mit wiederholten Querschnittserhebungen, wie sie in TIMSS realisiert werden, nicht in den Blick genommen werden.

In diesem Band werden bei den Analysen der Trenduntersuchungen die Staaten, die sich – wie auch Deutschland – an TIMSS 2007, 2019 und 2023 beteiligt haben, einbezogen. Anders als in den vorherigen TIMSS-Berichtsbänden werden damit nicht mehr durchgängig alle Zyklen dargestellt, stattdessen wird der Fokus neben der aktuellen Erhebung auf die erste (TIMSS 2007) sowie die unmittelbar vorangegangene Erhebung (TIMSS 2019) gelegt, um Wiederholungen zu vermeiden.[4] In Abbildung 2.1 sind die an TIMSS 2023 teilnehmenden Staaten und Regionen aufgeführt und in Abbildung 2.2 die Teilnehmerstaaten, die für Trendvergleiche herangezogen werden. In Abbildung 2.2 wird entsprechend der zuvor beschriebenen Auswahl der Fokus auf Staaten mit einer Teilnahme in TIMSS 2007, 2019 und 2023 gelegt. Diese Beschränkung hat jedoch keinen Einfluss auf die Anzahl der im Trend dargestellten Staaten und Regionen, da diese durchgängig auch an TIMSS 2011 und 2015 beteiligt waren. Armenien, Kasachstan, Katar, Kuwait, Marokko und Norwegen werden nicht in Trendvergleichen aufgeführt, da eine Vergleichbarkeit zwischen den Zyklen aufgrund von sehr starken Veränderungen in den Übersetzungen in den Testinstrumenten oder den Erhebungsbedingungen nicht gegeben ist (siehe Anhang).

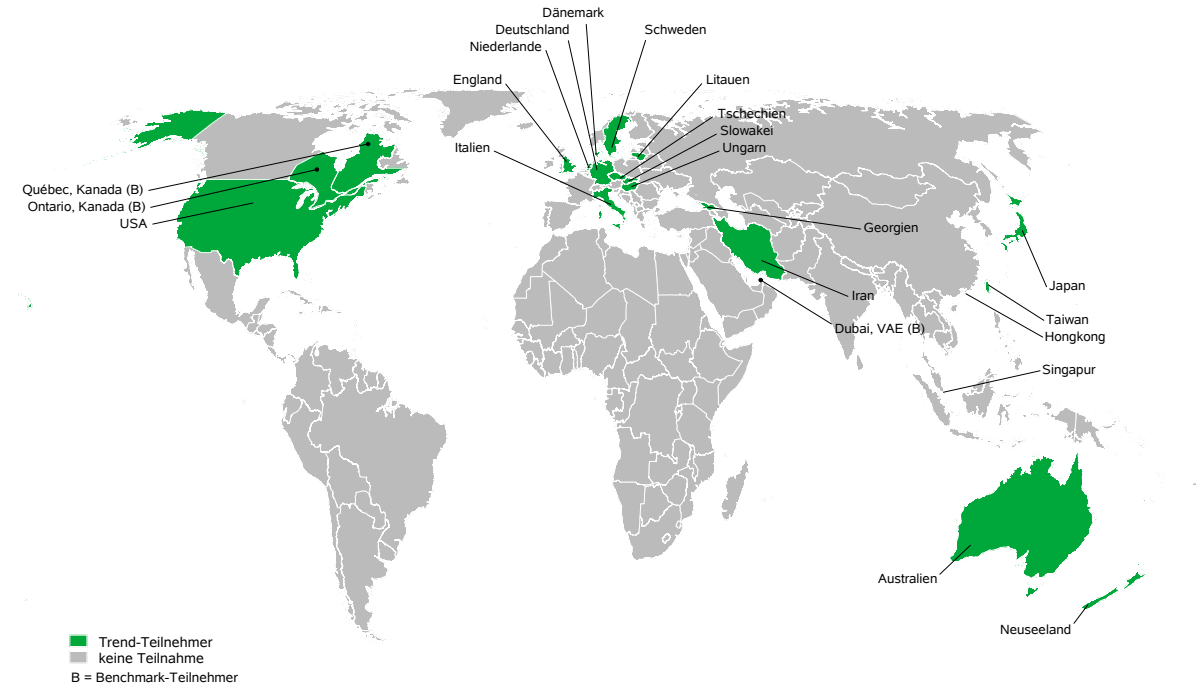
4 Die Ergebnisse zu TIMSS 2011 und TIMSS 2015 können in den jeweiligen Berichtsbänden oder im Überblick im Berichtsband von TIMSS 2019 nachgelesen werden (Bos et al., 2012; Schwippert, Kasper et al., 2020; Wendt et al., 2016).

Abbildung 2.1: Staaten und Regionen, die in TIMSS 2023 teilgenommen haben



IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Abbildung 2.2: Teilnehmer an TIMSS 2007, 2019 und 2023



IEA: Trends in International Mathematics and Science Study © TIMSS 2023

2.3.3 Inhaltlicher Schwerpunkt und Forschungsfragen in TIMSS 2023

Nachdem in TIMSS 2019 in Deutschland der Schwerpunkt auf die Betrachtung von Unterrichtsqualität in der Grundschule gelegt wurde, stehen als Reaktion auf eine sich stetig wandelnde Gesellschaft und Umwelt digitale Medien im Fokus von TIMSS 2023 (siehe Kapitel 5 in diesem Band). Diese sind Realität im Alltag von Viertklässler:innen in Deutschland und gewinnen auch in Schule und Unterricht immer mehr an Bedeutung. Mit dieser Schwerpunktsetzung wird gleichzeitig dem Umstand Rechnung getragen, dass 2023 parallel zu TIMSS auch die *International Computer and Information Literacy Study* (ICILS) stattfand. Somit bietet sich für die weiterführende Forschung die besondere Gelegenheit, die unterrichtsbezogene Nutzung von digitalen Medien sowohl in Grund- als auch in Sekundarschulen parallel zu betrachten.

Für TIMSS 2023 ergeben sich folgende übergeordnete Forschungsfragen:

- Welche Kompetenzen zeigen Schüler:innen in den Domänen Mathematik und Naturwissenschaften am Ende der vierten Jahrgangsstufe in Deutschland im Jahr 2023? Welche Ergebnisse erzielen die Viertklässler:innen in Deutschland dabei in den domänenspezifischen Inhaltsbereichen und den kognitiven Anforderungsbereichen der TIMS-Studie? Wie lassen sich die Ergebnisse im internationalen Vergleich bewerten? Wie haben sich die Ergebnisse seit 2007 und 2019 verändert? Welche Trends zeichnen sich ab?
- Wie lassen sich die TIMSS-Leistungskennwerte auf den Kompetenzstufen einordnen? Wie groß sind die Gruppen der auffällig leistungsstarken und leistungsschwachen Schüler:innen? Wie unterscheiden sich die Ergebnisse zwischen TIMSS 2007, 2019 und 2023?
- Welche Bedeutung haben die individuellen Lernvoraussetzungen und Kontextfaktoren für Kompetenzen in den Bereichen Mathematik und Naturwissenschaften? Welche Veränderungen von Lehr- und Lernbedingungen lassen sich seit 2007 beobachten?
- Wie werden digitale Medien in den Fächern Mathematik und Sachunterricht im internationalen Vergleich und in Deutschland genutzt? Welche Veränderungen sind im Vergleich zu vorherigen Zyklen festzustellen? Welche digitalen Ressourcen sind im häuslichen und schulischen Umfeld von Viertklässler:innen vorhanden?

2.4 Anlage der Studie

Um groß angelegte Schulleistungsstudien durchführbar sowie für viele Staaten finanzierbar zu halten und darüber hinaus auch die Inanspruchnahme der beteiligten Personen im Zusammenhang mit der Testdurchführung in den Schulen soweit es geht zu reduzieren, haben sich zwei Herangehensweisen bei der Konzeption der Studie bewährt: Zum einen werden Schülerstichproben gezogen, die so groß sind wie nötig, um die internationale Vergleichbarkeit sicherzustellen, aber auch so klein wie möglich, um die Belastung für die Schulen möglichst gering zu halten. Zum anderen werden systematisch entwickelte Testaufgaben zur Feststellung mathematischer und naturwissenschaftlicher Kompetenzen der Viertklässler:innen eingesetzt. Dabei wird jedem befragten Kind nur eine begrenzte Auswahl an Testaufgaben vorgelegt, um den Umfang der Testsitzungen zu minimieren.

Im Folgenden werden die Stichprobenziehung sowie die eingesetzten Test- und Befragungsinstrumente in TIMSS näher beleuchtet.

2.4.1 Stichprobenziehung

Damit die Zielpopulationen in den teilnehmenden Staaten in Studien wie TIMSS verglichen werden können, wird sie einheitlich anhand einer Systematisierung definiert, die unabhängig von den Bildungssystemen, den Bildungsgängen und den Klassenstufenbeschreibungen der jeweiligen Staaten ist. In TIMSS wird dazu die *International Standard Classification of Education* (ISCED) nach der Definition des UNECSO-Instituts von 2011 (UNESCO Institute for Statistics, 2012) genutzt, die sich an der Anzahl der formalen Unterrichtsjahre orientiert.

Die Definition der Zielpopulation der TIMSS-Grundschuluntersuchung gibt vor, dass die Kinder des vierten regulären Unterrichtsjahres befragt werden sollen – dies entspricht in Deutschland der vierten Klassenstufe. Zwar werden im Rahmen von TIMSS auf internationaler Ebene auch Schüler:innen höherer Klassenstufen befragt. Da sich Deutschland jedoch nur an der Grundschuluntersuchung beteiligt, werden die Aspekte von TIMSS, die sich auf die Untersuchung der Sekundarstufe beziehen, hier nicht näher beleuchtet.

Neben der einheitlichen Definition der Zielpopulation als Bedingung für die Vergleichbarkeit der beteiligten Staaten gibt es in TIMSS eine weitere wichtige Voraussetzung, die die Stichprobe erfüllen muss: Damit generalisierte Aussagen über die beteiligten Bildungssysteme getroffen werden können, müssen die in den teilnehmenden Staaten erhobenen Daten repräsentativ für die Zielpopulationen der Staaten sein. Daher kommt der Stichprobenziehung in TIMSS, die über ein systematisches Ziehungsverfahren[5] erfolgt, eine besondere Bedeutung zu. Für eine Berücksichtigung der Besonderheiten der teilnehmenden Bildungssysteme werden dabei unterschiedliche Stratifizierungsmerkmale genutzt. In Deutschland betrifft dies z. B. die Struktur des föderalen Bildungssystems und der 16 beteiligten Länder der Bundesrepublik Deutschland. Bei der Stichprobenziehung wurde deshalb in Deutschland mehrstufig vorgegangen: In einem ersten Schritt wurden zunächst für jedes Land der Bundesrepublik Deutschland Schulen per Zufallsverfahren unter Wahrung der Repräsentativität unterschiedlicher Schulgrößen ausgewählt. Gab es innerhalb der Schulen mehr als eine vierte Klasse, wurde in einem zweiten Schritt eine davon per Zufall ausgewählt. Da sich die Ziehungswahrscheinlichkeiten der einzelnen Schüler:innen je nach Land oder Schule unterscheiden können und damit die angestrebte repräsentative Stichprobe leicht verzerrt werden würde, werden zudem für die späteren Analysen Stichproben-Korrekturfaktoren (Gewichte) berechnet. Analysen unter Nutzung dieser Gewichte ermöglichen es, die in der Grundgesamtheit bestehenden Verteilungen auch tatsächlich bei den Auswertungen abzubilden.[6] Unabhängig von den spezifischen nationalen Designs der an TIMSS teilnehmenden Staaten ist durch die internationale Studienleitung vorgegeben, dass pro Staat mindestens 150 Schulen sowie mindestens 4 000 Schüler:innen für die Untersuchung ausgewählt werden sollten.

---

5 Das Ziehungsverfahren ist eine Kombination aus *probability proportional by size* (PPS), systematischer Zufallsauswahl (*random start – equal interval*), stratifiziertem Clustersample und mehrstufiger Stichprobe (Martin et al., 2016; Särndal et al., 1992).

6 Alle in diesem Band berichteten Ergebnisse basieren – soweit nicht anders gekennzeichnet – auf gewichteten Daten.

Die föderale Struktur im Bildungssystem führt in Deutschland darüber hinaus dazu, dass die Untersuchung in allen 16 Ländern der Bundesrepublik Deutschland separat beantragt und genehmigt werden muss. Dies hat zur Folge, dass sich die Verpflichtungsgrade zur Teilnahme an der TIMS-Studie zwischen den Ländern der Bundesrepublik Deutschland unterscheiden können. Während sich die Kultusministerien der Länder auf eine verpflichtende Teilnahme an den Kompetenztests zur Generierung einer für Deutschland repräsentativen Stichprobe verständigt haben, variieren die Verpflichtungsgrade der Teilnahme an der Erhebung der Hintergrundfragebögen für Schüler:innen, Lehrkräfte, Erziehungsberechtigte und Schulleitungen. In Tabelle 2.2 sind die verschiedenen Regelungen der Teilnahme an den Erhebungen in den einzelnen Ländern, die sich im Rahmen von TIMSS 2023 ergeben haben, aufgeführt. Die Verpflichtungsgrade der vorherigen TIMSS-Zyklen sind zusammenfassend im Berichtsband von TIMSS 2019 aufgeführt (Schwippert, Scholz et al., 2020).

Tabelle 2.2: Länderspezifische Modalitäten der Teilnahme an TIMSS 2023

	Verpflichtungsgrade der Teilnahme bei Erhebungen an öffentlichen Schulen				
Länder	**Leistungstest**	**Schüler-fragebogen**	**Lehrkräfte-fragebogen**	**Schulleitungs-fragebogen**	**Eltern-fragebogen**
Baden-Württemberg	verpflichtend	verpflichtend	verpflichtend	verpflichtend	freiwillig
Bayern	verpflichtend	freiwillig	freiwillig	freiwillig	freiwillig
Berlin	verpflichtend	verpflichtend	teilverpflichtend	teilverpflichtend	freiwillig
Brandenburg	verpflichtend	verpflichtend	verpflichtend	verpflichtend	freiwillig
Bremen	verpflichtend	verpflichtend	teilverpflichtend	teilverpflichtend	freiwillig
Hamburg	verpflichtend	freiwillig	teilverpflichtend	teilverpflichtend	freiwillig
Hessen	verpflichtend	verpflichtend	teilverpflichtend	teilverpflichtend	freiwillig
Mecklenburg-Vorpommern	verpflichtend	verpflichtend	verpflichtend	verpflichtend	freiwillig
Niedersachsen	verpflichtend	verpflichtend	verpflichtend	verpflichtend	freiwillig
Nordrhein-Westfalen	verpflichtend	freiwillig	verpflichtend	verpflichtend	freiwillig
Rheinland-Pfalz	verpflichtend	freiwillig	teilverpflichtend	teilverpflichtend	freiwillig
Saarland	verpflichtend	freiwillig	teilverpflichtend	teilverpflichtend	freiwillig
Sachsen	verpflichtend	freiwillig	freiwillig	freiwillig	freiwillig
Sachsen-Anhalt	verpflichtend	verpflichtend	verpflichtend	verpflichtend	freiwillig
Schleswig-Holstein	verpflichtend	verpflichtend	verpflichtend	verpflichtend	freiwillig
Thüringen	verpflichtend	verpflichtend	verpflichtend	verpflichtend	freiwillig

Bei einer Teilverpflichtung sind Fragen zur Person freiwillig zu beantworten, die Fragen zu Schule und Unterricht hingegen sind verpflichtend.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Auch wenn für Deutschland im Rahmen der TIMS-Studie länderspezifische Stichproben gezogen wurden, sind diese in den meisten Fällen zu klein, um empirisch verlässliche Ländervergleiche vornehmen zu können. Solche Vergleiche bietet die speziell darauf ausgelegte Bildungstrendstudie des *Instituts zur Qualitätsentwicklung im Bildungswesen* (IQB) in Berlin, die jedoch wiederum keine internationalen Vergleiche ermöglicht.

Optimalerweise nehmen bei der Durchführung der TIMSS-Untersuchung auch tatsächlich alle Schüler:innen der gezogenen Stichprobe teil (volle Ausschöpfung der Stichprobe). Abweichungen in der geplanten Stichprobe können sich jedoch

beispielsweise durch kurzfristige Schulschließungen oder durch Krankheit einzelner Schüler:innen am Testtag ergeben. Größeren Ausfällen wird durch Nachtestungen an denselben Schulen zu einem späteren Zeitpunkt oder der Erhebung an vorher systematisch ausgewählten sogenannten Ersatzschulen begegnet. Geringe Abweichungen von der geplanten Stichprobe sind in einem von der internationalen Studienleitung vorgegebenen Rahmen akzeptabel, wenn dadurch die Repräsentativität der Untersuchung nicht gefährdet wird (von Davier et al., 2024). Sind diese Abweichungen jedoch zu groß, ist dies bei den Analysen und der Berichtslegung zu berücksichtigen. Die Ausschöpfungsquoten für Deutschland sowie weitere Stichprobencharakteristika von TIMSS 2007, 2019 und 2023 sind in Tabelle 2.3 aufgeführt.

Tabelle 2.3: Beteiligungszahlen und zentrale Stichprobencharakteristika in TIMSS 2007, 2019 und 2023

	TIMSS 2007	TIMSS 2019	TIMSS 2023
Schulen			
Anzahl	246	203	230
Beteiligung am Schulleitungsfragebogen (%)	95	93	87
Klassen			
Anzahl	246	211	234
Lehrkräfte			
Anzahl (Mathematik/Sachunterricht)	240/231	216/218	244/243
Beteiligung am Lehrkräftefragebogen (%)	95	91	87
Schüler:innen			
Anzahl	5200	3437	4442
Durchschnittsalter	10.4	10.4	10.4
Mädchen/Jungen/anderes (%)	49/51/-	50/50/-	48/50/2
Kinder aus armutsgefährdeten Familien (%)	34	22	20
Kinder mit Migrationshintergrund (%) (beide Elternteile im Ausland geboren)	17	22	26
Testteilnahme der Schüler:innen			
Anzahl (Mathematik/Naturwissenschaften)	5194/5193	3437/3437	4442/4442
Teilnahmequote (%)	96	97	95
Beteiligung am Schülerfragebogen (%)	87	87	93
Beteiligung am Elternfragebogen (%)	77	67	60

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

Alle in Tabelle 2.3 dargestellten Abweichungen vom internationalen Idealwert (Teilnahmequote = 100 %) liegen im Bereich der akzeptablen Abweichungen. Die Ausschöpfungsquoten für alle teilnehmenden Staaten und Regionen seit 2007 sowie weitere Informationen sind im Anhang aufgeführt. Dass die Anzahl der Klassen die der Schulen übersteigt, obwohl ursprünglich nur eine Klasse je Schule vorgesehen war, liegt daran, dass bei besonders kleinen Klassen weitere Parallelklassen zusätzlich getestet wurden, um den notwendigen Stichprobenumfang zu erreichen. Zusätzliche Informationen und Hinweise zu den Ergebnisdarstellungen in diesem Band finden sich in Abschnitt 2.7.

Ein weiterer zu berücksichtigender Aspekt ist, dass Kinder, die auf Basis bestimmter Kriterien von der TIMSS-Testung befreit sind, nicht in die Ausschöpfungsquote einfließen. Gründe für eine Testbefreiung sind (1) bestimmte schulorganisatorische Aspekte, die eine Teilnahme verhindern (i. d. R. auf Schulebene, z. B. wegen zu geringer Größe), (2) Beeinträchtigungen oder Behinderungen der Kinder, die dazu führen, dass diese der Testung nicht gewachsen wären und (3) nicht hinreichende Sprachkenntnisse in der Testsprache (weniger als ein Jahr in der Testsprache unterrichtet). Die Testbefreiungsquoten unterscheiden sich in den teilnehmenden Staaten aufgrund verschiedener nationaler Herausforderungen und unterschiedlicher Traditionen integrativer und inklusiver Beschulung.

Die Möglichkeit der Testbefreiung führt immer wieder zu der Frage, ob sich in Deutschland das Gesamtergebnis verändern würde, wenn sich eine höhere Anzahl von Schüler:innen nicht an der Untersuchung beteiligen würde. Bei dieser Frage schwingt das Misstrauen mit, dass die beobachteten Stichprobenausfälle nicht aus den zuvor genannten Gründen erfolgen, sondern um die leistungsschwächsten Schüler:innen auszuschließen und so ein besseres Gesamtergebnis zu erzielen. Auch wenn systematische Stichprobenmanipulationen durch die umfangreichen Kontrollen der internationalen Studienleitung bei der Untersuchungsdurchführung praktisch ausgeschlossen werden können, wurde der hypothetische Effekt veränderter Testbefreiungsquoten auf den Rangplatz Deutschlands untersucht. Dazu werden in Tabelle 2.4 für Mathematik und in Tabelle 2.5 für Naturwissenschaften die Mittelwerte für Deutschland dargestellt, die sich ergeben würden, wenn eine höhere Anzahl leistungsschwächerer Schüler:innen ausgeschlossen werden würde. In der ersten Spalte sind jeweils die aktuellen Ergebnisse von TIMSS 2023 für Mathematik und Naturwissenschaften mit den tatsächlichen Ausschlussquoten für einen direkten Vergleich mit den verschiedenen hypothetischen Fällen für Deutschland dargestellt. Wie bereits in den vergangenen Zyklen werden dafür in den nachfolgenden Spalten als Referenz die Testbefreiungsquoten von Singapur (13.1 %), Schweden (6.2 %) und USA (5.5 %) herangezogen. Darüber hinaus wird zusätzlich die Ausschlussquote der Türkei als Referenz aufgenommen, da diese mit 21.9 Prozent die höchste Ausschlussquote aller Teilnehmer in TIMSS 2023 aufweist. Bei Ausschlussquoten von über 10 Prozent muss von einer in Frage zu stellenden Repräsentativität sowie einer deutlich eingeschränkten internationalen Vergleichbarkeit der Daten ausgegangen werden, weshalb inhaltliche Interpretationen nur unter Vorbehalt erfolgen sollten. Als Vergleichsstaaten wurden in den Tabellen jeweils die Bildungssysteme mit aufgenommen, die signifikant höhere oder nicht signifikant von Deutschland abweichende Kompetenzwerte in den Domänen Mathematik bzw. Naturwissenschaften aufweisen. Die Benchmark-Teilnehmer werden dabei nicht berücksichtigt.

Dabei zeigen sich in 2023 unter Berücksichtigung der hypothetischen Ausschlussquoten von Schweden und USA zwar minimale Verschiebungen und dadurch auch kleinere Veränderungen in der Rangreihe der berichteten Staaten. Insbesondere der signifikante Unterschied zu Staaten wie Singapur, Taiwan und der Republik Korea (Südkorea), deren Schüler:innen jeweils die höchsten Kompetenzen aufweisen, bleibt jedoch bestehen. Veränderungen ergeben sich lediglich bei den mittleren Rangplätzen, diese sind allerdings nicht signifikant. Ähnliche Befunde zeigten sich in TIMSS 2015 und TIMSS 2019. Unter Annahme der Ausschlussquoten von Singapur und der Türkei ergeben sich für Deutschland signifikant höhere hypothetische Leistungswerte und damit bedeutsame Änderungen in den Rangreihen für Mathematik und Naturwissenschaften. Nichtsdestotrotz bleibt in beiden Domänen ein großer Abstand zu den Spitzenreitern bestehen, die zumeist niedrigere Ausschlussquoten aufweisen als Deutschland (4.4 %).

Tabelle 2.4: Veränderung des Mittelwertes von Viertklässler:innen in Deutschland beim Vergleich der Leistung auf der Gesamtskala Mathematik in Abhängigkeit von der Ausschlussquote

Teilnehmer	Ausschlussquote in %	*M*	(*SE*)	Signifikanz zu Deutschland bei einer Ausschlussquote in Prozent von 4.4	5.5	6.2	13.1[B]	21.9[B]
2 *Singapur*	13.1	615	(2.3)	+	+	+	+	+
Taiwan	2.6	607	(1.4)	+	+	+	+	+
Republik Korea (Südkorea)	3.4	594	(2.6)	+	+	+	+	+
3 *Hongkong*	1.6	594	(4.0)	+	+	+	+	+
Japan	2.5	591	(2.3)	+	+	+	+	+
Macau	3.4	582	(0.9)	+	+	+	+	+
2 *Litauen*	8.6	561	(2.4)	+	+	+	+	+
1 2 3 *Türkei (5. Jgst.)*	21.9	553	(4.1)	+	+	+	+	
1 2 3 *England*	5.5	552	(2.7)	+	+	+	+	
Deutschland[A]	21.9	546	(1.7)	+	+	+	+	
2 *Polen*	6.5	546	(2.0)	+	+	+	+	
Irland	3.7	546	(2.9)	+	+	+	+	
2 3 *Rumänien*	6.5	542	(4.8)	+	+	+		
3 *Niederlande*	4.4	537	(2.0)	+	+	+		-
Deutschland[A]	13.1	536	(1.5)	+	+	+		-
Lettland	4.7	534	(2.6)	+	+	+		-
1 2 *Norwegen (5. Jgst.)*	6.1	531	(2.0)	+			-	-
2 *Tschechien*	7.3	530	(2.0)	+			-	-
2 *Schweden*	6.2	530	(2.6)				-	-
Bulgarien	3.7	530	(3.6)					-
Finnland	2.4	529	(2.3)				-	-
Deutschland[A]	6.2	527	(1.5)				-	-
Deutschland[A]	5.5	526	(1.6)				-	-
Australien	3.7	525	(2.0)				-	-
Deutschland	4.4	524	(1.9)				-	-
2 3 *Dänemark*	5.2	524	(2.1)				-	-
2 *Serbien*	10.2	523	(3.3)				-	-
2 3 *Belgien (Fläm. Gem.)*	5.7	521	(2.4)			-	-	-
2 *Ungarn*	5.1	520	(3.5)				-	-
Portugal	4.3	517	(2.8)		-	-	-	-

+ = Statistisch signifikant positive Abweichungen ($p \leq .05$) zu dem für Deutschland ermittelten Leistungswert unter Berücksichtigung der in der Spaltenüberschrift angegebenen Ausschlussquote
– = Statistisch signifikant negative Abweichungen ($p \leq .05$) zu dem für Deutschland ermittelten Leistungswert unter Berücksichtigung der in der Spaltenüberschrift angegebenen Ausschlussquote
Kursiv gesetzt sind Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Berechnung für Deutschland nach einem höheren Ausschluss von Schüler:innen mit geringen Kompetenzwerten
B = Informationen zu den Ursachen der hohen Ausschlussquoten für Singapur und die Türkei lagen zum Zeitpunkt der Veröffentlichung des Berichtsbandes noch nicht vor.

© TIMSS 2023

Tabelle 2.5: Veränderung des Mittelwertes von Viertklässler:innen in Deutschland beim Vergleich der Leistung auf der Gesamtskala Naturwissenschaften in Abhängigkeit von der Ausschlussquote

	Ausschlussquote			Signifikanz zu Deutschland bei einer Ausschlussquote in Prozent von				
Teilnehmer	in %	*M*	(*SE*)	4.4	5.5	6.2	13.1[B]	21.9[B]
[2] *Singapur*	13.1	607	(2.3)	+	+	+	+	+
Republik Korea (Südkorea)	3.4	583	(2.5)	+	+	+	+	+
Taiwan	2.6	573	(1.4)	+	+	+	+	+
[1 2 3] *Türkei (5. Jgst.)*	21.9	570	(3.4)	+	+	+	+	+
[1 2 3] *England*	5.5	556	(2.6)	+	+	+	+	+
Japan	2.5	555	(2.4)	+	+	+	+	+
[2] *Polen*	6.5	550	(2.2)	+	+	+	+	+
Australien	3.7	550	(1.8)	+	+	+	+	+
[3] *Hongkong*	1.6	545	(3.8)	+	+	+	+	
Finnland	2.4	542	(2.8)	+	+	+	+	
Deutschland[A]	21.9	541	(2.8)	+	+	+	+	
[2] *Litauen*	8.6	537	(2.4)	+	+	+	+	
Macau	3.4	536	(1.3)	+	+	+	+	
[2] *Schweden*	6.2	533	(3.1)	+	+	+		-
[2 3] *USA*	5.5	532	(2.4)	+	+	+		-
Irland	3.7	532	(3.2)	+	+	+		-
[1 2] *Norwegen (5. Jgst.)*	6.1	530	(2.5)	+	+	+		-
Deutschland[A]	13.1	530	(2.6)	+	+	+		-
Bulgarien	3.7	530	(4.8)	+	+			-
[2 3] *Rumänien*	6.5	526	(4.8)	+				-
[2] *Tschechien*	7.3	526	(2.2)	+	+	+		-
[2] *Slowenien*	5.1	526	(2.3)	+	+			-
Lettland	4.7	526	(2.9)	+	+			-
[2] *Ungarn*	5.1	524	(3.1)	+				-
[2 3] *Dänemark*	5.2	522	(2.5)				-	-
[2 3] *Kanada*	14.3	521	(1.7)				-	-
Slowakei	2.5	521	(3.3)				-	-
Deutschland[A]	6.2	519	(2.3)				-	-
Deutschland[A]	5.5	518	(2.4)				-	-
[1 2 3] *Neuseeland*	8.5	517	(2.8)				-	-
[3] *Niederlande*	4.4	517	(2.9)				-	-
Deutschland	4.4	515	(2.6)				-	-
Portugal	4.3	511	(2.3)		-	-	-	-
[2] *Italien*	7.0	511	(2.5)		-	-	-	-
[2] *Serbien*	10.2	510	(3.2)		-	-	-	-

\+ = Statistisch signifikant positive Abweichungen ($p \leq .05$) zu dem für Deutschland ermittelten Leistungswert unter Berücksichtigung der in der Spaltenüberschrift angegebenen Ausschlussquote
– = Statistisch signifikant negative Abweichungen ($p \leq .05$) zu dem für Deutschland ermittelten Leistungswert unter Berücksichtigung der in der Spaltenüberschrift angegebenen Ausschlussquote
Kursiv gesetzt sind Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Berechnung für Deutschland nach einem höheren Ausschluss von Schüler:innen mit geringen Kompetenzwerten
B = Informationen zu den Ursachen der hohen Ausschlussquoten für Singapur und die Türkei lagen zum Zeitpunkt der Veröffentlichung des Berichtsbandes noch nicht vor.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

2.4.2 Besonderheiten in TIMSS 2023

Mit Trendstudien wie TIMSS können langfristige Entwicklungen in Bildungssystemen, aber auch kurzfristige Ereignisse nachvollzogen werden. Bei solchen langen Zeiträumen, wie sie inzwischen mit TIMSS abgedeckt werden, kann es vorkommen, dass bestimmte gesellschaftliche, politische oder humanitäre Ereignisse Auswirkungen auf die Durchführung von Studien, auf die untersuchten Stichproben oder die Einordnung und Interpretation der Ergebnisse in einzelnen oder sogar allen teilnehmenden Staaten haben. In TIMSS 2023 betrifft dies vornehmlich zwei Ereignisse, die seit der Durchführung der vorherigen Erhebung im Jahr 2019 als besonders bedeutsam eingestuft werden können und zum Teil weltweite Auswirkungen hatten: zum einen die COVID-19-Pandemie, zum anderen

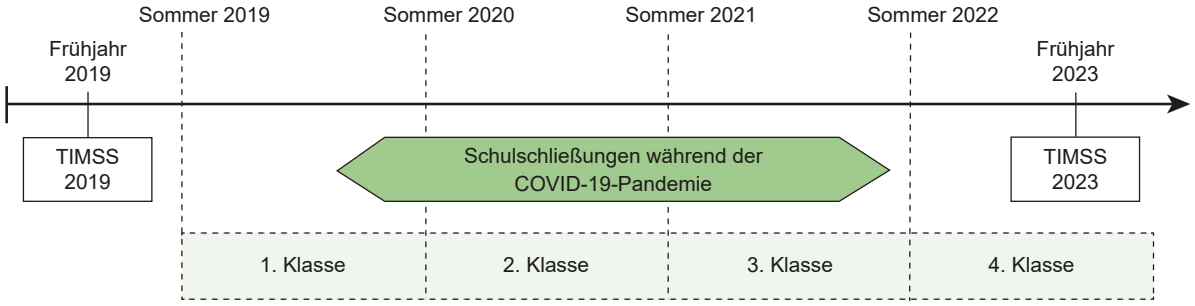
der Ukraine-Krieg. Die Bedeutung beider Ereignisse für TIMSS wird im Folgenden dargestellt, wobei die Betrachtungen vorrangig für Deutschland erfolgen.

Im März 2020 kam es in Folge der COVID-19-Pandemie in Deutschland zum ersten Lockdown, einer Abriegelung bzw. Einschränkung des öffentlichen Lebens, die auch die Schließung von Schulen umfasste. Die darauffolgenden Jahre der Pandemie 2021 und 2022 waren überwiegend geprägt von Lockdown-Phasen und zum Teil strengen Regulierungen zur Eindämmung der Pandemie, bis im Jahr 2023 COVID-19 zunehmend als endemisch eingestuft wurde (RKI, 2024). Im Frühling 2023 erklärte die *World Health Organization* (WHO) den Gesundheitsnotstand aufgrund der Pandemie für beendet (WHO, 2024). Während dieser Jahre waren auch die Schulen in Deutschland immer wieder teilweise oder vollständig geschlossen und von Fern- und Wechselunterricht geprägt, sodass die Beschulung von Kindern und Jugendlichen zu großen Teilen zu Hause erfolgte (Autor:innengruppe Bildungsberichterstattung, 2022). Die jeweiligen Regelungen unterschieden sich dabei zwischen den Ländern der Bundesrepublik Deutschland. Länderspezifische Analysen sind in TIMSS jedoch nicht möglich, sodass die unterschiedlichen Regelungen in den Analysen keine Berücksichtigung finden können.

Bisherige Untersuchungen konnten zeigen, dass die Lernzeit der Schüler:innen während der pandemiebedingten Beschulung zu Hause deutlich geringer ausfällt als bei regulärem Schulbesuch (Huber & Helm, 2020; Wacker et al., 2020), die schulischen Leistungen zum Teil schlechter ausfallen (Gasteiger et al., 2023) und häufiger psychosoziale Auffälligkeiten wie Probleme im Sozialverhalten oder emotionalen Bereich auftreten (Schneider et al., 2023). Neben den Schüler:innen waren auch Lehrkräfte und Erziehungsberechtigte von der veränderten schulischen Situation betroffen. Die Umstellung des gesamten Unterrichts, die häusliche Betreuung und Unterstützung der eigenen Kinder sowie erschwerte Kommunikation mit Schulen und Lehrkräften stellten besondere Herausforderungen für alle Beteiligten dar (Gasteiger et al., 2023; Schneider et al., 2021).

Die Schüler:innen, die 2023 im Rahmen von TIMSS getestet wurden, wurden im Sommer 2019 eingeschult und haben somit große Teile ihrer Schulzeit unter den Bedingungen der Pandemie erlebt. Zum Zeitpunkt der TIMSS-Erhebung im Frühjahr 2023 besuchten die Schüler:innen jedoch bereits seit einiger Zeit wieder regulär die Schulen in Deutschland, sodass die Testung von TIMSS 2023 wie in TIMSS 2019 im normalen Schulbetrieb stattfinden konnte und nicht unmittelbar von der COVID-19-Pandemie beeinflusst wurde. Dass die Schüler:innen einen großen Teil ihrer bisherigen Schulzeit unter dem Einfluss der COVID-19-Pandemie erlebten, muss bei der Interpretation der Ergebnisse in TIMSS 2023 berücksichtigt werden. Mit diesen Eigenschaften bieten die Zyklen von TIMSS 2019 und 2023 die Besonderheit, dass ein Vergleich der Situation unmittelbar vor und nach der COVID-19-Pandemie in Deutschland möglich ist. Der Zeitpunkt der Befragung der TIMSS-2023-Stichprobe bezogen auf die COVID-19-Pandemie kann Abbildung 2.3 entnommen werden.

Abbildung 2.3: Zeitliche Einordnung von TIMSS 2023 vor dem Hintergrund der COVID-19-Pandemie



IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Im Februar 2022 gipfelten bereits länger andauernde Konflikte in einen Angriffskrieg Russlands gegen die Ukraine. Dieser führte zu sprunghaft ansteigenden Zahlen an Geflüchteten in Deutschland und anderen Staaten (Statista, 2024a). Zum Zeitpunkt der Haupterhebung in TIMSS 2023 befanden sich in Deutschland ca. 205 000 ukrainische Schüler:innen an allgemein- und berufsbildenden Schulen und damit ca. 185 000 mehr als zu Beginn des Krieges (Statista, 2024b). Deshalb ist anzunehmen, dass eine große Anzahl an Viertklässler:innen, die keine oder nur marginale Deutschkenntnisse aufweisen, zum Zeitpunkt der Haupttestung Schulen in Deutschland besuchte. Aufgrund des in TIMSS definierten Testbefreiungskriteriums, dass Schüler:innen für eine Teilnahme mindestens ein Jahr in der Testsprache Unterricht erhalten haben sollten, fließt ein Großteil dieser ukrainischen Schüler:innen jedoch nicht mit in die Datenbasis von TIMSS 2023 ein (Anteil der in der Ukraine geborenen Schüler:innen in TIMSS 2023 unter 1 %). Es ist somit nicht davon auszugehen, dass die für Deutschland erhobenen Kompetenzstände durch diesen Anstieg von Schüler:innen aus der Ukraine beeinflusst sind.

2.4.3 Entwicklung der Testinstrumente

Die TIMSS-Untersuchung hat zum Ziel, die Kompetenzen von Schüler:innen in den Domänen Mathematik und Naturwissenschaften zu erfassen. Naturwissenschaften (in TIMSS mit den Inhaltsbereichen *Biologie*, *Physik/Chemie* und *Geografie*) werden in den Grundschulen in Deutschland im Rahmen des Sachunterrichts[7] vermittelt. Dazu werden Kompetenztests eingesetzt, die jeweils eine Auswahl von Aufgaben aus den Bereichen Mathematik und Naturwissenschaften aus einem größeren Aufgabenpool enthalten. Erstmals wurde dabei in TIMSS 2023 ein sogenanntes gruppenadaptives Testdesign eingesetzt, indem zwei unterschiedlich schwere Testhefte konzipiert und diese je nach vorherigen Leistungen der Staaten in unterschiedlicher Wahrscheinlichkeit eingesetzt wurden. So bekommen beispielsweise Staaten, die in vorherigen Zyklen oder dem Feldtest als leistungsstark eingestuft wurden, für eine höhere Messgenauigkeit anteilig mehr schwere Testhefte als Staaten mit durchschnittlichen oder niedrigeren Leistungen (Yin & Foy, 2021). Deutschland wurde als Staat mit durchschnittlicher

7 In Bayern und Thüringen wird das Fach als Heimat- und Sachunterricht bzw. Heimat- und Sachkunde bezeichnet. Der Einfachheit halber wird in diesem Band von dem Fach Sachunterricht gesprochen.

Leistungsstärke klassifiziert, weshalb beide Testheftversionen zu gleichen Anteilen eingesetzt wurden.

In Tabelle 2.6 ist die Gesamtzahl an Aufgaben für die Jahre 2007, 2019 und 2023 aufgeführt. Ein Großteil der Aufgaben, die in einem Zyklus eingesetzt wurden, wird nach der Testung nicht veröffentlicht, um in nachfolgenden Erhebungen wieder eingesetzt werden zu können. Die wiederholte Nutzung von Aufgaben in verschiedenen TIMSS-Zyklen ermöglicht den direkten Vergleich der Untersuchungen und bildet somit die notwendige Voraussetzung für Trendanalysen. Ein Teil der Aufgaben wird jedoch zur Illustration der Kompetenztests veröffentlicht, weshalb in jedem Zyklus neue Aufgaben entwickelt werden, um die Gesamtzahl über die Zeit hinweg auf einem ähnlichen Niveau zu halten.

Tabelle 2.6: Anzahl von Testaufgaben für Mathematik und Naturwissenschaften in TIMSS 2007, 2019 und 2023

Domänen	2007	2019	2023
	n	*n*	*n*
Mathematik	179	171	183
Naturwissenschaften	174	169	173
Gesamt	353	340	356

Die Werte geben die Anzahl der Items an, die für die internationale Skalierung der Leistungsergebnisse berücksichtigt wurden. In einzelnen Staaten und Zyklen standen zum Teil geringfügig mehr Items für nationale Auswertungen zur Verfügung.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Da die Bearbeitung aller in einem TIMSS-Zyklus eingesetzten Aufgaben durch alle Schüler:innen zu zeitaufwendig wäre und darüber hinaus auch über einer zumutbaren Belastung der Grundschüler:innen liegen würde, wurden die Aufgaben in kleinere Aufgabenblöcke aufgeteilt, von denen die Kinder jeweils nur eine Auswahl bearbeiten. Durch eine systematische Zusammenstellung dieser Aufgabenblöcke kann sichergestellt werden, dass der Mathematik- und Naturwissenschaftstest in TIMSS aufgrund der hohen Zahl an Aufgaben insgesamt hinreichend valide ist, die einzelnen Schüler:innen jedoch nicht überlastet werden und schließlich durch die Testung von ausreichend vielen Schüler:innen auch Aussagen über das Bildungssystem getroffen werden können. Die Einschränkung, die man bei dieser Art von Systemmonitoringstudien in Kauf nimmt, ist, dass keine differenzierten Befunde für die einzelnen Kinder – im Sinne einer Individualdiagnostik – vorliegen. Die Summe aller Informationen ist jedoch gut geeignet, um Aussagen über die nationale Grundgesamtheit der Untersuchung treffen zu können (Populationsschätzer). Damit sind die in TIMSS gesammelten Informationen auch dafür geeignet, Analysen zu Lehr- und Lernbedingungen vorzunehmen und Aussagen auf Ebene der Schule und des Unterrichts zu formulieren.

Obwohl die Schüler:innen jeweils unterschiedliche Teile des TIMSS-Gesamttests bearbeiten, ist es durch die systematische Variation der Testblöcke möglich, für alle Schüler:innen Kompetenzwerte auf derselben Kompetenzskala zu ermit-

teln.[8] Bei der Bestimmung der Leistungsdaten werden auch Informationen aus den Hintergrundfragebögen hinzugezogen, um die bestmöglichen Schätzer für die Kompetenzwerte in der Grundgesamtheit (Population) zu erhalten. Insgesamt werden je teilnehmendem Kind und Domäne fünf Kompetenzwerte (*plausible values*) bestimmt, die im Rahmen simultaner Berechnungen zusammengeführt werden und die bestmöglichen Schätzer für Aussagen über das jeweilige Bildungssystem erlauben. Dadurch, dass die vergangenen Zyklen durch eine gemeinsame Bestimmung der Aufgabenschwierigkeiten auf derselben Metrik liegen und gemeinsam skaliert werden, sind Trendanalysen für die teilnehmenden Bildungssysteme möglich.

2.4.4 Entwicklung der Fragebögen

Anspruch von TIMSS ist es, neben Informationen zu den Kompetenzen von Grundschüler:innen auch Informationen zu den Lehr- und Lernbedingungen in den verschiedenen teilnehmenden Staaten bereitzustellen. Diese werden über Hintergrundfragebögen für die Schüler:innen, deren Erziehungsberechtigten, ihre Lehrkräfte und die Schulleitungen erhoben. Um sicherzustellen, dass die erfragten Informationen international vergleichbar sind, gibt es für alle teilnehmenden Staaten einen gemeinsamen Teil, welcher von der internationalen Studienleitung vorgegeben wird. Der internationale Teil in den Hintergrundfragebögen kann darüber hinaus durch sogenannte nationale Erweiterungen von den teilnehmenden Staaten individuell ergänzt werden. Diese Möglichkeit hat Deutschland seit der ersten Durchführung von TIMSS 2007 genutzt, um auf aktuelle Fragestellungen im Kontext der Schul- und Unterrichtsforschung zu reagieren sowie Merkmale und Wirkungen aktueller Herausforderungen, die sich beispielsweise durch gesellschaftliche Veränderungen ergeben haben, zu erfassen. So wurde eine Datenbasis geschaffen, die es ermöglicht, spezifische Fragen zum deutschen Schulsystem sowohl im internationalen Vergleich als auch im nationalen Kontext zu beantworten (Hornberg et al., 2007; Lankes et al., 2003; Schwippert, Kasper et al., 2020; Schwippert et al., 2015; Wendt et al., 2016; Wendt et al., 2012). In den Fragebögen von TIMSS 2023 wurde deshalb beispielsweise nicht mehr nach *Mutter* und *Vater*, sondern nach *Elternteil A* und *Elternteil B* gefragt, um vielfältige Familienkonstellationen einzuschließen. Weiterhin wurde die Möglichkeit genutzt, zusätzlich zu den binären Geschlechterkategorien eine dritte Kategorie anzubieten (im Schülerfragebogen: *anderes*; im Lehrkräfte- und Schulleitungsfragebogen: *divers*).

Um einen möglichst umfassenden Eindruck der Rahmenbedingungen des Lehrens und Lernens bei Grundschulkindern zu erhalten, werden unterschiedliche Informationsebenen durch die TIMSS-Befragungen adressiert. Neben den Kompetenztests und den Hintergrundfragebögen gibt es deshalb zusätzlich einen Fragebogen zu den Lehrplänen, der von den nationalen Studienleitungen ggf. in Zusammenarbeit mit Expert:innen bearbeitet wird. In Tabelle 2.7 sind die Adressat:innen der Erhebungsinstrumente den jeweils ermittelten Informationen zugeordnet. Die Informationen in der Tabelle geben dabei eine Übersicht über die

8 Dabei werden mit einem rotierten Multi-Matrix-Design anhand von Modellen der *Item-Response-Theory* basierend auf den Tests und Hintergrunddaten die Schülerkompetenzen für die beiden Domänen bestimmt. Bei Multiple-Choice-Fragen wird das *3-PL-Model* angewendet, bei offenen Fragen mit einer richtigen Antwortmöglichkeit das *2-PL-Model* und bei mehr als zwei möglichen Antwortabstufungen das *Partial-Credit-Model*.

Inhalte in den Instrumenten. Dadurch lässt sich in TIMSS das bereits beschriebene Curriculum-Modell abbilden, wobei sich dazu eine mehrstufige Betrachtung bewährt hat (Blömeke, 2014; Bos & Postlethwaite, 2005): Zunächst wird in den Blick genommen, welche Ziele im Bildungssystem durch ministerielle Vorgaben erreicht werden sollen (intendiertes Curriculum). Das implementierte Curriculum als zweiter Schritt beschreibt, was durch eine individualisierte Vorbereitung durch Lehrkräfte sowie unter der Nutzung von Unterrichtsmaterialien (insbesondere Lehrbücher, Arbeitsblätter etc.) im Unterricht realisiert wird. Drittens werden vor dem Hintergrund schulischer Rahmenbedingungen sowie individueller Voraussetzungen der Schüler:innen die vermittelten Inhalte individuell verarbeitet und aufgenommen und als realisiertes Curriculum sichtbar. So nachvollziehbar dieser Prozess vom intendierten bis hin zum realisierten Curriculum ineinandergreift, so schwierig ist es auch, Veränderungen am Anfang dieses Prozesses (im Sinne der Veränderung von Zielen durch die Ministerien) bis zu den konkreten Kompetenzen der Schüler:innen nachzuvollziehen.

Tabelle 2.7: Erhebungsinstrumente und Charakterisierung der erfragten Informationen in TIMSS 2023

Instrument	Adressat:innen	Information	Informationsebene
Curriculum-Fragebogen	Nationale Studienleitung	Inhalte und Entwicklung der ministeriellen Vorgaben für den Unterricht	Intendiertes Curriculum
TIMSS-Test	Für TIMSS 2023 ausgewählte Schüler:innen	Kompetenzen in Mathematik und in den Naturwissenschaften	Realisiertes Curriculum
Schülerfragebogen	Für TIMSS 2023 ausgewählte Schüler:innen	Selbst- und fachbezogene Einstellungen, bildungsrelevante Ressourcen, individuelles inner- und außerschulisches Lernverhalten	
Elternfragebogen	Erziehungsberechtigte der ausgewählten Schüler:innen	Häusliche Lernumgebung und Ressourcen, Erziehungs- und Unterstützungsverhalten, familiäre Merkmale	
Lehrkräftefragebogen	Lehrkräfte der ausgewählten Schüler:innen	Unterrichtsqualität und -gestaltung, Nutzung digitaler Medien, Qualifizierung und Fortbildung	Implementiertes Curriculum
Schulleitungsfragebogen	Schulleitungen der ausgewählten Schulen	Strukturelle, organisatorische und personelle Rahmenbedingungen der Schule	Implementiertes Curriculum

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Bei der Konzeption der Hintergrundfragebögen werden zum Teil komplexe Abwägungsprozesse durchlaufen, um eine Balance zwischen der Sicherstellung des bestehenden Trends und der Berücksichtigung aktueller Herausforderungen zu finden. Die Reaktion auf aktuelle Herausforderungen bedeutet i. d. R. die Berücksichtigung neuer Fragen bzw. Frageblöcke in den Hintergrundfragebögen (z. B. Fragen zur COVID-19-Pandemie), für deren Beantwortung aber nur eine begrenzte Zeit zur Verfügung steht. Um die Befragten und insbesondere die Kinder zeitlich nicht zu überlasten, müssen für eine Ergänzung neuer Fragen ältere Fragen gestrichen werden, auch wenn dies das Ende eines mit diesen Fragen dokumentierten Trends bedeutet. Die hierzu notwendigen Entscheidungen wurden

im deutschen TIMSS-2023-Konsortium gemeinsam getroffen. Darüber hinaus hat das deutsche TIMSS-2023-Konsortium entschieden, nach den Kompetenztests, die bereits in TIMSS 2019 auf eine digitale Erhebung umgestellt wurden, auch die Befragung der Schüler:innen, der Lehrkräfte und der Schulleitungen in diesem Zyklus erstmals digital durchzuführen. Der Elternfragebogen wurde nach wie vor papierbasiert erhoben, um möglichst hohe Ausschöpfungsquoten zu erzielen. Unterstützend wurden zusätzlich im Vorfeld der Erhebung Informationsflyer in sieben verschiedenen Sprachen von den Schulen an die Erziehungsberechtigten verteilt.

2.5 Vorbereitung und Durchführung von TIMSS 2023

Als Reaktion auf aktuelle gesellschaftliche Veränderungen und deren Konsequenzen für die Bildungssysteme werden in jedem TIMSS-Zyklus die Instrumente und Prozeduren angepasst und weiterentwickelt. Um die Trenduntersuchung ohne Qualitätsverlust fortschreiben zu können, wird empirisch überprüft, ob diese Weiterentwicklungen inhaltlich geeignet und praktisch anwendbar sind. Den jeweiligen Haupttests sind daher in jedem TIMSS-Zyklus Feldtests vorangestellt, um neue Instrumente und Abläufe zu entwickeln und zu erproben.

2.5.1 Feldtest im Frühjahr 2022

Zu Beginn eines neuen TIMSS-Zyklus und damit mehr als zwei Jahre vor der eigentlichen Haupterhebung wird mit der Entwicklung und Anpassung der Test- und Befragungsinstrumente begonnen. Bei der (Weiter-)Entwicklung der Instrumente werden vier zentrale Zielsetzungen verfolgt: (1) Entwicklung neuer inhaltlicher Testaufgaben, (2) Entwicklung neuer innovativer Testformate, (3) Entwicklung neuer Fragen zu Rahmenbedingungen des Lehrens und Lernens sowie (4) Sicherung des TIMSS-Trends. Damit sichergestellt wird, dass neu entwickelte Testaufgaben und Befragungsinstrumente anschlussfähig an den bisher dokumentierten TIMSS-Trend sind, werden diese im Rahmen des Feldtests einer empirischen Überprüfung unterzogen. Neben der psychometrisch-empirischen Analyse der Instrumente werden im Feldtest auch die Organisation und Durchführung der Untersuchung systematisch erprobt.

Am TIMSS-Feldtest nahmen die meisten der an TIMSS 2023 beteiligten Staaten und Regionen teil, damit die weiter- und neuentwickelten Testaufgaben und Befragungsinstrumente in den nationalen Sprachvarianten geprüft werden konnten. In Deutschland wurde der Feldtest zudem zur Kürzung der nationalen Teile in den Hintergrundfragebögen und zur Erprobung der neuen Fragen im Rahmen des TIMSS-Schwerpunktes zu digitalen Medien in Grundschulen genutzt. Der TIMSS-Feldtest fand im Frühjahr 2022 in fünf Ländern der Bundesrepublik Deutschland (Berlin, Bayern, Niedersachsen, Rheinland-Pfalz und Sachsen-Anhalt) unter Beteiligung von ca. 1 100 Schüler:innen sowie den dazugehörigen Erziehungsberechtigten, Lehrkräften und Schulleitungen statt.

2.5.2 Haupttest im Frühjahr 2023

Die Haupterhebung von TIMSS 2023 wurde im Frühjahr 2023 durchgeführt. Das gewählte Testzeitfenster lag in allen Ländern der Bundesrepublik Deutschland vor den Sommerferien, sodass mit der TIMSS-Erhebung die Kompetenzen der Schüler:innen in Mathematik und den Naturwissenschaften am Ende der Grundschulzeit erhoben werden konnten. Die Erfassung der Kompetenzen der Schüler:innen sowie die Befragung der Schüler:innen, der Mathematik- und Sachunterrichtslehrkräfte und der Schulleitungen erfolgte in digitaler Form. Die Schüler:innen wurden in den Schulen mithilfe von Laptops getestet. Die Zugangsdaten für die digitalen Lehrkräfte- und Schulleitungsfragebögen wurden den Lehrkräften und Schulleitungen zusammen mit weiteren Informationen zur Befragung in TIMSS vor dem Testtag postalisch übermittelt. Die Befragung der Erziehungsberechtigten wurde wie in den vorherigen Zyklen in papierbasierter Form durchgeführt. Der Elternfragebogen wurde dazu am Testtag allen teilnehmenden Schüler:innen mit der Bitte ausgehändigt, diesen an ihre Erziehungsberechtigten weiterzugeben. In einem Anschreiben wurden die Erziehungsberechtigten darum gebeten, den Fragebogen zu bearbeiten und diesen ihrem Kind in einem verschlossenen Umschlag wieder mit in die Schule zu geben.

2.5.3 Übersetzung, Testdurchführung und Qualitätssicherung

Damit sichergestellt werden kann, dass die Instrumente und Prozesse in allen teilnehmenden Staaten und Regionen einheitlich sind, werden diese von der internationalen Studienleitung standardisiert vorgegeben. Dies betrifft in der Vorbereitung der Testung insbesondere die Übersetzung von Materialien, während der Erhebung die Testdurchführung und zusätzlich übergreifende Prozesse zur Sicherung der Qualität.

Da die internationale Studiensprache in TIMSS Englisch ist, werden alle Materialien durch das ISC in dieser Sprache bereitgestellt. Zu diesen Materialen gehören neben den Kompetenztests und Hintergrundfragebögen auch Manuale und verschiedene Handreichungen zur Durchführung der Studie. In den Staaten und Regionen, die mit den englischsprachigen Unterlagen nicht direkt arbeiten können, sind diese durch Übersetzer:innen in die jeweilige(n) Landessprache(n) zu übersetzen. Da der spätere Vergleich der Bildungssysteme insbesondere auf den getesteten Kompetenzen und erfragten Hintergrundinformationen basiert, werden diese Übersetzungen besonderen Qualitätskontrollen unterzogen. Aufwendige Rückübersetzungen und ein Vergleich mit den englischen Originalvarianten sorgen im Rahmen der internationalen Kontrolle der Übersetzungen (*Translation Verification*) und des Layouts (*Layout Verification*) für international vergleichbare Befragungsinstrumente. Mögliche Abweichungen werden mit den jeweiligen Staaten und Regionen kommuniziert und im gegenseitigen Austausch angepasst. Notwendige nationale Abweichungen von der internationalen Vorlage werden durch das ISC dokumentiert und bei der anschließenden Berichtslegung berücksichtigt.

In der Vorbereitung und Durchführung der Erhebungen sind in allen teilnehmenden Staaten standardisierte Abläufe in den Testsitzungen und zeitliche Vorgaben einzuhalten. Dazu werden externe Testleiter:innen umfassend geschult und zusätzlich detaillierte Manuale und Skripte während der Erhebung eingesetzt. Am Testtag bearbeiten die Schüler:innen die Kompetenztests in altersge-

Tabelle 2.8: Untersuchungsablauf der Testung im TIMSS-Haupttest 2023

Ablauf der Testsitzung	TIMSS 2023
Vor Beginn der Testsitzung: Vorbereitung der Testsitzung zusammen mit der Testleitung	
Beginn der Testsitzung: Begrüßung, Einweisung	ca. 30 Min.
Bearbeitung des Testhefts Teil I (Mathematik oder Naturwissenschaften)	**36 Min.**
Pause	10 Min.
Einweisung in den zweiten Testteil	ca. 5 Min.
Bearbeitung des Testhefts Teil II (Mathematik oder Naturwissenschaften)	**36 Min.**
Pause	20 Min.
Einweisung in die Bearbeitung des kognitiven Fähigkeitstests Teil I A+B	5 Min.
Bearbeitung des kognitiven Fähigkeitstests Teil I A (CFT 20-R) Matrizen	**4 Min.**
Bearbeitung des kognitiven Fähigkeitstests Teil I B (CFT 20-R) Reihenfortsetzen	**3 Min.**
Einweisung in die Bearbeitung des kognitiven Fähigkeitstests Teil II	3 Min.
Bearbeitung des kognitiven Fähigkeitstests Teil II (KFT, V3)	**7 Min.**
Einweisung in die Bearbeitung des Schülerfragebogens	ca. 5 Min.
Bearbeitung des Schülerfragebogens Teil I	**25 Min.**
Pause	10 Min.
Bearbeitung des Schülerfragebogens Teil II	**20 Min.**
Beenden der Testsitzung: Einsammeln der Testmaterialien	
Reine Bearbeitungszeit Gesamtzeit	131 Min. ca. 219 Min.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

rechten Abschnitten, unterbrochen durch regelmäßige und aktive Pausen (siehe Tabelle 2.8). Abgeschlossen wird die Testsitzung durch das Bearbeiten des digitalen Schülerfragebogens. An den für die TIMSS-Untersuchung ausgewählten Förderschulen konnten die Test- und Befragungsteile, in Absprache mit der internationalen Studienleitung, abweichend auf zwei Testtage verteilt werden, um die zeitliche Belastung der Schüler:innen an diesen Schulen zu verringern.

Um sicherzustellen, dass den internationalen Vorgaben für die Testung in den Schulen gefolgt wird, hospitieren durch die IEA Amsterdam geschulte unabhängige Prozessbeobachter:innen in ausgewählten Schulen. Auch der Ablauf dieser Qualitätssicherungsmaßnahme in den Schulen ist standardisiert und wird von den internationalen Prozessbeobachter:innen im Rahmen der Qualitätssicherung (*Quality Control Monitoring*) schrittweise durchgeführt und dokumentiert. Bei zu starken Abweichungen von den Vorgaben der internationalen Studienleitung kann es zu Ausschlüssen von Schüler:innen oder sogar ganzen Klassen für nachfolgende Auswertungen kommen.

Darüber hinaus durchläuft auch die Dateneingabe eine mehrstufige Qualitätskontrolle. Im Anschluss an die Testung mit den mehrfach geprüften Instrumenten werden die erhobenen Daten aus der Befragung der Erziehungsberechtigten in ein elektronisches Format überführt. Dazu werden die Daten in den nationalen Testzentren digitalisiert und gemeinsam mit den Informationen der digitalen Befragungen auf Eingabefehler kontrolliert. Anschließend übernimmt die internationale Datenverarbeitungsabteilung der IEA Hamburg die Zusammenführung und Kontrolle der Daten aller teilnehmenden Staaten und Regionen. Nach Kontrolle und Korrekturen von Eingaberoutinen werden darüber hinaus national

notwendige Anpassungen (z. B. Rekodierungen) so vorgenommen und dokumentiert, dass sie in eine gemeinsame internationale Datenbank einfließen können. Diese Datenbank kann anschließend genutzt werden, um die Kompetenzstände der Schüler:innen zu bestimmen und damit die Basis für die internationalen und nationalen Berichtslegungen zu schaffen.

2.6 Technische und empirische Voraussetzungen für die Auswertung

Um die Besonderheiten in der Anlage und Durchführung der Untersuchung angemessen zu berücksichtigen, müssen das Stichproben- und das Testdesign bei den Analysen beachtet werden. Dies erfolgt über die Nutzung von Stichprobengewichten, die Bestimmung von Messfehlern durch die Verwendung von *plausible values* (im Rahmen von Kompetenzanalysen) und die stichprobengerechte Bestimmung der Stichprobenfehler bei den Analysen (durch das sogenannte *Jackknifing*-Verfahren[9], Foy & LaRoche, 2016).

Zusätzlich führt das gewählte Stichprobendesign zu einer weiteren Besonderheit in Bezug auf die Daten der Lehrkräfte und Schulleitungen. Weil das Design die Ziehung einer repräsentativen Schülerstichprobe gewährleisten soll, können keine repräsentativen Aussagen über die Lehrkräfte bzw. Schulleitungen in Deutschland getroffen werden. Grund dafür ist, dass in TIMSS die Lehrkräfte befragt werden, die die gezogenen Schüler:innen im Mathematik- und Sachunterricht unterrichten, und somit die Lehrkräfte nicht als Zufallsauswahl aller Lehrkräfte einer Schule ausgewählt wurden. In dieser Logik können somit nur repräsentative Aussagen über Schüler:innen gemacht werden, nicht aber über die Lehrkräfte im Allgemeinen. So ist die Formulierung, dass z. B. ‚63 Prozent aller Schüler:innen von weiblichen Lehrkräften unterrichtet werden' zwar umständlich, aber korrekt – im Gegensatz zu der Aussage, dass ‚63 Prozent der Lehrkräfte weiblich sind'. Da im Gegensatz dazu jedoch die Ziehungswahrscheinlichkeiten der Schulen und somit der Schulleitungen bestimmbar sind, können auch Aussagen über Schulleitungen in Deutschland ohne Bezug auf die Schüler:innen gemacht werden. Wird bei Analysen einer dieser Aspekte nicht berücksichtigt, kann es zu Fehlinterpretationen sowohl bei der Beschreibung der Befunde im internationalen Vergleich als auch im Trend kommen.

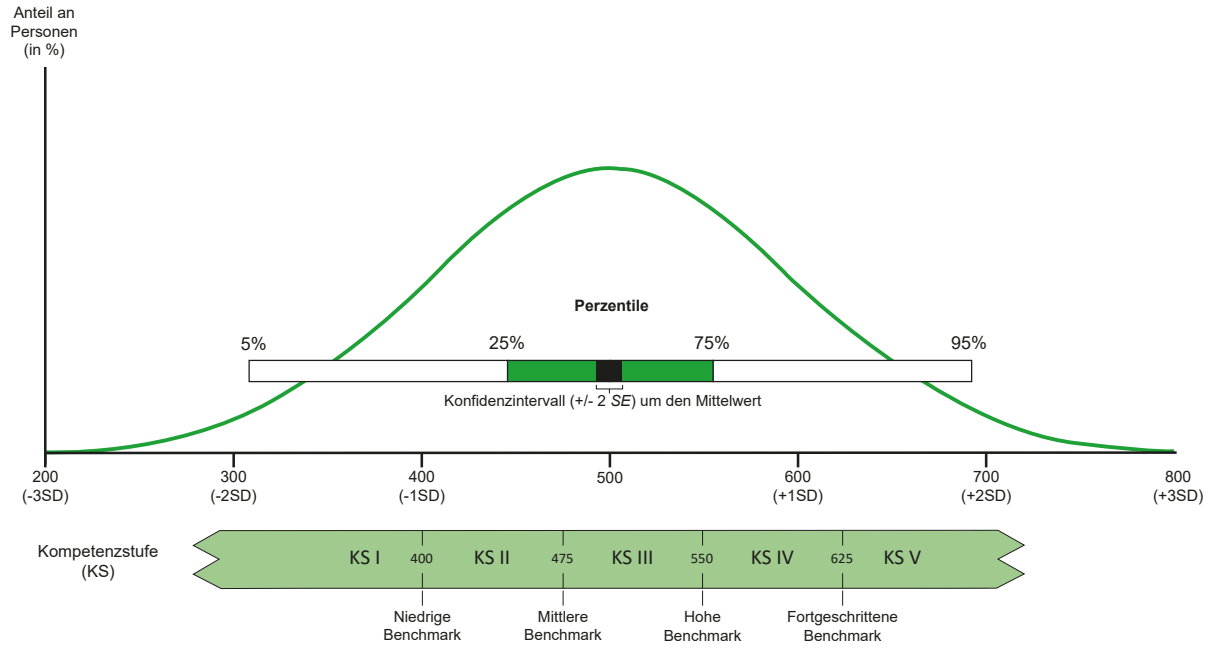
2.7 Hinweise für die Ergebnisdarstellung

Zur Darstellung der Ergebnisse werden im vorliegenden Band Tabellen und Abbildungen genutzt, um die verschiedenen komplexen Befundlagen möglichst anschaulich zu visualisieren. Dabei werden neben Histogrammen (Säulendiagramme) zur Darstellung von prozentualen Anteilen auch statistische Kennwerte (wie z. B. Mittelwerte) aufgeführt. Eine besondere Form der Darstellung in diesem Band stellt das Perzentilband dar (siehe Abbildung 2.4). Dieses berücksichtigt die stichprobenbasierten Unsicherheiten (Standardfehler und Konfidenzintervall), um beurteilen zu können, ob beobachtete Unterschiede bzw. Abweichungen zufällig oder signifikant sind. Gerade bei der vergleichenden

9 Abweichend von der internationalen Berichtslegung wird in diesem Band das gleiche *Jackknifing*-Verfahren wie in TIMSS 2019 verwendet, sodass minimale Abweichungen zu den international berichteten Standardfehlern auftreten. Weiterführende Informationen finden sich bei von Davier et al. (2024).

Gegenüberstellung der Kompetenzen der Grundschüler:innen wird diese Abbildungsvariante genutzt, da diese den Mittelwert, dessen statistische Streuung und die allgemeine Verteilung der Kompetenzen (über die mittleren 90 % einer Merkmalsverteilung) abbildet. Die jeweils verbleibenden 5 Prozent am unteren und oberen Ende der Verteilung werden nicht abgebildet, da diese mit besonders großen statistischen Unsicherheiten behaftet sind. Ein Perzentilband kann wie Höhenlinien einer von oben betrachteten statistischen Verteilung interpretiert werden. Der mittlere (schwarze) Abschnitt kennzeichnet den für die entsprechende Gruppe festgestellten Mittelwert inklusive der stichprobenbasierten Schwankung (Konfidenzintervall). Die dunkelgrün hinterlegten Balken links und rechts davon kennzeichnen die mittleren 50 Prozent der Beobachtungen bei dem untersuchten Merkmal. Die links und rechts anschließenden Balkenabschnitte (ohne farbige Füllung) kennzeichnen jeweils die 20 Prozent niedrigeren bzw. 20 Prozent höheren Werte. Werden entsprechende Perzentilbänder für verschiedene Gruppen (z. B. Staaten im Hinblick auf die Kompetenzen der Viertklässler:innen oder besondere Schülergruppen innerhalb von Deutschland) dargestellt, so gibt die parallele Darstellung der Konfidenzintervalle der Staaten einen ersten visuellen Eindruck, ob sich diese Gruppen signifikant voneinander unterscheiden (keine Überlappung der Konfidenzintervalle) oder nicht (Überlappung der Konfidenzintervalle). Neben den visuellen Hinweisen auf signifikante Unterschiede zeigen Symbole in den daneben dargestellten Tabellen explizit statistisch signifikante Unterschiede an.

Abbildung 2.4: Normalverteilung mit Perzentilband, Benchmarks und Kompetenzstufen



IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Die Perzentilbänder geben, insbesondere in Verbindung mit weiteren statistischen Kennwerten wie Mittelwerten und Standardabweichungen, empirisch ermittelte Werte wieder, jedoch sind diese Angaben als Zahlenwerte sehr abstrakt. Um die Skalenwerte der Kompetenztests inhaltlich intuitiver nachvollziehbar zu machen, wurden Bereiche der Kompetenzverteilung zu sogenannten Kompetenzstufen zusammengefasst, die durch international vorgegebene Grenzwerte (Benchmarks)

getrennt werden. In Abbildung 2.4 sind die Benchmarks und Kompetenzstufen unter der Verteilungskurve und dem korrespondierenden Perzentilband dargestellt.

Bezogen auf die einzelnen Stichproben der an TIMSS teilnehmenden Staaten und Regionen gibt es zusätzliche Hinweise für die Interpretation der in diesem Band dargestellten Ergebnisse. In diesem Zusammenhang werden Abweichungen von den durch die internationale Studienleitung definierten Stichprobenumfängen nachvollzogen und dokumentiert. Diese Dokumentation ist notwendig, damit bei vertiefenden Vergleichen von Staaten und Regionen nur solche einander gegenübergestellt werden, die tatsächlich aufgrund gleicher Teilnahmebedingungen auch vergleichbar sind. Abweichungen vom international vorgegebenen Stichprobenplan sind in diesem Band jeweils als Fußnoten in den Abbildungen vermerkt. Nachfolgend sind die Fußnoten aufgeführt, die in den Auswertungen im vorliegenden Band berichtet werden. Diese werden seit TIMSS 2011 in gleicher Weise dokumentiert.

1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten

Die Fußnote 4 wird notwendig, wenn für Schüler:innen keine *plausible values* vorliegen. Darüber hinaus wird sie bei Bedarf in Fußnote 4 (nicht skalierbare Leistungswerte in Mathematik) und Fußnote 5 (nicht skalierbare Leistungswerte in den Naturwissenschaften) differenziert. Zudem gibt es drei weitere, zyklusspezifische Fußnoten. Die Fußnote 6 verweist darauf, dass die Testung zeitlich verzögert durchgeführt wurde. Für die Kennzeichnung der Teilnahme von Staaten bzw. Benchmark-Teilnehmern an TIMSS *Numeracy* (in 2019: *Less Difficult Mathematics*) wurde in TIMSS 2015 die Fußnote 7 eingeführt. Sowohl Fußnote 6 als auch Fußnote 7 werden in TIMSS 2023 jedoch nicht benötigt. Besonderheiten in den Trendvergleichen werden durch die Fußnote 8 gekennzeichnet (siehe Anhang). Zusätzlich wird durch die Kursivschreibung von Staaten und Benchmark-Teilnehmern kenntlich gemacht, dass Besonderheiten in den Erhebungsbedingungen vorliegen, sodass eine Interpretation der Ergebnisse nur unter deren Berücksichtigung erfolgen sollte. Erläuternd ist dazu unter den Abbildungen der Hinweis ‚Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss' ergänzt (siehe Anhang).[10]

10 Die hier vorgestellten internationalen Ergebnisse basieren auf dem Stand der Daten vom 06.09.2024, die Ergebnisse der nationalen Analysen auf dem Stand der Daten vom 18.07.2024.

Literatur

Autor:innengruppe Bildungsberichterstattung. (2022). *Bildung in Deutschland 2022. Ein indikatorengestützter Bericht mit einer Analyse zum Bildungspotential.* https://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2022/pdf-dateien-2022/bildungsbericht-2022.pdf

Beese, C., Scholz, L. A., Steffensky, M., & Walter, D. (2024). Germany. In K. A. Reynolds, C. E. A. Aldrich, A. Bookbinder, A. Gallo, M. von Davier & A. Kennedy (Hrsg.), *TIMSS 2023 encyclopedia: Education policy and curriculum in mathematics and science*. TIMSS & PIRLS International Study Center.

Blömeke, S. (2014). Forschung zur Lehrerbildung im internationalen Vergleich. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 441–467). Waxmann.

Bos, W., & Postlethwaite, T. N. (2005). Möglichkeiten, Grenzen und Perspektiven internationaler Schulleistungsforschung. In R. Tippelt (Hrsg.), *Handbuch Bildungsforschung* (S. 241–261). Springer. https://doi.org/10.1007/978-3-322-99634-3_13

Bos, W., Wendt, H., Köller, O., & Selter, C. (Hrsg.). (2012). *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich.* Waxmann.

Foy, P., & LaRoche, S. (2016). Estimating standard errors in the TIMSS 2015 results. In M. O. Martin, I. V. S. Mullis & M. Hooper (Hrsg.), *Methods and procedures in TIMSS 2015* (S. 4.1–4.69). TIMSS & PIRLS International Study Center. http://timssandpirls.bc.edu/publications/timss/2015-methods.html

Gasteiger, H., Sachse, K. A., Schumann, K., Gerve, M., Schulz, A., & Engelbert-Kocher, M. (2023). COVID-19-related school closures and mathematical performance – findings from a study with grade 3 students in Germany. *Frontiers in Education, 8*, 1213857. https://doi.org/10.3389/feduc.2023.1213857

Hornberg, S., Bos, W., Buddeberg, I., Potthoff, B., & Stubbe, T. C. (2007). Anlage und Durchführung von IGLU 2006. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 21–45). Waxmann.

Huber, S. G., & Helm, C. (2020). COVID-19 and schooling: evaluation, assessment and accountability in times of crises – reacting quickly to explore key issues for policy, practice and research with the school barometer. *Educational Assessment, Evaluation and Accountability, 32,* 237–270. https://doi.org/10.1007/s11092-020-09322-y

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2016). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring.* https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2015/2015_06_11-Gesamtstrategie-Bildungsmonitoring.pdf

Lankes, E.-M., Bos, W., Mohr, I., Plaßmeier, N., Schwippert, K., Sibberns, H., & Voss, A. (2003). Anlage und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU) und ihrer Erweiterung um Mathematik und Naturwissenschaften (IGLU-E). In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 7–28). Waxmann.

Martin, M. O., Mullis, I. V. S., & Hooper, M. (Hrsg.). (2016). *Methods and procedures in TIMSS 2015*. http://timssandpirls.bc.edu/publications/timss/2015-methods.html

Reynolds, K. A., Aldrich, C. E. A., Bookbinder, A., Gallo, A., von Davier, M., & Kennedy, A. (Hrsg.). (2024). *TIMSS 2023 encyclopedia: Education policy and curriculum in mathematics and science*. TIMSS & PIRLS International Study Center.

RKI (Robert-Koch-Institut). (2024). *Antworten auf häufig gestellte Fragen zur COVID-19-Pandemie.* https://www.rki.de/SharedDocs/FAQ/NCOV2019/FAQ-Liste-COVID-19-Pandemie.html?nn=16765520

Särndal, C.-E., Swensson, B., & Wretman, J. (1992). *Model assisted survey sampling*. Springer. https://doi.org/10.1007/978-1-4612-4378-6

Schneider, R., Sachse, K. A., Schipolowsky, S., & Enke, F. (2021). Teaching in times of COVID-19: The evaluation of distance teaching in elementary and secondary schools in Germany. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.702406

Schneider, R., Wittig, J., Enke, F., & Schipolowski, S. (2023). Schulische und außerschulische Lernbedingungen während der Coronavirus-Pandemie. In P. Stanat, S. Schipolowski, R. Schneider, S. Weirich, S. Henschel & K. A. Sachse (Hrsg.), *IQB-Bildungstrend 2022. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im dritten Ländervergleich* (S. 207–229). Waxmann. https://doi.org/10.31244/9783830997771

Schwippert, K., Kasper, D., Köller, O., McElvany, N., Selter, C., Steffensky, M., & Wendt, H. (Hrsg.). (2020). *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Waxmann. https://doi.org/10.31244/9783830993193

Schwippert, K., Scholz, L. A., Beese, C., Kasper, D., Schulz-Heidorf, K., & Girelli, A.-L. (2020). Ziele, Anlage und Durchführung der Trends in International Mathematics and Science Study (TIMSS 2019). In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 25–55). Waxmann. https://doi.org/10.31244/9783830993193

Schwippert, K., Stubbe, T. C., & Wendt, H. (2015). IGLU/TIMSS: International vergleichende Schulleistungsuntersuchungen im Grundschulbereich von 2001 bis 2011. In H. Wendt, T. C. Stubbe, K. Schwippert & W. Bos (Hrsg.), *IGLU & TIMSS. 10 Jahre international vergleichende Schulleistungsforschung in der Grundschule. Vertiefende Analysen zu IGLU und TIMSS 2001 bis 2011* (S. 19–34). Waxmann.

Statista. (2024a). *Schätzungen der Gesamtanzahl der Flüchtlinge aus der Ukraine nach Grenzübertritten in Folge des Krieges von Februar 2022 bis April 2024*. https://de.statista.com/statistik/daten/studie/1293762/umfrage/anzahl-der-kriegsfluechtlinge-aus-der-ukraine/

Statista. (2024b). *Gesamtzahl der an allgemeinbildenden und berufsbildenden Schulen in Deutschland aufgenommenen geflüchteten Kinder und Jugendlichen aus der Ukraine*. https://de.statista.com/statistik/daten/studie/1308090/umfrage/gefluechtete-kinder-und-jugendliche-aus-der-ukraine-an-deutschen-schulen/

UNESCO Institute for Statistics. (2012). *International Standard Classification of Education. ISCED 2011*. UNESCO Institute for Statistics.

von Davier, M., Fishbein, B., & Kennedy, A. (Hrsg.). (2024). *TIMSS 2023 technical report (methods and procedures)*. TIMSS & PIRLS International Study Center.

Wacker, A., Unger, V., & Rey, T. (2020). „Sind doch Corona-Ferien, oder nicht?“. Befunde einer Schüler*innenbefragung zum „Fernunterricht“. In D. Fickermann & B. Edelstein (Hrsg.), *„Langsam vermisse ich die Schule …“. Schule während und nach der Corona-Pandemie* (S. 79–94). https://doi.org/10.31244/9783830992318.04

Weinert, F. E. (2001). Vergleichende Leistungsmessungen in Schulen. Eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17–31). Beltz.

Wendt, H., Bos, W., Kasper, D., Walzebug, A., Goy, M., & Jusufi, D. (2016). Ziele, Anlage und Durchführung der Trends in International Mathematics and Science Study (TIMSS 2015). In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland* (S. 31–78). Waxmann. https://doi.org/10.25656/01:14022

Wendt, H., Tarelli, I., Bos, W., Frey, K., & Vennemann, M. (2012). Ziele, Anlage und Durchführung der Trends in International Mathematics and Science Study (TIMSS 2011). In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 27–68). Waxmann.

WHO (World Health Organisation). (2024). *Coronavirus disease (COVID-19) pandemic*. https://www.who.int/europe/emergencies/situations/covid-19

Yin, L., & Foy, P. (2021). TIMSS 2023 assessment design. In I. V. S Mullis, M. O. Martin & M. von Davier (Hrsg.). *TIMSS 2023 assessment frameworks*. (S. 71–85). https://timssandpirls.bc.edu/timss2023/frameworks/pdf/T23_Frameworks.pdf

Kapitel 3
Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse

Christoph Selter, Daniel Walter, Aiso Heinze, Johanna Brandt und Daniel Kasper

3.1 Einleitung

Deutschland hat im Jahr 2023 zum fünften Mal an der *Trends in International Mathematics and Science Study* (TIMSS) teilgenommen. Die Hauptergebnisse der letzten Studie von 2019 lassen sich unter Einbezug der Resultate der Studien aus 2007, 2011 und 2015 wie folgt zusammenfassen (Selter et al., 2020):

- *Leistungsmittelwerte*: Viertklässler:innen in Deutschland erzielten 2019 im internationalen Vergleich einen Leistungsmittelwert von 521 Punkten. Anders als in den Studienzyklen von 2007 und 2011 befanden sie sich damit nicht mehr im oberen Drittel der Rangreihe, sondern – wie im Jahr 2015 – nur noch im Mittelfeld der insgesamt 58 teilnehmenden Staaten. Dabei lagen die Testergebnisse für Deutschland signifikant über dem internationalen Mittelwert von 501 Punkten. Im Vergleich zu den Mittelwerten aller teilnehmenden Staaten der EU (527) sowie der beteiligten Staaten der OECD (529) schnitt Deutschland aber wie bei TIMSS 2015 signifikant schwächer ab.
- *Leistungsstreuung*: In den TIMSS-Zyklen 2007, 2011 und 2015 konnte für Deutschland eine im Vergleich zu vielen anderen teilnehmenden Staaten geringe Leistungsstreuung verzeichnet werden. Dieser Befund setzte sich tendenziell auch in TIMSS 2019 fort. Dafür sprach sowohl die Größe der Standardabweichung (70) als auch die Leistungsspanne zwischen den jeweils 5 Prozent schwächsten und 5 Prozent stärksten Viertklässler:innen. Es gab zwölf Staaten, die eine vergleichbare oder sogar höhere ‚Leistungshomogenität' aufwiesen, während in 45 teilnehmenden Staaten eine größere Streuung zu beobachten war.
- *Kompetenzstufen IV und V*: Vergleicht man die Ergebnisse zur prozentualen Besetzung der Kompetenzstufen von 2019 mit denen der vorangegangenen Untersuchungszyklen, so lassen sich für die beiden höchsten Kompetenzstufen IV und V keine statistisch bedeutenden Unterschiede erkennen. Jeweils zwischen

29 und 32 Prozent der Schüler:innen erreichten die Kompetenzstufe IV, jeweils 5 bis 6 Prozent die Stufe V. Insbesondere im Vergleich zu TIMSS 2007 und TIMSS 2011 gelang es vielen Staaten, signifikante Steigerungen der Anteile der Schüler:innen in diesen beiden Kompetenzstufen aufzuweisen.

- *Kompetenzstufen I und II*: Mit Blick auf die Kompetenzstufen I und II ließen sich im Vergleich von TIMSS 2019 zu TIMSS 2015 keine statistisch bedeutsamen Unterschiede konstatieren. Im Vergleich zu TIMSS 2007 und TIMSS 2011 wurde jedoch festgestellt, dass es zu einem signifikanten Anstieg des Anteils der Schüler:innen in beiden Kompetenzstufen (2011) bzw. in der Kompetenzstufe II (2007) gekommen war. Einer ganzen Reihe von Staaten gelang hingegen eine signifikante Reduktion dieser Anteile. Während es in 2007 und 2011 22 bzw. 19 Prozent der Schüler:innen waren, waren es in 2015 und 2019 23 bzw. 25 Prozent der Schüler:innen in Deutschland, deren Testergebnisse den Rückschluss zuließen, dass sie sich in den Kompetenzstufen I und II befanden.
- *Inhaltsbereiche:* Im Jahr 2019 zeigten sich im Bereich *Arithmetik* signifikant schwächere Testleistungen als in 2007, was im Vergleich von 2019 mit den Erhebungen in 2011 bzw. 2015 nicht der Fall war. Im Bereich *Messen und Geometrie* waren im jeweiligen Vergleich zu den vorangegangenen TIMSS-Zyklen keine statistisch bedeutsamen Veränderungen zu verzeichnen. Hingegen waren im Bereich *Daten* im Vergleich der Resultate von 2019 mit den anderen drei Zeitpunkten signifikant niedrigere Punktwerte festzustellen. Insgesamt zeigten die Viertklässler:innen damit eine relative Stärke im Bereich *Messen und Geometrie*. Für die anderen beiden Bereiche zeigten sich relative Schwächen, die auch im Trendvergleich nicht ignoriert werden sollten.
- *Kognitive Anforderungsbereiche*: Im Vergleich der Studienzyklen wurde deutlich, dass 2019 beim *Reproduzieren* signifikant bessere Testergebnisse erzielt wurden als im Jahr 2007, während im Vergleich zu den anderen beiden Erhebungszeitpunkten keine nennenswerten Differenzen zu berichten waren. Im Bereich *Anwenden* ergab der Vergleich mit 2015 keine signifikant unterschiedlichen, der Vergleich zu 2007 und 2011 jedoch signifikant schwächere Ergebnisse. Für den Bereich *Problemlösen* waren im Vergleich von TIMSS 2019 mit den vorangegangenen Studienzyklen keine statistisch signifikanten Unterschiede auszumachen.
- *Einstellung*: Der Vergleich der mittleren Skalenwerte der vier TIMSS-Zyklen ergibt, dass in TIMSS 2019 ein signifikant niedrigerer Wert als in allen anderen Studienzyklen vorlag. Es konnte zudem gezeigt werden, dass in 2019 anteilig signifikant weniger Kinder eine hohe positive Einstellung aufwiesen als noch in 2007. In der Gruppe der Schüler:innen mit niedriger positiver Einstellung zur Mathematik lag für 2019 ein signifikant größerer Leistungswert als für 2011 und 2007 vor.
- *Mathematikbezogenes Selbstkonzept*: Der mittlere Skalenwert für das mathematikbezogene Selbstkonzept hatte sich 2019 im Vergleich zu 2015 kaum verändert. Im Vergleich zu 2011 und 2007 war der Mittelwert jedoch signifikant gesunken, wobei er aber nach wie vor als hoch bezeichnet werden konnte. Darüber hinaus war ein positiver Zusammenhang zwischen dem mathematikbezogenen Selbstkonzept und den mathematischen Kompetenzen für Deutschland in TIMSS 2019 zu verzeichnen.

Vor diesem Hintergrund befasst sich dieses Kapitel mit den Ergebnissen Deutschlands in Mathematik im Rahmen der TIMS-Studie 2023. In Abschnitt 3.2 wer-

den zunächst zentrale Aspekte mathematischer Grundbildung in Deutschland beschrieben, bevor in Abschnitt 3.3 die TIMSS-Rahmenkonzeption zur Erfassung mathematischer Kompetenzen dargelegt wird. Abschnitt 3.4 beschreibt den TIMSS-Test zur Erfassung der mathematischen Kompetenzen. In Abschnitt 3.5 schließlich werden die Ergebnisse von TIMSS 2023 entlang folgender Fragestellungen beschrieben, wobei in jedem Abschnitt auch untersucht wird, wie sich die Ergebnisse von TIMSS 2023 von denen aus dem ersten (2007) und dem letzten (2019) Studienzyklus, an dem Deutschland teilnahm, unterscheiden:

- Welche Mathematikleistungen zeigen Schüler:innen in Deutschland am Ende der vierten Jahrgangsstufe im internationalen Vergleich? (Abschnitt 3.5.1)
- Wie lassen sich die Leistungskennwerte bezüglich der Kompetenzstufen in Deutschland im internationalen Vergleich einordnen? Wie groß sind die Gruppen der auffällig leistungsstarken und leistungsschwachen Kinder? (Abschnitt 3.5.2)
- Welche Ergebnisse erzielen Schüler:innen differenziert nach den mathematischen Teilgebieten *Arithmetik*, *Messen und Geometrie* und *Daten* im internationalen Vergleich? Wie unterscheiden sich die Leistungen der Schüler:innen in Bezug auf die mathematischen Teilgebiete? (Abschnitt 3.5.3)
- Welche Ergebnisse erzielen Schüler:innen differenziert nach den kognitiven Anforderungsbereichen *Reproduzieren*, *Anwenden* und *Problemlösen* im internationalen Vergleich? Wie unterscheiden sich die Leistungen der Schüler:innen in Bezug auf die kognitiven Anforderungsbereiche? (Abschnitt 3.5.4)
- Über welche Einstellung zur Mathematik und über welches mathematische Selbstkonzept verfügen die Schüler:innen am Ende der vierten Jahrgangsstufe? (Abschnitt 3.5.5)

Weitere Befunde zu den in TIMSS 2023 erfassten mathematischen Kompetenzen werden in den Kapiteln 6 bis 8 in diesem Band dargelegt. Diese enthalten Analysen der in TIMSS 2023 untersuchten Kompetenzen vor dem Hintergrund der individuellen Merkmale *Geschlecht*, *soziale Herkunft* sowie *Migrationshintergrund*. Aus diesem Grund werden diese Variablen in diesem Kapitel nicht näher in den Blick genommen.

3.2 Mathematische Grundbildung in Deutschland

In diesem Abschnitt werden wesentliche Aspekte des Konzepts der mathematischen Grundbildung in Deutschland beschrieben. So wird die Grundlage geschaffen, um in den nachfolgenden Abschnitten die Beschreibung der TIMSS-Testkonzeption sowie die Testergebnisse von TIMSS 2023 einordnen zu können.

Im Jahr 2004 wurden die Bildungsstandards Mathematik der *Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland* (KMK) für den Primarbereich verabschiedet und in den folgenden Jahren in den Lehrplänen der einzelnen Länder der Bundesrepublik Deutschland abgebildet. Die KMK hat im Jahr 2019 im Rahmen einer Bedarfsanalyse einen deutlichen Überarbeitungsbedarf festgestellt. Daher wurde im Juni 2020 beschlossen, alle Bildungsstandards für den Primarbereich und die Sekundarstufe I weiterzuentwickeln und stufenübergreifend besser aufeinander abzustimmen. Die überarbeiteten Fassungen der Bildungsstandards Mathematik wurden im Juni 2022 von der KMK verabschiedet (KMK, 2022). Die Länder verpflichteten sich, die weiterentwickelten Bildungsstandards in ihren länderspezifischen Vorgaben zu implementieren. Die

Erhebungen zu TIMSS 2023 fanden im Frühjahr 2023 statt. Da die angestrebte Implementation der Bildungsstandards von 2022 in den Lehrplänen zu diesem Zeitpunkt noch nicht stattgefunden hatte, erfolgt im Weiteren wie bei den vorangegangenen Zyklen der Bezug auf die während der Haupterhebung noch gültigen Bildungsstandards von 2004.

Die Bildungsstandards Mathematik für die Primarstufe legen fest, welche mathematischen Kompetenzen Schüler:innen am Ende der vierjährigen Grundschulzeit erworben haben sollten. Neben der Beschreibung und Operationalisierung inhaltsbezogener und allgemeiner mathematischer Kompetenzen werden die Kompetenzen auch in drei Anforderungsbereichen auf mathematische Konzepte bezogen.

Inhaltsbezogene mathematische Kompetenzen: Die inhaltsbezogenen mathematischen Kompetenzen sind mit den folgenden fünf mathematischen Leitideen (KMK, 2005) verbunden:

- Zahlen und Operationen,
- Raum und Form,
- Muster und Strukturen,
- Größen und Messen sowie
- Daten, Häufigkeit und Wahrscheinlichkeit.

Zur Leitidee *Daten, Häufigkeiten und Wahrscheinlichkeit* beispielsweise sind folgende zwei inhaltsbezogene Kompetenzen formuliert (KMK, 2005):

- Daten erfassen und darstellen sowie
- Wahrscheinlichkeiten von Ereignissen in Zufallsexperimenten vergleichen.

Diese Kompetenzen wiederum werden in Form von Standards konkretisiert. Die Kompetenz *Daten erfassen und darstellen* beispielsweise ist wie folgt gegliedert (KMK, 2005):

- in Beobachtungen, Untersuchungen und einfachen Experimenten Daten sammeln, strukturieren und in Tabellen, Schaubildern und Diagrammen darstellen sowie
- aus Tabellen, Schaubildern und Diagrammen Informationen entnehmen.

Allgemeine mathematische Kompetenzen: Neben den inhaltsbezogenen Kompetenzen beschreiben die Bildungsstandards auch allgemeine, inhaltsübergreifende mathematische Kompetenzen. Die Berücksichtigung allgemeiner mathematischer Kompetenzen erscheint zentral, um einen auf Verständnis ausgerichteten Mathematikunterricht zu gestalten sowie positive Grundhaltungen zum Fach Mathematik zu entwickeln. Die nachfolgenden fünf allgemeinen mathematischen Kompetenzen sind in den Bildungsstandards festgeschrieben (KMK, 2005):

- Problemlösen,
- Kommunizieren,
- Argumentieren,
- Modellieren,
- Darstellen.

Auch die allgemeinen mathematischen Kompetenzen sind in Form von Bildungsstandards konkretisiert. Für *Darstellen* beispielsweise lauten diese (KMK, 2005):

- für das Bearbeiten mathematischer Probleme geeignete Darstellungen entwickeln, auswählen und nutzen,
- eine Darstellung in eine andere übertragen sowie
- Darstellungen miteinander vergleichen und bewerten.

Anforderungsbereiche: Die Bildungsstandards formulieren zudem drei Bereiche, die die kognitiven Anforderungen an Schüler:innen bei der Bearbeitung von Aufgaben beschreiben (KMK, 2005):

- Anforderungsbereich I – *Reproduzieren*: Das Lösen der Aufgabe erfordert Grundwissen und das Ausführen von Routinetätigkeiten.
- Anforderungsbereich II – *Zusammenhänge herstellen*: Das Lösen der Aufgabe erfordert das Erkennen und Nutzen von Zusammenhängen.
- Anforderungsbereich III – *Verallgemeinern und Reflektieren*: Das Lösen der Aufgabe erfordert komplexe Tätigkeiten wie Strukturieren, Entwickeln von Strategien, Beurteilen und Verallgemeinern.

Berücksichtigung der Bildungsstandards in den Lehrplänen der Länder der Bundesrepublik Deutschland: Bezüglich der Umsetzung der Bildungsstandards in den Lehrplänen der einzelnen Länder der Bundesrepublik Deutschland – in einigen auch als *Rahmenpläne*, *Kerncurricula* oder *Fachanforderungen* bezeichnet – kann mit Rückbezug auf die Curriculum-Analyse von TIMSS 2019 eine gute Passung konstatiert werden, welche sich in den folgenden Ergebnissen zusammenfassen lässt (Selter et al., 2020):

- *Inhaltsbezogene mathematische Kompetenzen*: Die Kompetenzbereiche *Zahlen und Operationen*, *Raum und Form* sowie *Größen und Messen* fanden sich in den Lehrplänen aller Länder wieder. Kompetenzen im Bereich *Daten, Häufigkeiten und Wahrscheinlichkeiten* sowie *Muster und Strukturen* wurden nicht in allen Ländern als separate Bereiche aufgeführt, aber als integrale oder übergeordnete Bestandteile in den weiteren Bereichen mitberücksichtigt.
- *Allgemeine mathematische Kompetenzen*: In nahezu allen Ländern der Bundesrepublik Deutschland wurden die prozessbezogenen Kompetenzen – teils zusammengefasst – in eigenständigen Kompetenzbereichen der Lehrpläne aufgeführt. Weniger explizit, aber dennoch unter Berücksichtigung wesentlicher Charakteristika, wurden die allgemeinen mathematischen Kompetenzen inzwischen auch in den Lehrplänen der drei Länder einbezogen, die im Mai 2019 noch keine neuen Lehrpläne mit konkretem Bezug zu den Bildungsstandards eingeführt hatten (Schleswig-Holstein, Bremen, Mecklenburg-Vorpommern).

Auch im Hinblick auf den TIMSS-Zyklus 2023 ist eine gute Passung der Bildungsstandards zu den Lehrplänen der Länder der Bundesrepublik Deutschland zu konstatieren. Dies ist darin begründet, dass lediglich in drei Ländern der Bundesrepublik Deutschland zum Erhebungszeitpunkt (Ende April bis Anfang Juni 2023) für das vierte Schuljahr neue Lehrpläne unterrichtswirksam waren: Sachsen (2019), Sachsen-Anhalt (2019) und Schleswig-Holstein (2018). Ergänzend ist anzumerken, dass in vier weiteren Ländern neue Lehrpläne in Kraft getreten sind: Berlin (2017/2024), Brandenburg (2017/2023), Hamburg (2011/2023) und Nordrhein-Westfalen (2008/2021). Diese traten jedoch teils mit den Schuljahren aufwachsend oder gestaffelt in Kraft und waren zum Zeitpunkt der Erhebung von TIMSS 2023 für die Schüler:innen des vierten Schuljahres noch nicht gültig.

In diesem Abschnitt wurde gezeigt, dass zum Erhebungszeitpunkt, wie bereits zur Erhebung von TIMSS 2019, die in den KMK-Bildungsstandards festgelegten inhalts- und prozessbezogenen Kompetenzen in den Lehrplänen der Länder der Bundesrepublik Deutschland für das Fach Mathematik der Grundschule (4. Jahrgangsstufe) nahezu flächendeckend explizit benannt und hierzu Kompetenzerwartungen beschrieben werden. Die bereits in TIMSS 2019 festgestellte Entwicklung, dass sich die Lehrpläne der 16 Länder zunehmend einander angenähert

haben, setzt sich in den neuen Lehrplänen weiter fort. Einschränkend ist anzumerken, dass die Verankerung entsprechender Kompetenzen in den Lehrplänen kein Garant für die adäquate Implementation in Schulen ist. Hier zeigt sich die typische Übergangsproblematik von intendiertem zu implementiertem Curriculum (siehe Kapitel 2 in diesem Band).

3.3 TIMSS-Rahmenkonzeption zur Erfassung mathematischer Kompetenzen

Eine zentrale Herausforderung der TIMS-Studie besteht darin, die Testinstrumente so zu gestalten, dass die mathematischen Kompetenzen der Schüler:innen in den verschiedenen teilnehmenden Staaten und Regionen möglichst präzise erfasst werden können. Die Analyse der erfassten Kompetenzen ermöglicht dann differenzierte Aussagen über die spezifischen Stärken und Schwächen von Viertklässler:innen aus den verschiedenen Bildungssystemen. Entsprechend sollen die eingesetzten Testaufgaben wesentliche Aspekte der Ziele mathematischer Grundbildung in Deutschland (Abschnitt 3.2) abbilden, um anhand der gewonnenen Daten Erkenntnisse über die mathematischen Kompetenzen von Schüler:innen in Deutschland gewinnen zu können. Um dies zu überprüfen, wird in diesem Abschnitt die TIMSS-Rahmenkonzeption zur Erfassung mathematischer Kompetenzen von Schüler:innen in der Erhebungsrunde 2023 (Mullis et al., 2021) beschrieben. Insbesondere wird dargelegt, inwiefern eine Passung zu den in Deutschland festgeschriebenen Kompetenzerwartungen am Ende der vierjährigen Grundschulzeit vorliegt.

Wie in den bisherigen TIMSS-Zyklen unterscheidet die TIMSS-Rahmenkonzeption einerseits zwischen drei *mathematischen Inhaltsbereichen* (Abschnitt 3.3.1) und andererseits zwischen drei *kognitiven Anforderungsbereichen* (Abschnitt 3.3.2). In beiden Bereichen sind im Vergleich zu den vorangegangenen Zyklen lediglich kleine Veränderungen vorgenommen worden. Eine Vergleichbarkeit der TIMSS-Rahmenkonzeption zwischen TIMSS 2023 und den bisherigen TIMSS-Zyklen ist damit gegeben. Wie bei TIMSS 2019 ist die Rahmenkonzeption für digitale Erhebungsformate geeignet, sodass die Vorteile computerbasierter Testformate genutzt werden können. Während in TIMSS 2019 in einigen Staaten noch papierbasiert erhoben wurde, fand die Erhebung in TIMSS 2023 nahezu vollständig digital statt.

3.3.1 Mathematische Inhaltsbereiche

Die TIMSS-Rahmenkonzeption sieht vor, dass die Testaufgaben eine möglichst große Übereinstimmung mit den Curricula der teilnehmenden Staaten und Regionen aufweisen. Die relevanten mathematischen Inhalte für die Grundschule werden dazu in die folgenden Inhaltsbereiche gegliedert:

- Arithmetik (*Number*),
- Messen und Geometrie (*Measurement and Geometry*) sowie
- Daten (*Data*).

Jeder dieser Inhaltsbereiche ist noch einmal in verschiedene Teilgebiete unterteilt. So enthält etwa der Inhaltsbereich *Daten* die Teilgebiete *Daten lesen und darstel-*

len sowie *Daten interpretieren, kombinieren und vergleichen*. Für jedes Teilgebiet sind schließlich inhaltsbezogene Kompetenzerwartungen spezifiziert.

Tabelle 3.1 zeigt die drei Inhaltsbereiche mit ihren Teilgebieten und den darauf bezogenen Kompetenzerwartungen. Die Differenzierung in drei Inhaltsbereiche bei TIMSS entspricht dabei nur teilweise der feineren Strukturierung der KMK-Bildungsstandards in fünf Leitideen (siehe Abschnitt 3.2). So sind etwa

Tabelle 3.1: Inhaltsbereiche und darauf bezogene Kompetenzerwartungen

Inhaltsbereiche	Teilgebiete
Arithmetik	**1. Natürliche Zahlen** - Stellenwerte bei Zahlen mit bis zu sechs Ziffern erkennen; Zusammenhänge zwischen Zahlendarstellungen herstellen (Worte, Symbole, geometrische Modelle einschließlich Zahlenstrahl); Zahlen vergleichen - Bis zu vierstellige Zahlen addieren und subtrahieren - Multiplizieren (max. dreistellig mal einstellig oder zweistellig mal zweistellig) und dividieren (max. dreistellig durch einstellig) - Aufgaben mit geraden und ungeraden Zahlen oder mit Vielfachen und Teilern lösen; Zahlen runden; Überschlagsrechnungen durchführen - Zwei oder mehr Eigenschaften von Zahlen oder Operationen nutzen **2. Terme, Gleichungen und Beziehungen** - Fehlende Zahlen oder Operationen in einem Zahlensatz finden (z.B. 17 + __ = 29) - Terme oder Gleichungen zur Darstellung von Problemsituationen mit Unbekannten zuordnen oder aufschreiben - Beziehungen in klar definierten Mustern erkennen, beschreiben und nutzen (z. B. die Beziehung zwischen benachbarten Zahlen beschreiben; Zahlenpaare erzeugen, die einer vorgegebenen Regel folgen) **3. Brüche und Dezimalzahlen** - Brüche als Teile eines Ganzen oder als Teile mehrerer Ganzer beschreiben; verschiedene Darstellungen von Brüchen (Worte, Symbole oder geometrische Modelle) in Beziehung setzen; Brüche vergleichen; einfache gleichnamige Brüche mit Nennern 2, 3, 4, 5, 6, 8, 10, 12 oder 100 addieren und subtrahieren - Verschiedene Darstellungen von Dezimalzahlen (Worte, Zahlen oder geometrische Modelle) in Beziehung setzen; Dezimalzahlen vergleichen, ordnen und runden; Dezimalzahlen mit Brüchen in Beziehung setzen; Dezimalzahlen addieren und subtrahieren (mit bis zu zwei Nachkommastellen)
Messen und Geometrie	**1. Messen** - Längen messen, schätzen, addieren und subtrahieren (mm, cm, m, km) - Addieren und Subtrahieren von Massen (g, kg), Volumina (ml, l) und Zeit (min, h); geeignete Einheiten auswählen und Skalen lesen - Umfang von Vielecken, Flächeninhalt von Rechtecken, Flächeninhalt von Figuren, die (von Teilen) von Einheitsquadraten abgedeckt sind, sowie Volumen von Figuren, die mit Einheitswürfeln ausgelegt sind, bestimmen **2. Geometrie** - Zueinander parallele und zueinander senkrechte Geraden erkennen und zeichnen; rechte Winkel und Winkel, die kleiner oder größer sind als der rechte Winkel, erkennen; Winkel gemäß ihrer Größe vergleichen - Elementare Eigenschaften nutzen (einschließlich Achsen- und Drehsymmetrie), um zweidimensionale geometrische Figuren zu beschreiben und zu erzeugen (Kreise, Dreiecke, Vierecke und andere Vielecke) - Elementare Eigenschaften zur Beschreibung von dreidimensionalen geometrischen Figuren (Würfel, Quader, Kegel, Pyramiden und Kugeln), ihrer Unterschiede und den Beziehungen zu ihren zweidimensionalen Repräsentationen nutzen
Daten	**1. Daten lesen und darstellen** - Daten aus Tabellen, Piktogrammen, Balken-, Linien- und Kreisdiagrammen entnehmen - Tabellen, Piktogramme, Balken-, Linien- und Kreisdiagramme erstellen oder ergänzen **2. Daten interpretieren, kombinieren und vergleichen** - Daten interpretieren und sie zur Beantwortung von Fragen nutzen, die über das bloße Entnehmen von Daten hinausgehen - Daten aus zwei oder mehr Quellen kombinieren oder vergleichen; Schlussfolgerungen auf der Grundlage von zwei oder mehr Datensätzen ziehen

 © TIMSS 2023

Aufgaben, die bei den Bildungsstandards der Leitidee *Größen und Messen* zugeordnet werden (z. B. Aufgaben zu Geldwerten, Massen/Gewichten, Längen, Flächen, Volumina), bei TIMSS 2023 entweder dem Inhaltsbereich *Arithmetik* oder dem Inhaltsbereich *Messen und Geometrie* zuzuordnen.

An TIMSS 2023 nehmen 58 Staaten und fünf sogenannte Benchmark-Regionen teil (siehe Kapitel 2 in diesem Band), die unterschiedliche Curricula für die Grundschule aufweisen. Einen Test zu entwickeln, der zu allen Curricula am Ende der vierten Jahrgangsstufe passt, ist nicht möglich. Gleichzeitig wäre ein Test, der sich nur auf die Inhalte bezieht, die in allen Curricula genannt werden, eine zu große Einschränkung und würde den Zielen von TIMSS nicht gerecht. Es wird in der TIMSS-Rahmenkonzeption und folglich auch bei den Testaufgaben also immer Teile geben, die im Curriculum eines teilnehmenden Staates gar nicht oder nur ansatzweise enthalten sind. Für Deutschland trifft dies vor allem auf das Teilgebiet *Brüche und Dezimalzahlen* zu. Die in den Testaufgaben abgebildeten Kompetenzen dieses Bereiches sind überwiegend nicht in den Grundschulcurricula verankert (siehe Abschnitt 3.4.2).

3.3.2 Kognitive Anforderungsbereiche

Die Rahmenkonzeption von TIMSS 2023 umfasst neben den mathematischen Inhaltsbereichen auch kognitive Anforderungsbereiche. Diese beschreiben Denkprozesse, die für die Aufgabenbearbeitung erforderlich sind. Wie in den vorherigen TIMSS-Zyklen werden die folgenden drei kognitiven Anforderungsbereiche unterschieden:

- *Reproduzieren*: Reproduzieren von Wissen, Fertigkeiten und Grundvorstellungen (*Knowing*),
- *Anwenden*: Anwenden von Wissen, Fertigkeiten und Grundvorstellungen in vielfältigen Situationen (*Applying*) sowie
- *Problemlösen*: Problemlösungen finden und begründen, Schlussfolgerungen ziehen und sich mit komplexen Beziehungen zwischen mathematischen Objekten auseinandersetzen (*Reasoning*).

Die Unterteilung der Anforderungsbereiche dient dazu, die kognitive Leistungsfähigkeit der Schüler:innen möglichst breit abzubilden. Dazu würden Aufgaben, die beispielsweise nur das *Reproduzieren* von Wissen, Fertigkeiten und Grundvorstellungen abbilden, nicht ausreichen (Mullis et al., 2021). Im Folgenden werden die kognitiven Anforderungsbereiche – beginnend mit dem *Reproduzieren* in Tabelle 3.2 – durch kognitive Aktivitäten beschrieben und beispielhaft konkretisiert.

Tabelle 3.3 beschreibt die kognitiven Aktivitäten des Bereiches *Anwenden*. Diese gehen über das Abrufen von Wissen hinaus, da auch der situative Gebrauch von Wissen und Fertigkeiten in vielfältigen Anforderungssituationen gefordert ist.

Tabelle 3.4 stellt abschließend die kognitiven Aktivitäten des Anforderungsbereiches *Problemlösen* dar. Diese umfassen vor allem das systematische und logische Denken sowie schlussfolgernd-begründende Aktivitäten.

Die in TIMSS 2023 betrachteten kognitiven Aktivitäten der drei Anforderungsbereiche stimmen weitgehend mit den drei Anforderungsbereichen der KMK-Bildungsstandards (KMK, 2005) überein. Abweichungen bestehen beispielsweise hinsichtlich der Aktivität *Effiziente und geeignete Operationen und Strategien wählen* des TIMSS-Anforderungsbereiches *Anwenden*, die in den

Tabelle 3.2: Kognitive Aktivitäten im Anforderungsbereich *Reproduzieren*

Kognitive Aktivität *Reproduzieren*	**Beschreibung**
Standardwissen abrufen	Definitionen, Begriffe, Zahleigenschaften, Maßeinheiten, geometrische Eigenschaften und Notationen abrufen
Einfache mathematische Objekte und Beziehungen identifizieren	Zahlen, Terme, Mengen und Figuren identifizieren; mathematisch gleichwertige Einheiten identifizieren; Graphen, Tabellen, Texten oder anderen Quellen Informationen entnehmen
Ordnen	Zahlen, Terme, Mengen und Figuren nach gemeinsamen Eigenschaften ordnen und klassifizieren
Rechnen	Arithmetische Operationen mit natürlichen Zahlen, Brüchen, Dezimalzahlen und ganzen Zahlen unter Nutzung algorithmischer Prozeduren ausführen; einfache algebraische Umformungen durchführen

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Tabelle 3.3: Kognitive Aktivitäten im Anforderungsbereich *Anwenden*

Kognitive Aktivität *Anwenden*	**Beschreibung**
Effiziente und geeignete Operationen und Strategien wählen	Effiziente bzw. geeignete Operationen, Strategien und Werkzeuge für das Lösen von Standardaufgaben bestimmen
Informationen darstellen	Daten in Tabellen oder Graphen darstellen; Gleichungen, Ungleichheiten, geometrische Figuren oder Diagramme erstellen, die Problemsituationen modellieren; einander entsprechende Repräsentationen eines gegebenen mathematischen Objektes oder einer Beziehung erzeugen
Durchführen	Geeignete Strategien und Operationen für das Lösen von Aufgaben nutzen

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Tabelle 3.4: Kognitive Aktivitäten im Anforderungsbereich *Problemlösen*

Kognitive Aktivität *Problemlösen*	**Beschreibung**
Analysieren	Beziehungen zwischen Zahlen, Termen, Mengen und Figuren analysieren, beschreiben oder nutzen
Integrieren	Verschiedene Wissenselemente, zueinander zugehörige Darstellungen und Prozeduren miteinander verknüpfen
Verallgemeinern	Verallgemeinernde Aussagen treffen, die Beziehungen in breit anwendbarer Weise darstellen
Begründen	Mathematische Begründungen für eine Strategie oder Lösung geben

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

KMK-Bildungsstandards eher dem Anforderungsbereich I (siehe Abschnitt 3.2) zuzuordnen wäre. Auch fällt die dem höchsten TIMSS-Anforderungsbereich *Problemösen* zugeordnete Aktivität *Integrieren*, die das Verknüpfen verschiedener Wissenselemente, Darstellungen und Prozeduren umfasst, eher in den Anforderungsbereich II der Bildungsstandards (siehe Abschnitt 3.2).

Im Gegensatz zu den KMK-Bildungsstandards wird in der TIMSS-Rahmenkonzeption keine explizite Beschreibung der allgemeinen mathematischen bzw. prozessbezogenen Kompetenzerwartungen als drittes Strukturelement mathematischer Kompetenz angegeben (KMK, 2005). Diese prozessbezogenen Kompetenzen im Sinne der Bildungsstandards finden sich in der TIMSS-Rahmenkonzeption aber in wesentlichen Teilen in den kognitiven Anforderungsbereichen *Anwenden* und vor allem *Problemlösen* wieder (z. B. *Modellieren* im Bereich *Anwenden*, Tabelle 3.3 oder *Argumentieren* als *Begründen* im Bereich *Problemlösen*, Tabelle 3.4).

3.3.3 Aufgabenbeispiele

Da die drei mathematischen Inhaltsbereiche und kognitiven Anforderungsbereiche der TIMSS-Rahmenkonzeption unabhängig voneinander sind, resultieren daraus neun Kombinationen als Grundlage für die Entwicklung bzw. Auswahl der Testaufgaben für TIMSS 2023. Abbildung 3.1 zeigt zu jeder dieser Kombinationen eine Beispielaufgabe.

3.4 Der TIMSS-Test zur Erfassung der mathematischen Kompetenzen

Im Weiteren werden wesentliche Aspekte des TIMSS-Tests zur Erfassung der mathematischen Kompetenzen von Schüler:innen am Ende der Grundschulzeit dargelegt. Zunächst wird der Testaufbau (Abschnitt 3.4.1) beschrieben, gefolgt von einer Betrachtung der curricularen Validität der Testaufgaben (Abschnitt 3.4.2). Anschließend wird in Abschnitt 3.4.3 das Kompetenzstufenmodell vorgestellt, das in dieser Studie zur Beschreibung mathematischer Fähigkeiten verwendet wird.

3.4.1 Testentwicklung und Testaufbau

Die in TIMSS 2023 gestellten Testaufgaben wurden in einem kooperativen Prozess entwickelt. Neben dem *TIMSS & PIRLS International Study Center* am Boston College wirkten die nationalen Studienleitungen sowie Fachdidaktiker:innen und Expert:innen für die Domänen Mathematik und Naturwissenschaften aus den verschiedenen Staaten bei der Itementwicklung mit. Zunächst wurden Aufgaben- und Auswertungsvorschläge durch die jeweiligen Studienleitungen der teilnehmenden Staaten und Regionen entwickelt. Im Nachgang wurden diese durch das *Item Review Committee* (Yin & Foy, 2021) begutachtet und falls notwendig fachgerecht überarbeitet. Abschließend erfolgten Übersetzungsarbeiten, in denen vor allem fach- und bildungssprachliche Spezifika der jeweiligen Sprachen Berücksichtigung fanden. Der daraus entstandene vorläufige Aufgabenpool wurde in

Abbildung 3.1: Beispielaufgaben zu Inhaltsbereichen und kognitiven Anforderungsbereichen

Arithmetik

Reproduzieren

Welche Zahl ist gerade **und** ein Vielfaches von 7?

- (A) 21
- (B) 24
- (C) 42
- (D) 77

ME61246 (int. .53/nat. .57)

Bei dieser Aufgabe wird Grundwissen zum Zahlverständnis erhoben, indem zwei Zahleigenschaften der passenden Zahl zugeordnet werden müssen (Standardwissen abrufen).

Anwenden

Ziehe in jedes Kästchen eine Zahlenkarte, sodass die Aussagen stimmen.

3 5 7 8

5 + [8] **ist größer als 12.**

20 - [3] **ist größer als 15.**

ME71036 (int. .62/nat. .76)

Bei dieser Aufgabe sollen die jeweils passenden Zahlenkarten ausgewählt werden, damit die darunter stehenden Aussagen korrekt sind (Durchführen).

Problemlösen

Mark und Ole spielen ein Spiel mit diesen 8 Karten.

1 2 3 4 5 6 7 8

Sie legen die Karten mit der Vorderseite nach unten, sodass man die Zahlen nicht sieht. Dann nimmt sich jeder Spieler 2 Karten. Der Spieler mit der größten Summe gewinnt.

Mark nimmt die Karte mit der **8** und die Karte mit der **2**. Sie ergeben die Summe 10.

Ole nimmt zuerst die Karte mit der **4**.

Kann Ole das Spiel gewinnen?
(Klicke das Kästchen an.)

- (A) Ja
- (B) Nein

Erkläre deine Antwort.

Er braucht die 7, weil 4+7 gleich 11 ist und Mark nur 10 hatte. Mit den anderen Karten hätte Ole eine kleinere Summe als Mark.

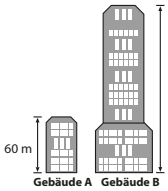
ME61256 (int. .43/nat. .56)

Hier soll im Rahmen einer beschriebenen Spielsituation, in der die Addition einstelliger Zahlen Lerngegenstand ist, ein Spielausgang antizipiert werden. Dazu müssen Beziehungen zwischen Zahlen und Termen genutzt werden (Analysieren).

Messen/Geometrie

Reproduzieren

Die Höhe von Gebäude A beträgt 60 Meter.



Welche Schätzung der Höhe von Gebäude B ist die beste?

- (A) 120 m
- (B) 150 m
- (C) 180 m
- (D) 210 m

ME71068 (int. .47/nat. .52)

Bei dieser Aufgabe soll die Höhe eines Gebäudes auf der Grundlage der angegebenen Höhe eines anderen Gebäudes geschätzt werden. Hierzu ist Grundwissen über die Verwendung und die Größenordnung von Maßeinheiten erforderlich (Standardwissen abrufen).

Anwenden

Kimia und Christin messen die Größe ihres Klassenzimmers aus: Sie zählen ihre Schritte von einem Ende des Zimmers bis zum anderen.

Sie gehen dabei ohne Abstand zwischen den Füßen, so wie hier:



A. Kimias Füße sind 20 cm lang. Sie braucht 40 Schritte, um die Länge des Klassenzimmers abzugehen.

Was ist die Länge ihres Klassenzimmers?

Antwort: [800] cm

ME71075A (int. .42/nat. .50)

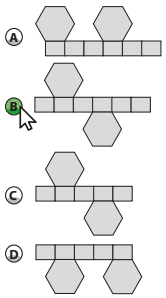
Hier muss die Länge eines Klassenzimmers, das zuvor durch nahtlos aneinandergereihte Schritte ausgemessen wurde, in cm angegeben werden. Hierzu sind die nötigen Angaben vorgegeben (Anzahl der Schritte sowie die Länge des Fußes des ausmessenden Kindes). Das Wissen zum Verrechnen der Angaben ist in dieser Aufgabe anzuwenden (effiziente und geeignete Operationen und Strategien wählen).

Problemlösen

Julia möchte so eine Schachtel basteln wie auf dem Bild unten.



Aus welchem Stück Pappe kann Julia die Schachtel falten?



ME61108 (int. .48/nat. .51)

Bei dieser Aufgabe soll entschieden werden, bei welchem der vier Netze es sich um ein Schachtelnetz handelt. Dazu müssen die vorliegenden Einzelinformationen der Darstellungen miteinander verknüpft und mit der Prozedur des Faltens in der Beziehung gesetzt werden (Integrieren).

Daten

Reproduzieren

In dieser Tabelle stehen die Punktzahlen von einigen Schülerinnen und Schülern aus drei Runden eines Spiels.

	Marco	Theo	Junie	Jenny
Runde 1	19	18	20	11
Runde 2	17	12	14	9
Runde 3	15	16	13	16

A. Wer hat in Runde 2 die meisten Punkte gemacht?

- (A) Marco
- (B) Theo
- (C) Junie
- (D) Jenny

ME71135A (int. .81/nat. .86)

Hier sollen relevante Werte aus einer Tabelle abgelesen und in Beziehung gesetzt werden, wozu Grundwissen im Umgang mit Tabellen notwendig ist (einfache mathematische Objekte und Beziehungen identifizieren).

Anwenden

Schülerinnen und Schüler in einer Klasse haben drei verschiedene Origami-Tiere aus blauem, rotem und gelbem Papier gebastelt. In der Tabelle steht die Anzahl der Tiere, die aus dem Papier in den einzelnen Farben gebastelt wurden.

Tier	Papierfarbe: Blau	Rot	Gelb
Schildkröte	8	4	3
Giraffe	3	2	10
Fisch	10	6	11

Fülle die Tabelle aus, indem du dieses Rätsel löst:

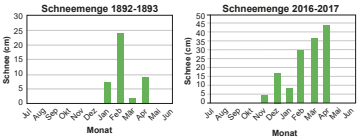
- Es gibt dieselbe Anzahl von blauen Fischen wie von gelben Giraffen.
- Es gibt dieselbe Anzahl von roten Fischen wie von den anderen zwei roten Tieren zusammen.
- Es gibt insgesamt 24 gelbe Tiere.

ME71175 (int. .51/nat. .58)

Bei dieser Aufgabe ist eine Tabelle zu vervollständigen. Hierzu müssen die dazu zu nutzenden Vorgaben mit den vorhandenen Informationen verknüpft werden (Informationen darstellen).

Problemlösen

Diese Diagramme zeigen die Schneemenge in Zedland in den Jahren 1892-1893 im Vergleich zu 2016-2017.



Julia glaubt, dass es im Februar 1893 mehr Schnee gab als im Februar 2017.

Erkläre, warum Julia sich irrt.

Julia könnte sich nur die Größe der Balken angesehen haben. Vielleicht hat sie nicht gesehen, dass das erste Diagramm bis 30 cm und das zweite bis 50 cm geht.

ME71201 (int. .18/nat. .20)

Bei dieser Aufgabe sind Daten aus zwei Liniendiagrammen miteinander zu verknüpfen. Hierbei sind insbesondere die unterschiedlichen Skalierungen zu beachten (Begründen).

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

Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: https://www.iea.nl/data-tools/permission-requests

© TIMSS 2023

einem Feldtest eingesetzt, der die Grundlage für die erneute Begutachtung und Überarbeitung der Testaufgaben bildete (siehe Kapitel 2 in diesem Band).

Der TIMSS-Test 2023 bestand aus 187 Items. Im Vergleich zu den vergangenen Studienzyklen liegen damit etwas mehr Testaufgaben vor. Dies ist im Kern darin begründet, dass mit der Umstellung auf ein digitales Erhebungsformat neuerdings auch komplexere Erkundungsaufgaben, die lebensweltliche oder Laborsituationen simulieren, mit in die Skalierung aufgenommen wurden. Dabei handelt es sich um insgesamt 29 Items. Bereits in TIMSS 2019 wurden entsprechende Testaufgaben eingesetzt, jedoch noch nicht in die Skalierung aufgenommen. Dies hatte das Ziel, die Konstruktvalidität der Aufgaben zunächst zu prüfen. Da diese im Rahmen von TIMSS 2019 nachgewiesen werden konnte, wurde für TIMSS 2023 entschieden, auch die komplexeren Erkundungsaufgaben mit in die Skalierung aufzunehmen. Wie bereits in TIMSS 2019 wurden vier Testaufgaben aufgrund nicht geeigneter statistischer Kennwerte aus der Skalenbildung ausgeschlossen. Daher basieren die in diesem Kapitel dargestellten Befunde auf den Bearbeitungen von 183 Mathematikaufgaben.

Damit aussagekräftige Trendvergleiche zwischen dem aktuellen und den vorherigen Studienzyklen gezogen werden können, muss zumindest ein Teil der Testaufgaben über die Erhebungsrunden hinweg identisch sein. Für TIMSS 2023 bedeutet dies, dass 108 der 187 Testaufgaben bereits in vergangenen Erhebungsrunden eingesetzt wurden. Analog beinhaltete der TIMSS-Test des Jahres 2019 in ähnlichem Umfang Testaufgaben aus den vorangegangenen Erhebungsrunden. Dementsprechend kann eine Vergleichbarkeit der Daten durch eine geeignete Kalibrierung der studienspezifischen Skalierungen über die Erhebungsrunden hinweg als weitgehend gewährleistet betrachtet werden. Die Antwortformate der 187 Items waren jeweils zur Hälfte im Multiple-Choice-Format bzw. in einem offenen Format zu beantworten.

In Tabelle 3.5 ist die Verteilung der Testaufgaben auf die mathematischen Inhaltsbereiche für die Studienzyklen 2007, 2019 und 2023 dargestellt. Die auf die Studienzyklen der Jahre 2007 und 2019 bezogenen Daten wurden aus dem Berichtsband von TIMSS 2019 entnommen (Selter et al., 2020). Es zeigt sich, dass für den aktuellen Studienzyklus 51 Prozent der Testaufgaben dem Bereich *Arithmetik*, 27 Prozent dem Bereich *Messen und Geometrie* sowie 22 Prozent dem Bereich *Daten* zuzuordnen sind. Damit stimmt die Verteilung der Testaufgaben auf die drei Inhaltsbereiche im Wesentlichen mit denen aus den bisherigen Studienzyklen überein. Leichte Verschiebungen der Anteile in den Bereichen *Messen und Geometrie* sowie *Daten* sind durch veränderte Richtlinien der TIMSS-Rahmenkonzeption begründet. Diese sieht seit 2019 vor, die Verteilung der Testaufgaben 50-30-20 statt bislang 50-35-15 für die drei Bereiche anzuwenden (Philpot et al., 2021).

Bezogen auf die kognitiven Anforderungsbereiche ist festzustellen, dass in TIMSS 2023 32 Prozent der Testaufgaben das *Reproduzieren*, 46 Prozent das *Anwenden* sowie 22 Prozent das *Problemlösen* adressieren. Im Vergleich zu den Vorgaben der TIMSS-Rahmenkonzeption, die eine Verteilung von 40-40-20 vorsieht, ist das *Anwenden* damit etwas häufiger und der Bereich *Reproduzieren* etwas seltener gefordert.

Tabelle 3.6 zeigt die Verteilung der Testaufgaben auf die drei kognitiven Anforderungsbereiche für die drei ausgewählten TIMSS-Erhebungsrunden.

Tabelle 3.5: Verteilung der Testaufgaben auf die mathematischen Inhaltsbereiche

Mathematische Inhaltsbereiche	2007		2019		2023	
	n	%	*n*	%	*n*	%
Arithmetik	93	52	83	49	94	51
Messen/Geometrie	60	34	52	30	49	27
Daten	26	15	36	21	40	22
Gesamt	179	100	171	100	183	100

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Tabelle 3.6: Verteilung der Testaufgaben auf die kognitiven Anforderungsbereiche

Kognitive Anforderungsbereiche	2007		2019		2023	
	n	%	*n*	%	*n*	%
Reproduzieren	69	39	59	35	58	32
Anwenden	70	39	74	43	85	46
Problemlösen	40	22	38	22	40	22
Gesamt	179	100	171	100	183	100

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

3.4.2 Curriculare Validität[1]

In den Abschnitten 3.3.1 und 3.3.2 wurde auf die festgestellte hohe Passung zwischen den curricularen Vorgaben der Bildungsstandards für das Fach Mathematik sowie der Rahmenkonzeption in TIMSS 2023 eingegangen. Dies war bereits für die vergangenen Studienzyklen zu verzeichnen, da die inhaltliche Ausrichtung der TIMSS-Rahmenkonzeption weitgehend unverändert geblieben ist und die KMK-Bildungsstandards von 2004 zu den jeweiligen Erhebungszeitpunkten durchgehend in Kraft waren.

Gleichwohl bleibt es in internationalen Schulleistungsstudien wie TIMSS nicht aus, dass in den teilnehmenden Staaten und Regionen einige Testaufgaben als curricular nicht valide einzuschätzen sind. Zwar wird im Zuge der Testentwicklung darauf geachtet, dass die Testaufgaben durch die Curricula möglichst vieler Staaten hinreichend abgedeckt sind. Allerdings ist die Entwicklung eines Aufgabenpools, der ausschließlich curricular valide Testaufgaben für alle teilnehmenden Staaten und Regionen enthält, nicht realisierbar. Ursächlich hierfür ist die Diversität curricularer Vorgaben. Daher erfolgt in TIMSS für jede ein-

1 Wesentliche Teile dieses Abschnittes sind bereits im Kapitel *Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse* im Berichtsband von 2019 (Selter et al., 2020) bearbeitet worden und werden im vorliegenden Berichtsband für den Studienzyklus von 2023 adaptiert.

zelne Testaufgabe eine Überprüfung der curricularen Validität, die auf nationaler Ebene durch Fachdidaktiker:innen vorgenommen wird. Dieser Arbeitsschritt wird in den TIMSS-Zyklen als *Test-Curriculum Matching Analysis* (TCMA) bezeichnet. Die Testaufgaben wurden, wie in vorherigen Zyklen, für die Bundesrepublik Deutschland dahingehend überprüft, ob sie mit den KMK-Bildungsstandards von 2004, dem Lehrplan Mathematik für die Grundschule des bevölkerungsstärksten Landes der Bundesrepublik Deutschland, Nordrhein-Westfalen (Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen, 2008), und einer Auswahl weit verbreiteter Schulbücher vereinbar sind (Selter et al., 2020).

Die TCMA ergab für Deutschland in TIMSS 2023 einen Anteil von 17 Prozent curricular nicht valider Aufgaben. Damit ist der Anteil curricular nicht valider Testaufgaben in der aktuellen Erhebungsrunde geringer als in den bisherigen Studienzyklen. So lag der Anteil für TIMSS 2007 bei 25 Prozent, für TIMSS 2011 bei 21 Prozent, für TIMSS 2015 bei 19 Prozent und beim vorangegangenen Zyklus bei 23 Prozent (Selter et al., 2020).

Wie in den vorherigen Erhebungsrunden wurde auch bei der aktuellen TCMA die curriculare Validität jeder einzelnen Aufgabe anhand der folgenden drei Merkmale überprüft:

- Grafische Darstellungsformen,
- Mathematische Begriffe, Verfahren und Sachverhalte sowie
- Bezeichnungen.

Seit TIMSS 2015 haben sich keine bedeutenden Unterschiede bezogen auf die in Tabelle 3.7 aufgeführten Merkmale curricular nicht valider Aufgaben ergeben. Daher wird, wie auch im vergangenen Berichtsband, an dieser Stelle auf eine detaillierte Darstellung der Merkmale verzichtet und auf den Berichtsband der Erhebungsrunde 2011 verwiesen, in dem entsprechende Erläuterungen zusammengestellt sind (Selter et al., 2012). Stattdessen sind in Abbildung 3.2 exemplarisch zwei curricular nicht valide Testaufgaben dargestellt. Während die Aufgabe aus dem Themengebiet *Terme, Gleichungen und Beziehungen* Kenntnisse über das Rechnen mit Klammern voraussetzt, zielt die Aufgabe aus der *Geometrie* auf den Vergleich von Winkelgrößen ab.

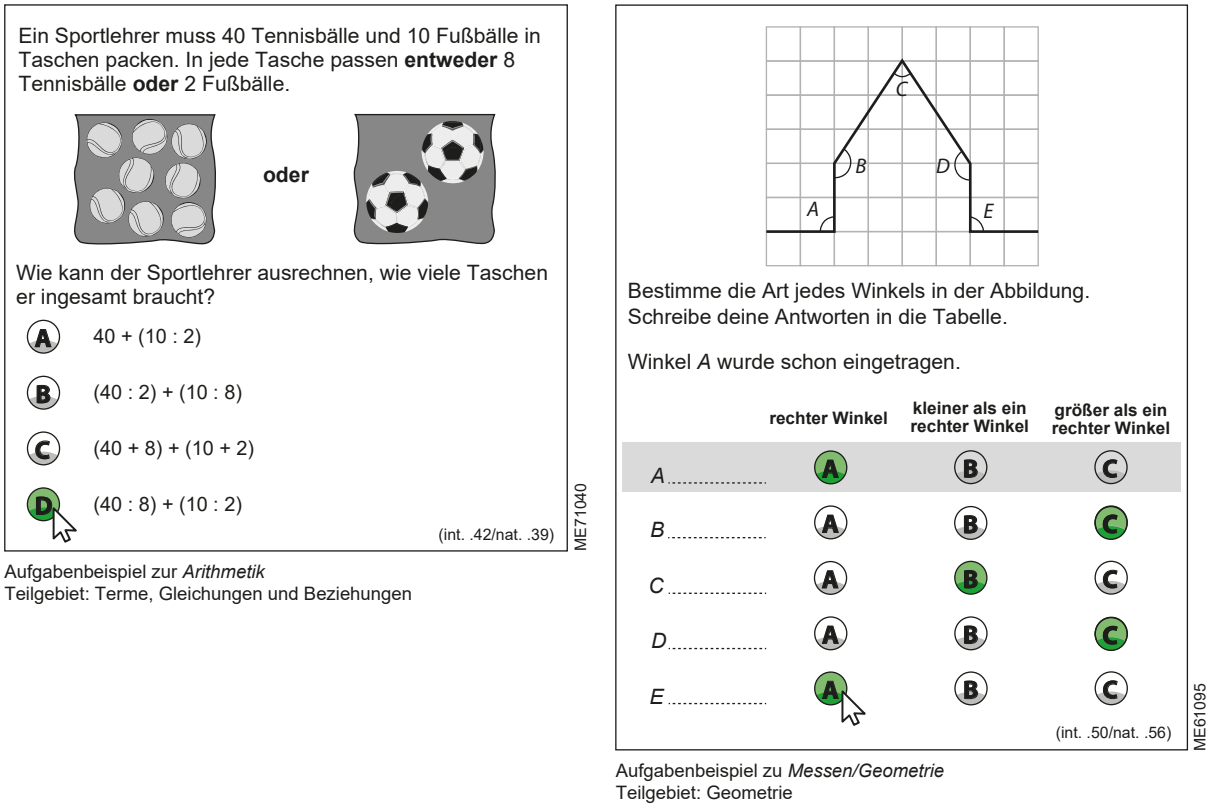
Tabelle 3.7: Merkmale von curricular nicht validen Testaufgaben in Mathematik

Aufgabenmerkmale	Curricular nicht valide Inhalte
Grafische Darstellungsformen	Kreisdiagramm, Liniendiagramm
Mathematische Begriffe, Verfahren und Sachverhalte	*Brüche*: Bruchteile, Bruchvergleich (kleiner, gleich, größer), Addition oder Subtraktion von Brüchen
	Dezimalbrüche: Vergleich von Dezimalbrüchen (kleiner, gleich, größer), Addition oder Subtraktion von Dezimalbrüchen
	Gleichungen: der Form a : b = b mit b als Unbekannter, Rechnen mit Klammern
	Winkel: Maßzahlen (90°, 180°), Vergleich von Winkelgrößen (kleiner, größer)
Bezeichnungen	Drehsymmetrisch, Quadratmeter, Quadratzentimeter, Seite A

 © TIMSS 2023

In Tabelle 3.8 ist die Verteilung curricular nicht valider Testaufgaben auf die drei mathematischen Inhaltsbereiche für die Studienzyklen 2007, 2019 und 2023 zusammengestellt.[2] Für die aktuelle Erhebungsrunde zeigt sich, dass für den Bereich *Arithmetik* 17 Prozent der Testaufgaben curricular nicht valide sind. Dieser Anteil ist niedriger als in den vorangegangenen Erhebungsrunden. Im Inhaltsbereich *Messen und Geometrie* liegt der Anteil curricular nicht valider Aufgaben bei 10 Prozent und ist damit nahezu identisch zur vergangenen Erhebungsrunde von 2019. Im Vergleich zu den anderen Inhaltsbereichen ist der Anteil bei *Daten* wie auch schon in TIMSS 2019 am größten. 23 Prozent der Testaufgaben wurden als curricular nicht valide eingeschätzt, was vorrangig darauf zurückzuführen ist, dass die betreffenden neun Aufgaben aus dem Bereich *Daten* Kreis- bzw. Liniendiagramme thematisierten.

Abbildung 3.2: Beispiele curricular nicht valider Testaufgaben



Aufgabenbeispiel zur *Arithmetik*
Teilgebiet: Terme, Gleichungen und Beziehungen

Aufgabenbeispiel zu *Messen/Geometrie*
Teilgebiet: Geometrie

Die Werte in Klammern geben die relativen internationalen und nationalen Lösungshäufigkeiten an.
Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: https://www.iea.nl/data-tools/permission-requests

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Tabelle 3.9 zeigt die Verteilung der curricular nicht validen Testaufgaben auf die drei kognitiven Anforderungsbereiche. 19 Prozent der Aufgaben aus dem Bereich *Reproduzieren*, 18 Prozent aus dem Bereich *Anwenden* sowie 10 Prozent aus dem Bereich *Problemlösen* wurden als curricular nicht valide eingeschätzt. Damit ist der Anteil curricular nicht valider Testaufgaben im Bereich *Problemlösen* erneut am geringsten, während für *Reproduzieren* sowie *Anwenden* die Anteile curricular nicht valider Testaufgaben etwas höher sind.

2 Aufgrund nachträglicher Ausschlüsse einzelner Items haben sich im Vergleich zur Berichtslegung von TIMSS 2007 (Walther et al., 2008) leichte Verschiebungen der Prozentwerte zur curricularen Validität der Testaufgaben ergeben.

Tabelle 3.8: Anteile curricular nicht valider Testaufgaben nach mathematischen Inhaltsbereichen

Mathematische Inhaltsbereiche	2007			2019			2023		
	Gesamt	nicht valide		Gesamt	nicht valide		Gesamt	nicht valide	
	n	*n*	%	*n*	*n*	%	*n*	*n*	%
Arithmetik	93	20	22	83	22	27	94	16	17
Messen/Geometrie	60	20	33	52	7	13	49	5	10
Daten	26	5	19	36	10	28	40	9	23
Gesamt	179	45	25	171	39	23	183	30	16

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Tabelle 3.9: Anteile curricular nicht valider Testaufgaben nach kognitiven Anforderungsbereichen

Kognitive Anforderungsbereiche	2007			2019			2023		
	Gesamt	nicht valide		Gesamt	nicht valide		Gesamt	nicht valide	
	n	*n*	%	*n*	*n*	%	*n*	*n*	%
Reproduzieren	69	25	36	59	15	25	58	11	19
Anwenden	70	13	19	74	19	26	85	15	18
Problemlösen	40	7	18	38	5	13	40	4	10
Gesamt	179	45	25	171	39	23	183	30	16

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

In den vergangenen Erhebungsrunden von 2007, 2011 und 2015 wurden die Leistungsmittelwerte der teilnehmenden Staaten bezogen auf die für Schüler:innen in Deutschland curricular validen Testaufgaben in separaten Berechnungen ermittelt und in den nationalen Berichtsbänden analysiert. Es wurde der Frage nachgegangen, inwiefern der Ausschluss curricular nicht valider Testaufgaben Einfluss auf das Testergebnis der Schüler:innen in Deutschland ausübt. Es ließen sich für Deutschland zwar nominelle Verbesserungen um wenige Rangplätze nachweisen, jedoch keine bedeutenden Unterschiede der Rangreihenfolge der Staaten bei den jeweiligen Subtests im Vergleich zu den Rangreihen der Gesamttests (Selter et al., 2016). Daher wird im nationalen Berichtsband zu TIMSS 2023 wie auch schon in der Erhebungsrunde von 2019 auf eine entsprechende Berechnung verzichtet. In der Projektdokumentation der internationalen Studienleitung sind diese Berechnungen für TIMSS 2023 einzusehen (von Davier et al., 2024). Hier wurden auf der Grundlage der für die einzelnen Staaten curricular validen Testaufgaben separate Skalen berechnet und die Leistungsdaten für jeden Staat ermittelt.

3.4.3 Kompetenzstufen in Mathematik

Wie in den vorherigen TIMSS-Zyklen werden die Leistungsstände der Schüler:innen der vierten Klasse zunächst mithilfe bestimmter statistischer Kennwerte beschrieben. Diese ermöglichen zwar Aussagen über relative Unterschiede der Leistungsstände, bieten jedoch keine inhaltlichen Einblicke in das mathematische Wissen und Können der teilnehmenden Schüler:innen. Dazu wird bei TIMSS das Benchmark-Verfahren genutzt. Es ermöglicht die Bestimmung von Fähigkeitswerten der Schüler:innen sowie von Schwierigkeitswerten der Aufgaben auf derselben Skala.

Eine Methode, die auch in Kapitel 2 in diesem Band thematisiert wird, unterteilt die TIMSS-Skala in fünf Intervalle. Diese Intervalle werden durch Teilungspunkte, die als Benchmarks bezeichnet werden, erzeugt. Jedes Intervall kann durch spezifische mathematische Kompetenzen beschrieben werden, die für die Lösung von Aufgaben auf dem entsprechenden Niveau notwendig sind. Diese fünf Intervalle werden als aufeinander aufbauende und hierarchisierte Kompetenzstufen bezeichnet.

Die Beschreibung einer solchen Kompetenzstufe erfolgt folgendermaßen: Eine Testaufgabe wird beispielsweise der Kompetenzstufe III zugeordnet, wenn mindestens 65 Prozent der Schüler:innen, die den unteren Benchmark-Wert dieser Stufe erreichen, die Aufgabe korrekt lösen, wobei höchstens 50 Prozent der Schüler:innen auf Kompetenzstufe II dies tun. Anschließend werden in einem weiteren Analyseprozess die charakteristischen Anforderungen der Kompetenz aus den Spezifika der identifizierten Testaufgaben abgeleitet.

Tabelle 3.10 beschreibt die fünf Kompetenzstufen für die Gesamtskala Mathematik. Eine detaillierte Erklärung des Kompetenzstufenmodells ist im internationalen Berichtsband verfügbar (von Davier et al., 2024). Die angegebenen mathematischen Kompetenzen stimmen bis auf geringfügige Änderungen mit denen der früheren TIMSS-Zyklen überein.

Die Abbildungen 3.3 und 3.4 enthalten Beispielaufgaben[3] zur Illustration der Kompetenzstufen. In den Abbildungen 3.3 und 3.4 sind für jede Beispielaufgabe zwei Werte in Klammern angegeben. Diese sind die relativen Lösungshäufigkeiten auf internationaler bzw. nationaler Ebene für TIMSS 2019. So wurde im Jahr 2019 die erste Aufgabe der Abbildung 3.3 von 14 Prozent aller an TIMSS teilnehmenden Schüler:innen gelöst, während 9 Prozent der Kinder in Deutschland zu einer korrekten Lösung gekommen sind. Um eine internationale Vergleichbarkeit des Kompetenzstufenmodells zu gewährleisten, erfolgt die Zuordnung der Testaufgaben zu einer der Kompetenzstufen auf Basis internationaler Leistungsdaten.

In Bezug auf die Zuordnung der Testaufgaben zu den Kompetenzstufen ist zu berücksichtigen, dass in einzelnen teilnehmenden Staaten und Regionen höhere Lösungshäufigkeiten bei den Testaufgaben beobachtet wurden als im internationalen Durchschnitt. Dies kann zur Konsequenz haben, dass Testaufgaben einer höheren Kompetenzstufe in den einzelnen Staaten und Regionen eigentlich einer niedrigeren Kompetenzstufe zuzuordnen wären.

---

3 Zum Zeitpunkt der Erstellung des nationalen Berichtsbandes standen die für die Zuordnung der TIMSS-2023-Aufgaben zu den Kompetenzstufen erforderlichen Informationen noch nicht zur Verfügung, sodass in den Abbildungen 3.3 und 3.4 Beispielaufgaben und Daten aus dem vorangegangenen Studienzyklus von 2019 (Selter et al., 2020) verwendet wurden.

Tabelle 3.10: Beschreibung der fünf Kompetenzstufen für die Gesamtskala Mathematik

Kompetenzstufe V (ab 625): fortgeschritten
Die Schüler:innen können Informationen auswählen und in Beziehung setzen, um geeignete Maßnahmen zur Lösung von Problemen zu entwickeln. Sie können die Ergebnisse von Berechnungen in Problemkontexten interpretieren, eine Vielzahl von Termen und Mustern formulieren sowie Brüche und Dezimalzahlen in Beziehung setzen. Sie können Maßangaben schätzen und in Beziehung setzen, ihr Wissen über zwei- und dreidimensionale Formen anwenden, einfache Eigenschaften von Linien und Winkeln identifizieren sowie ein grundlegendes Verständnis von Fläche und Umfang bei einfachen Formen zeigen. Die Schüler:innen können Daten interpretieren und Entscheidungen im Umgang mit Daten treffen, die in unterschiedlichen Kontexten gegeben werden.
Kompetenzstufe IV (550–624): hoch
Die Schüler:innen setzen Konzepte oder Darstellungen in erweiterten Kontexten in Beziehung. Sie können ihr Wissen über die Eigenschaften von natürlichen Zahlen anwenden, um eine Lösung zu begründen. Sie zeigen ein Verständnis für Zahlenstrahl, Vielfache, Faktoren, Runden von Zahlen und Rechnen mit Brüchen und Dezimalzahlen. Die Schüler:innen können Aufgaben zum Messen in unterschiedlichen Kontexten lösen. Sie können zweidimensionale Formen zu bislang unbekannten dreidimensionalen Figuren in Beziehung setzen und zeigen ein grundlegendes Verständnis für Winkel. Die Schüler:innen können Merkmale der Darstellungen von Daten interpretieren und Daten in einer Vielzahl von Diagrammen darstellen.
Kompetenzstufe III (475–549): durchschnittlich
Die Schüler:innen zeigen mathematische Kenntnisse in einfachen Situationen und stellen Zusammenhänge zwischen Repräsentationen her. Sie können Berechnungen mit dreistelligen natürlichen Zahlen in verschiedenen Situationen durchführen. Sie können einfache Dezimalzahlen addieren und ordnen. Die Schüler:innen können Strecken messen und dreidimensionale Formen beschreiben. Sie können Daten aus verschiedenen Quellen verwenden, um Zusammenhänge zwischen ihnen herzustellen.
Kompetenzstufe II (400–474): niedrig
Die Schüler:innen zeigen ein grundlegendes mathematisches Verständnis. Sie können natürliche Zahlen mit bis zu drei Stellen addieren und subtrahieren, einstellige natürliche Zahlen multiplizieren und dividieren sowie einfache Textaufgaben lösen. Sie können grundlegende Messideen und Eigenschaften einfacher geometrischer Formen anwenden. Die Schüler:innen können Daten aus verschiedenen Darstellungen ablesen und einfache Säulendiagramme ausfüllen.
Kompetenzstufe I (unter 400): rudimentär
Die Schüler:innen verfügen über rudimentäres schulisches Anfangswissen. Auf diesem Niveau können viele Kinder einfache Routineaufgaben mit Grundrechenarten ausführen, jedoch selbst einfache Aufgaben des Niveaus II nur gelegentlich oder ansatzweise lösen.

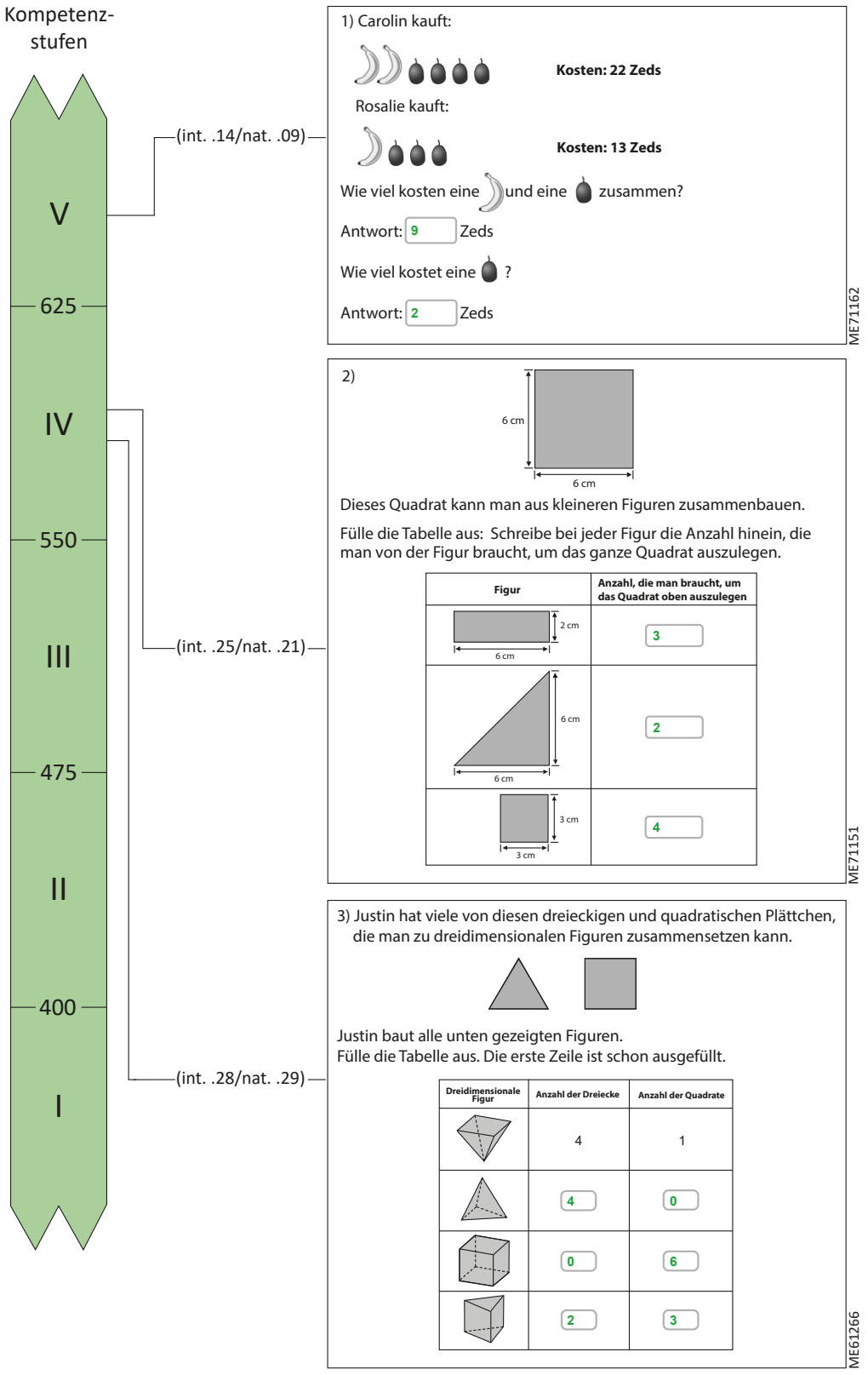
IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Nachfolgend wird erläutert, welche mathematischen Kompetenzen Schüler:innen benötigen, um die Aufgaben aus den Abbildungen 3.3 und 3.4 erfolgreich lösen zu können:

Kompetenzstufe V: Die erste Aufgabe besteht in der Bestimmung der Kosten zweier verschiedener Obstsorten unter Berücksichtigung vorgegebener Daten. Um die gesuchten Größen zu berechnen, sind mehrere Lösungsschritte erforderlich, da die Berechnung auf Basis grundlegender Routinetätigkeiten nicht möglich ist. Stattdessen ist ein geschicktes Ableiten der Ergebnisse notwendig.

Kompetenzstufe IV: In der zweiten Aufgabe werden den Kindern ein Quadrat sowie die Seitenlängen desselben präsentiert. Im Anschluss sollen die Kinder anhand dreier Figuren entscheiden, wie viele der jeweiligen Figuren notwendig sind, um das gegebene Quadrat vollständig auszulegen. Um diese Aufgabe adäquat bearbeiten zu können, müssen die Kinder über die Fähigkeit zum mentalen Operieren mit ebenen Figuren sowie über die Kompetenz verfügen, geometrische Eigenschaften der Figuren zu nutzen.

Abbildung 3.3: Kompetenzstufen und Beispielaufgaben (Gesamtskala Mathematik) I

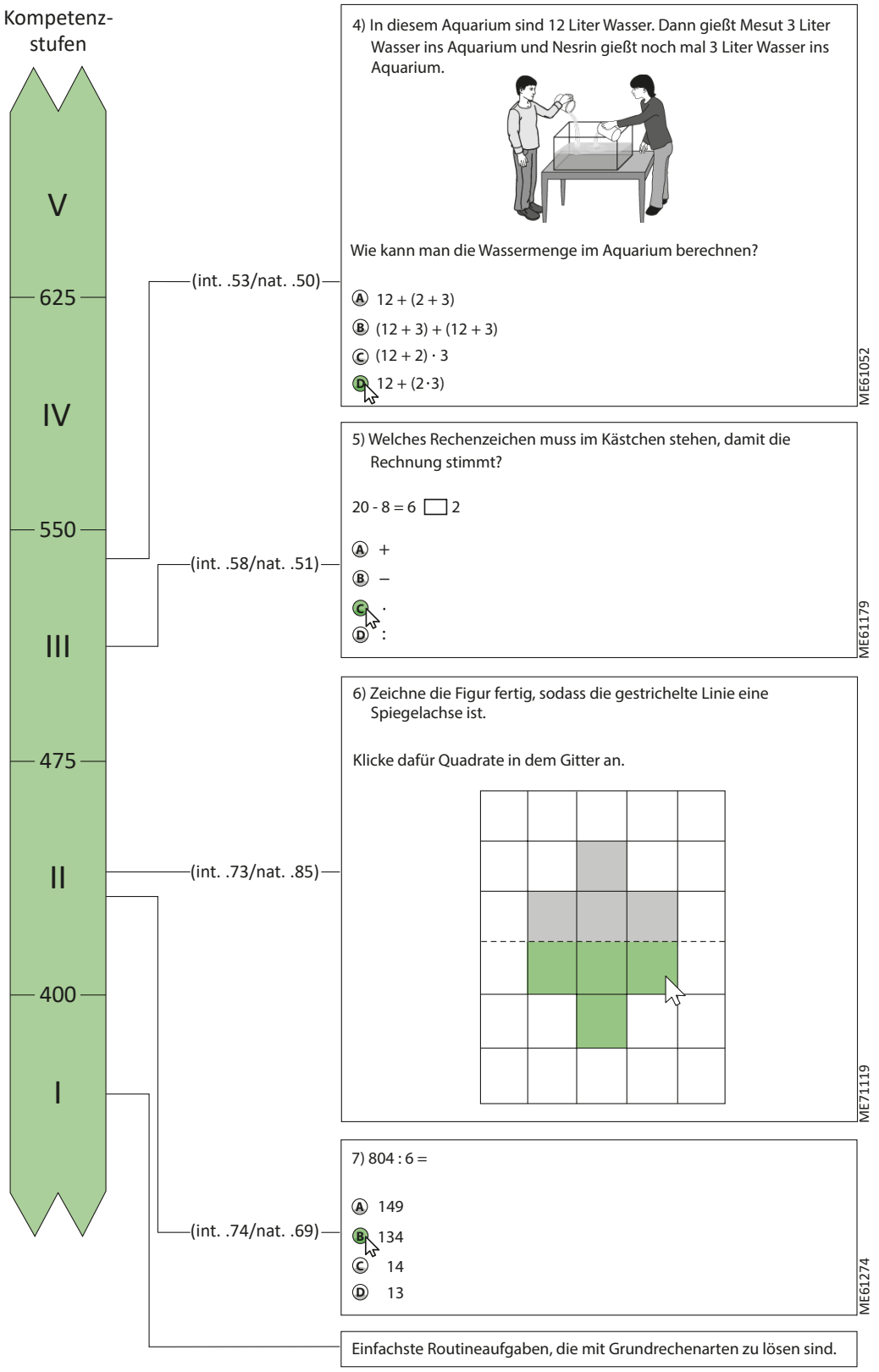


Da die erforderlichen Informationen zur Erstellung der Abbildung für TIMSS 2023 durch die internationale Studienleitung erst mit dem Datum der Veröffentlichung des Berichts bereitgestellt werden, wird zur Illustration der Aufgaben auf den verschiedenen Kompetenzstufen die Abbildung aus TIMSS 2019 (Selter et al., 2020) genutzt.

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

Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: https://www.iea.nl/publications/form/iea-permission-request-form

Abbildung 3.4: Kompetenzstufen und Beispielaufgaben (Gesamtskala Mathematik) II



Da die erforderlichen Informationen zur Erstellung der Abbildung für TIMSS 2023 durch die internationale Studienleitung erst mit dem Datum der Veröffentlichung des Berichts bereitgestellt werden, wird zur Illustration der Aufgaben auf den verschiedenen Kompetenzstufen die Abbildung aus TIMSS 2019 (Selter et al., 2020) genutzt.

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

Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: https://www.iea.nl/publications/form/iea-permission-request-form

Im Rahmen der dritten Aufgabe sollen die Schüler:innen die Anzahl der für die Erstellung vorgegebener dreidimensionaler Figuren benötigten quadratischen und dreieckigen Plättchen bestimmen. Dazu ist es erforderlich, dass die Schüler:innen die geometrischen Eigenschaften der Figuren nutzen und über Kompetenzen zum räumlichen Vorstellungsvermögen verfügen.

Kompetenzstufe III: Im Rahmen der vierten Aufgabe sind die Kinder gefordert, zu einer vorgegebenen einfachen Sachsituation den passenden Term zu bestimmen. Dabei ist insbesondere die Fähigkeit zur Übersetzung einer schriftlich-symbolischen und ikonischen Darstellung in eine mathematisch-symbolische Darstellung von zentraler Bedeutung.

In der fünften Aufgabe ist die Bestimmung eines fehlenden Symbols in einer Gleichung erforderlich, sodass eine Parität der zu vergleichenden Terme hergestellt ist. Dies setzt voraus, dass insbesondere ein Operationsverständnis zu den vier Grundrechenarten entwickelt wurde.

Kompetenzstufe II: In der sechsten Aufgabe sollen die Schüler:innen eine achsensymmetrische Ergänzung vornehmen. Dazu ist es erforderlich, dass die Kinder über ein grundlegendes Verständnis zur Achsensymmetrie verfügen und es auf eine einfache geometrische Form anwenden können.

Die siebte Aufgabe erfordert von den Kindern die Ermittlung des Ergebnisses einer Divisionsaufgabe, welche einen dreistelligen Dividenden und einen einstelligen Divisor aufweist. Dies bedingt das Abrufen und die Ausführung einer geeigneten Strategie oder des schriftlichen Algorithmus. Eine Möglichkeit stellt auch die Anwendung der multiplikativen Umkehroperation dar.

Kompetenzstufe I: Die Schüler:innen, die dieser Kompetenzstufe zuzuordnen sind, zeigen im TIMSS-Test lediglich basale Kompetenzen. Es kann angenommen werden, dass Schüler:innen, die dieser Kompetenzstufe zugeordnet sind, lediglich sehr einfache Routineaufgaben mit Grundrechenarten ausführen können.

3.5 Ergebnisse

Im vorliegenden Abschnitt werden die Ergebnisse der TIMS-Studie für das Fach Mathematik aus dem Jahr 2023 präsentiert und einer vergleichenden Analyse unterzogen. Als Referenzpunkte werden dazu die Ergebnisse der ersten TIMS-Studie aus dem Jahr 2007 und des vorangegangenen Studienzyklus aus dem Jahr 2019 herangezogen.

Dieser Abschnitt ist entlang der formulierten Forschungsfragen strukturiert, sodass der Abschnitt 3.5.1 mathematische Kompetenzen der Viertklässler:innen in Deutschland im internationalen Vergleich einordnet. Abschnitt 3.5.2 widmet sich der Darstellung der Verteilung der Lernenden auf die fünf Kompetenzstufen, wobei insbesondere auch auf die Anteile leistungsstarker und leistungsschwacher Schüler:innen eingegangen wird. In den Abschnitten 3.5.3 und 3.5.4 werden die Testergebnisse hinsichtlich der drei Inhaltsbereiche und der drei kognitiven Anforderungsbereiche präsentiert. Im Abschnitt 3.5.5 werden schließlich die nationalen Ergebnisse zur Einstellung von Viertklässler:innen zur Mathematik sowie zu deren mathematischem Selbstkonzept erläutert. Differenzierte Ergebnisse zu den mathematischen Kompetenzen vor dem Hintergrund der Variab-

len *Geschlecht*, *soziale Disparitäten* sowie *Migrationshintergrund* lassen sich darüber hinaus dem Kapitel 9 entnehmen.

Insgesamt nahmen 58 Staaten sowie fünf sogenannte Benchmark-Teilnehmer, also einzelne Regionen wie beispielsweise Abu Dhabi in den Vereinigten Arabischen Emiraten, an TIMSS 2023 teil. Darunter befanden sich 22 EU-Staaten und 29 OECD-Staaten (siehe Kapitel 2 in diesem Band). Verglichen mit dem TIMSS-Zyklus von 2019, an dem 64 teilnehmende Staaten beteiligt waren, ist die Teilnehmerzahl damit nahezu konstant, während bei TIMSS 2007 mit 43 Staaten noch eine geringere Beteiligung zu verzeichnen war.

Um die Testergebnisse der Schüler:innen in Deutschland international einzuordnen, werden im weiteren Verlauf dieses Abschnitts Verweise auf *Vergleichsgruppen* (VG) hergestellt. Die Testergebnisse werden verglichen mit den Mittelwerten der

- Teilnehmer, die Mitglieder der EU sind (VG_{EU}),
- Teilnehmer, die der OECD angehören (VG_{OECD}),
- Teilnehmer, die auf der Gesamtskala Mathematik signifikant bessere Ergebnisse erzielt haben, sowie der
- Teilnehmer, deren Ergebnisse sich auf der Gesamtskala Mathematik nicht statistisch signifikant von Deutschland unterscheiden.

Eine umfassende Übersicht aller Ergebnisse von TIMSS 2023 ist im Ergebnisbericht der internationalen Studienleitung (von Davier et al., 2024) verfügbar.

Abweichend von den Abbildungen, die die Testergebnisse der teilnehmenden Staaten und Regionen aus dem Jahr 2023 zeigen, wird bei allen Darstellungen, die die Unterschiede zwischen den verschiedenen TIMSS-Zyklen thematisieren, auf die Angabe des internationalen Mittelwertes sowie des Mittelwertes der VG_{EU} und der VG_{OECD} verzichtet. Grund dafür ist, dass sich die Zusammensetzung der Teilnehmer in den verschiedenen Vergleichsgruppen zwischen den TIMSS-Zyklen – wenn auch nur marginal – unterscheidet, sodass Interpretationen vor allem bezogen auf Veränderungen der Testergebnisse unzulässig wären.

3.5.1 Kompetenzen im internationalen Vergleich

In diesem Abschnitt wird auf die Testergebnisse der Viertklässler:innen in Deutschland im internationalen Vergleich eingegangen. Hierbei werden zunächst die Mittelwerte sowie die Streuungen der Kompetenzen beschrieben. Bei der Präsentation der Testergebnisse in der internationalen Rangreihe wird eine numerisch-formale Zusammenstellung der Testergebnisse gegeben. Differenziertere Analysen inhaltlicher Aspekte, wie dem Abschneiden in Bezug auf die Inhaltsbereiche und die kognitiven Anforderungsbereiche, folgen in den weiteren Abschnitten.

Leistungsmittelwerte der Staaten und Regionen: Abbildung 3.5 enthält für alle an TIMSS 2023 teilnehmenden Staaten und Regionen die festgestellten Leistungsmittelwerte (*M*) sowie die dazugehörigen Standardabweichungen (*SD*) und Standardfehler (*SE*). Des Weiteren sind in der Abbildung der Mittelwert aller teilnehmenden Staaten (internationaler Mittelwert) sowie die Mittelwerte der VG_{OECD} und VG_{EU} jeweils an der entsprechenden Position in der Rangreihe vermerkt.

Die Fußnoten links von den Teilnehmern weisen auf spezifische Merkmale der jeweiligen Untersuchungspopulationen und Stichproben hin. Die verschiedenen Bedeutungen der unterschiedlichen Fußnoten werden in Kapitel 2 sowie im Anhang dieses Bandes ausführlich erläutert.

Um die internationale Vergleichbarkeit der Testergebnisse sicherzustellen, hat die internationale Studienleitung strenge methodische Standards festgelegt. Werden diese Standards nicht eingehalten, sind die erzielten Ergebnisse des jeweiligen teilnehmenden Staates nur eingeschränkt vergleichbar. In TIMSS 2023 betrifft dies 37 der 58 Staaten. In den folgenden Abbildungen sind diese Staaten sowie die Benchmark-Teilnehmer jeweils kursiv dargestellt. Die übrigen 21 Staaten, zu denen auch Deutschland zählt, erfüllen die international festgelegten Kriterien, sodass die internationale Vergleichbarkeit dieser Testergebnisse vollständig gewährleistet ist.

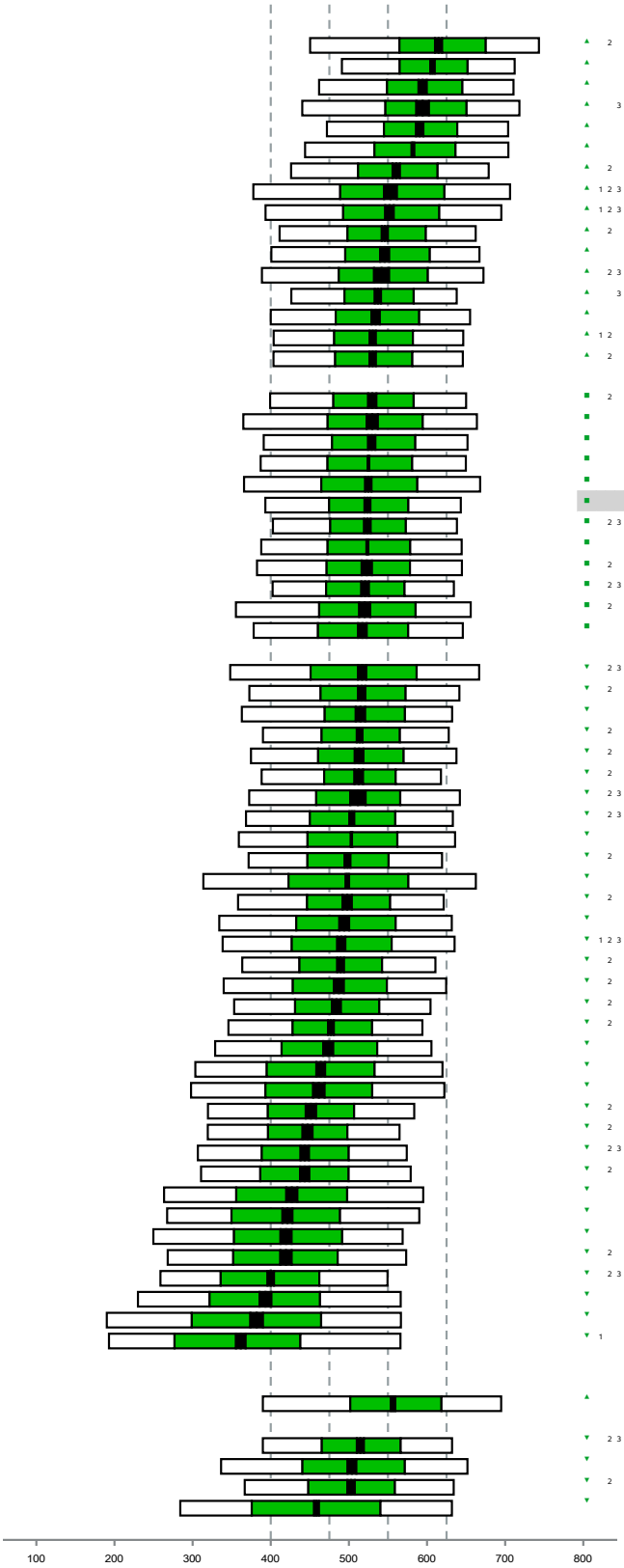
In Abbildung 3.5 wurden diejenigen Teilnehmer mit einem grünen Quadrat gekennzeichnet, deren Mittelwerte sich ggf. geringfügig, jedoch – auf einem Signifikanzniveau von 5 Prozent – nicht signifikant vom Leistungsmittelwert für Deutschland unterscheiden. Diese Gruppe wird im weiteren Verlauf dieses Abschnitts als *deutsche Leistungsgruppe* bezeichnet. Signifikant bessere oder schlechtere Werte wurden durch ein nach oben bzw. unten weisendes grünes Dreieck deutlich gemacht.

Viertklässler:innen in Deutschland erreichen in Mathematik einen durchschnittlichen Leistungswert von 524 Punkten, womit sie im internationalen Vergleich im oberen Mittelfeld liegen. Dieser Wert liegt über dem Skalenmittelwert von 500 und signifikant über dem internationalen Durchschnittswert von 503, der in TIMSS 2023 ermittelt wurde. Die Leistungswerte aller teilnehmenden Staaten und Regionen zeigen ein Spektrum von 362 bis 615 Punkten. Zu den Teilnehmern mit den höchsten Ergebnissen zählen Singapur (615), Taiwan (607) und Südkorea (594). Im Gegensatz dazu erzielten die Schüler:innen aus Marokko (393), Kuwait (382) und Südafrika (362) im Durchschnitt die niedrigsten Werte.

Signifikante Leistungsunterschiede: Die Darbietung der Ergebnisse in einer nach Leistungsmittelwerten geordneten Übersicht gibt zwar einen ersten Überblick über alle Teilnehmer hinweg, erlaubt jedoch keine differenzierten Aussagen über die mathematischen Kompetenzen der Schüler:innen in den teilnehmenden Staaten und Regionen. Es besteht insbesondere die Gefahr, dass kleine numerische Unterschiede in den Testergebnissen fälschlicherweise als bedeutende Unterschiede in den mathematischen Kompetenzen der Teilnehmer wahrgenommen werden, obwohl diese Unterschiede statistisch nicht signifikant sind.

Die *deutsche Leistungsgruppe* setzt sich aus Schweden (530), Bulgarien (530), Finnland (529), Australien (525), Dänemark (524), Serbien (523), der Flämischen Gemeinschaft in Belgien (521), Ungarn (520) und Portugal (517) zusammen. Auch zum Mittelwert der an der Studie teilnehmenden EU-Staaten (VG_{EU}, 524) sowie der OECD-Staaten (VG_{OECD}, 525) unterscheidet sich der Mittelwert für Deutschland nicht signifikant. Signifikant höhere Leistungsmittelwerte lassen sich für 16 teilnehmende Staaten bzw. Regionen feststellen, zu denen neben den bereits erwähnten führenden drei Nationen auch Hongkong (594), Japan (591), Macau (582), Litauen (561), die Türkei (553), England (552), Polen (546), Irland (546), Rumänien (542), die Niederlande (537), Lettland (534), Norwegen (531) und Tschechien (530) zählen. Im Gegensatz dazu wiesen 32 Staaten bzw. Regionen signifikant niedrigere Leistungsmittelwerte auf. Hierzu zählen u. a. die

Abbildung 3.5: Testleistung der Schüler:innen im internationalen Vergleich – Gesamtskala Mathematik



		Teilnehmer	*M*	*(SE)*	*SD*	*(SE)*
▲	2	*Singapur*	615	(2.3)	88	(1.4)
▲		Taiwan	607	(1.4)	67	(0.7)
▲		Republik Korea (Südkorea)	594	(2.6)	75	(2.1)
▲	3	*Hongkong*	594	(4.0)	84	(2.2)
▲		Japan	591	(2.3)	71	(1.1)
▲		Macau	582	(0.9)	79	(0.9)
▲	2	*Litauen*	561	(2.4)	76	(1.4)
▲	1 2 3	*Türkei (5. Jgst.)*	553	(4.1)	99	(2.0)
▲	1 2 3	*England*	552	(2.7)	92	(1.9)
▲	2	*Polen*	546	(2.0)	76	(1.1)
▲		Irland	546	(2.9)	82	(1.6)
▲	2 3	*Rumänien*	542	(4.8)	85	(2.2)
▲	3	*Niederlande*	537	(2.0)	65	(1.3)
▲		Lettland	534	(2.6)	77	(1.3)
▲	1 2	*Norwegen (5. Jgst.)*	531	(2.0)	74	(0.9)
▲	2	*Tschechien*	530	(2.0)	74	(1.0)
■	2	*Schweden*	530	(2.6)	76	(1.6)
■		Bulgarien	530	(3.6)	91	(3.1)
■		Finnland	529	(2.3)	80	(1.2)
■		**VG OECD**	**525**	**(0.5)**	**80**	**(0.3)**
■		Australien	525	(2.0)	91	(1.3)
■		Deutschland	524	(1.9)	75	(1.4)
■	2 3	*Dänemark*	524	(2.1)	72	(1.1)
■		**VG EU**	**524**	**(0.6)**	**78**	**(0.3)**
■	2	*Serbien*	523	(3.3)	80	(1.9)
■	2 3	*Belgien (Fläm. Gem.)*	521	(2.4)	71	(1.4)
■	2	*Ungarn*	520	(3.5)	91	(2.5)
■		Portugal	517	(2.8)	82	(1.4)
▼	2 3	*USA*	517	(2.6)	97	(0.9)
▼	2	*Zypern*	516	(2.4)	81	(1.3)
▼		Slowakei	515	(3.0)	81	(2.1)
▼	2	*Slowenien*	514	(1.8)	73	(0.9)
▼	2	*Italien*	513	(2.8)	80	(1.3)
▼	2	*Armenien*	513	(2.7)	70	(1.4)
▼	2 3	*Albanien*	512	(4.9)	81	(3.0)
▼	2 3	*Kanada*	504	(1.5)	81	(0.9)
▼		**Internationaler Mittelwert**	**503**	**(0.4)**	**84**	**(0.2)**
▼	2	*Spanien*	498	(1.8)	75	(1.2)
▼		Vereinigte Arabische Emirate (VAE)	498	(1.1)	107	(0.7)
▼	2	*Georgien*	498	(2.9)	79	(1.6)
▼		Aserbaidschan	494	(3.1)	91	(1.6)
▼	1 2 3	*Neuseeland*	490	(2.6)	91	(1.2)
▼	2	*Belgien (Frz. Gem.)*	489	(2.2)	76	(1.2)
▼	2	*Kasachstan*	487	(3.2)	87	(1.7)
▼	2	*Frankreich*	484	(2.8)	77	(1.2)
▼	2	*Montenegro*	477	(1.8)	76	(1.1)
▼		Nordmazedonien	474	(3.2)	85	(1.3)
▼		Katar	464	(2.8)	97	(1.7)
▼		Bahrain	462	(3.7)	99	(2.2)
▼	2	*Kosovo*	451	(3.3)	80	(2.0)
▼	2	*Bosnien und Herzegowina*	447	(3.2)	75	(1.6)
▼	2 3	*Chile*	444	(2.7)	80	(1.3)
▼	2	*Usbekistan*	443	(3.0)	82	(1.3)
▼		Jordanien	427	(3.5)	101	(2.4)
▼		Oman	421	(3.2)	99	(1.3)
▼	4	*Iran*	420	(3.6)	98	(1.9)
▼	2	*Saudi-Arabien*	420	(3.6)	94	(1.7)
▼	2 3 4	*Brasilien*	400	(2.1)	90	(1.2)
▼		Marokko	393	(3.8)	102	(2.0)
▼		Kuwait	382	(4.0)	115	(2.4)
▼	1	*Südafrika (5. Jgst.)*	362	(3.1)	114	(2.3)
		Benchmark-Teilnehmer				
▲		Dubai, VAE	557	(1.4)	91	(0.9)
▼	2 3	*Québec, Kanada*	515	(2.2)	74	(1.1)
▼		Schardscha, VAE	504	(2.8)	96	(1.7)
▼	2	*Ontario, Kanada*	503	(2.4)	82	(1.3)
▼		Abu Dhabi, VAE	459	(1.5)	108	(0.9)

▲ = Mittelwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Mittelwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu Deutschland nicht statistisch signifikant ($p > .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten

© TIMSS 2023

USA (517), die Slowakei (515), Slowenien (514), Italien (513), Kanada (504), Spanien (498), Neuseeland (490), Frankreich (484), Chile (444), der Iran (420), Brasilien (400) oder Südafrika (362).

Leistungsstreuung: Die bisher dargestellten Ergebnisse dienen der Beschreibung der Leistungsmittelwerte der teilnehmenden Staaten und Regionen sowie der Identifizierung von Unterschieden zwischen diesen Werten. Von besonderem Interesse ist zudem die Analyse der Verteilung mathematischer Leistungswerte innerhalb eines Staates bzw. einer Region. In Abbildung 3.5 erfolgt eine Darstellung der Leistungsstreuung der teilnehmenden Staaten mittels sogenannter Perzentilbänder. Dabei werden jeweils das 5., 25., 75. und 95. Perzentil angegeben (siehe Kapitel 2 in diesem Band). Der für das 25. Perzentil angegebene Leistungswert besagt demnach, dass 25 Prozent der Viertklässler:innen einen gleich großen oder kleineren Leistungswert erzielt haben. Demgegenüber lässt sich ableiten, dass 75 Prozent einen höheren Leistungswert erreicht haben.

Die dargestellten Perzentilbänder verdeutlichen, dass die Testergebnisse Deutschlands im internationalen Vergleich eine vergleichsweise homogene Verteilung aufweisen. Im Vergleich zu vielen anderen teilnehmenden Staaten und Regionen ist die Leistungsstreuung in Deutschland relativ gering, was der Breite der Perzentilbänder zu entnehmen ist, die den Unterschied der Leistungsmittelwerte zwischen den Perzentilen 5 und 95 beschreibt. Bezogen auf Deutschland beläuft sich dieser – in der Abbildung nicht numerisch abgebildete – Unterschied auf 250 Punkte. In nur wenigen Staaten und Regionen, etwa für Taiwan (221), die Niederlande (212) sowie Dänemark (235), ist eine geringere Leistungsstreuung zu verzeichnen.

Die Streuung der Leistungen ist zudem auch in der Standardabweichung (*SD*) vom jeweiligen Leistungsmittelwert ersichtlich. Für Deutschland ist diese in Abbildung 3.5 mit 75 Punkten vermerkt. Das bedeutet, dass ca. zwei Drittel der an TIMSS teilnehmenden Viertklässler:innen in Deutschland Leistungen erzielen, die zwischen 449 (524 – 75) und 599 (524 + 75) Punkten zu verorten sind. Ein Vergleich Deutschlands mit anderen europäischen sowie internationalen Staaten und Benchmark-Teilnehmern zeigt, dass die Testergebnisse der Viertklässler:innen in Deutschland – wie auch schon in TIMSS 2007 (68) und TIMSS 2019 (70) – durch eine eher geringe Standardabweichung gekennzeichnet sind.

Konkret bedeutet dies beispielsweise für einen Vergleich der Testergebnisse mit denen Rumäniens, dass Rumänien zwar einen signifikant höheren Mittelwert von 542 als Deutschland (524) erzielt hat, die Leistungswerte der Viertklässler:innen in Rumänien mit einer Standardabweichung von 85 jedoch eine weitaus stärkere Heterogenität aufweisen. Im Gegensatz dazu erreichen die Niederlande einen höheren Leistungsmittelwert (537) als Deutschland bei einer geringeren Leistungsstreuung (65).

Lediglich für zehn teilnehmende Staaten und Regionen liegen geringere Standardabweichungen vor. Höhere Standardabweichungen sind demgegenüber für 52 Staaten und Regionen zu verzeichnen. Die höchsten Leistungsstreuungen europäischer Staaten, gemessen in Standardabweichungen, liegen für die Türkei (99), England (92), Bulgarien (91) und Ungarn (91) vor.

Vergleich der Testleistungen 2007, 2019 und 2023 auf der Gesamtskala: Wie einleitend beschrieben, werden die Testleistungen aus TIMSS 2023 für die Trendvergleiche insbesondere mit der ersten TIMSS-Erhebungsrunde aus dem Jahr 2007 sowie dem vorangegangenen Studienzyklus aus dem Jahr 2019 verglichen. Neben Deutschland haben insgesamt 18 Staaten und drei Benchmark-Regionen an diesen Erhebungsrunden teilgenommen. Die zugehörigen Leistungsmittelwerte sind für die drei betrachteten Erhebungsrunden in Abbildung 3.6 zusammengestellt.

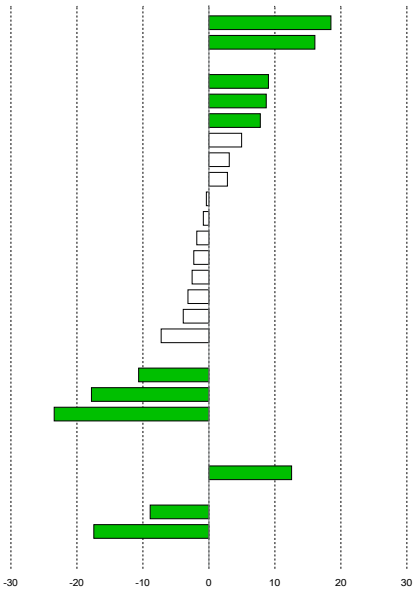
Wie in Abbildung 3.5 markieren grüne Dreiecke bzw. Quadrate das Vorliegen statistisch (nicht) signifikanter Abweichungen von den in Deutschland erzielten Differenzwerten, während Plus- bzw. Minuszeichen auf statistisch signifikant höhere bzw. niedrigere Differenzwerte im Vergleich der Ergebnisse der verglichenen Erhebungszeitpunkte hindeuten.

Die im rechten Teil der Abbildung 3.6 dargestellten, grün gefärbten Balken deuten auf signifikante Leistungsunterschiede zwischen den Studienzyklen von 2019 und 2023 hin. Sind diese hingegen weiß, so ist der jeweilige Unterschied statistisch nicht signifikant. Der Vergleich des aktuellen mit dem vorherigen Studienzyklus 2019 ist insbesondere vor dem Hintergrund der Auswirkungen der pandemiebedingten Schulschließungen zwischen 2020 und 2022 von besonderem Interesse. Es zeigt sich, dass Deutschland einer von elf Teilnehmern ist, für die keine signifikanten Leistungsunterschiede zwischen 2019 und 2023 zu verzeichnen sind. Fünf Staaten erzielten in 2023 signifikant bessere Leistungsmittelwerte als noch in 2019. Zu diesen Staaten zählen Litauen (+18), Georgien (+16), Australien (+9), Schweden (+9) sowie Taiwan (+8). Signifikant niedrigere Leistungsmittelwerte sind hingegen für die folgenden drei Staaten zu verzeichnen: Singapur (-11), USA (-18) sowie Iran (-23).

Der Vergleich der Erhebungsrunden von 2007 und 2023 zeigt, dass Deutschland zu den fünf teilnehmenden Staaten gehört, für die keine signifikanten Leistungsunterschiede zu verzeichnen sind. Schüler:innen aus zwölf Staaten schnitten in TIMSS 2023 signifikant besser ab als 2007. Die größten Zuwächse sind für Georgien (+59), Tschechien (+44) sowie Litauen und Taiwan (jeweils +31) festzustellen. Bei zwei Teilnehmern sind in TIMSS 2023 signifikant schwächere Leistungen zu verzeichnen als noch in 2007: Hongkong (-12) und USA (-12).

Abbildung 3.6: Vergleich der Testleistungen zwischen TIMSS 2007, TIMSS 2019 und TIMSS 2023 – Gesamtskala Mathematik

Fußnoten 2007	2019	2023	T	Teilnehmer	2007 M_{07}	(SE)	2019 M_{19}	(SE)	2023 M_{23}	(SE)	Veränderungen[A] M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		
2	2	2	8	Litauen	530	(2.4)	542	(2.8)	561	(2.4)	31	(3.3)	+	18	(3.6)	+	▲
2	2	2		Georgien	438	(4.2)	482	(3.7)	498	(2.9)	59	(5.1)	+	16	(4.7)	+	▲
				Australien	516	(3.5)	516	(2.8)	525	(2.0)	9	(4.0)	+	9	(3.4)	+	■
	2	2		Schweden	503	(2.5)	521	(2.8)	530	(2.6)	27	(3.6)	+	9	(3.8)	+	■
				Taiwan	576	(1.7)	599	(1.9)	607	(1.4)	31	(2.2)	+	8	(2.4)	+	■
	2			Slowakei	496	(4.5)	510	(3.5)	515	(3.0)	19	(5.4)	+	5	(4.6)		■
1 2	1 2	1 2 3	8	Neuseeland	492	(2.3)	487	(2.6)	490	(2.6)	-2	(3.5)		3	(3.7)		■
				Deutschland	525	(2.3)	521	(2.3)	524	(1.9)	-1	(2.9)		3	(2.9)		■
3	3	3	8	Niederlande	535	(2.1)	538	(2.2)	537	(2.0)	2	(2.9)		0	(2.9)		■
3	3	2 3	8	Dänemark	523	(2.4)	525	(1.9)	524	(2.1)	1	(3.2)		-1	(2.8)		■
2		2	8	Italien	507	(3.1)	515	(2.4)	513	(2.8)	6	(4.2)		-2	(3.7)		■
	3		8	Japan	568	(2.1)	593	(1.8)	591	(2.3)	23	(3.1)	+	-2	(2.9)		■
		2		Tschechien	486	(2.8)	533	(2.5)	530	(2.0)	44	(3.5)	+	-3	(3.2)		■
		2		Ungarn	510	(3.5)	523	(2.6)	520	(3.5)	11	(5.0)	+	-3	(4.4)		■
1 3	1 2	1 2 3	8	England	541	(2.9)	556	(3.0)	552	(2.7)	10	(4.0)	+	-4	(4.0)		■
2 3	3	3	8	Hongkong	607	(3.6)	602	(3.3)	594	(4.0)	-12	(5.4)	–	-7	(5.2)		■
	2	2	8	Singapur	599	(3.7)	625	(3.9)	615	(2.3)	15	(4.4)	+	-11	(4.5)	–	▼
2 3	2 3	2 3		USA	529	(2.4)	535	(2.5)	517	(2.6)	-12	(3.6)	–	-18	(3.7)	–	▼
		4	8	Iran	402	(4.1)	443	(3.9)	420	(3.6)	17	(5.4)	+	-23	(5.2)	–	▼
				Benchmark-Teilnehmer													
2 3 6	2		8	Dubai, VAE	444	(2.1)	544	(1.6)	557	(1.4)	112	(2.5)	+	13	(2.1)	+	▲
2		2	8	Ontario, Kanada	512	(3.1)	512	(3.3)	503	(2.4)	-9	(3.9)	–	-9	(4.1)	–	▼
2	3	2 3	8	Québec, Kanada	519	(3.0)	532	(2.3)	515	(2.2)	-4	(3.8)		-17	(3.2)	–	▼



□ Keine statistisch signifikanten Veränderungen zwischen 2019 und 2023 ($p > .05$)
■ Statistisch signifikante Veränderungen zwischen 2019 und 2023 ($p \leq .05$)
▲ = Differenzwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
\+ = Differenzwert in 2023 statistisch signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Differenzwert in 2023 statistisch signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorga
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

Extremgruppen: Die Tabelle 3.11 zeigt die Mathematikleistungen aufgeschlüsselt nach Perzentilen, um insbesondere die Anteile an leistungsstarken und leistungsschwachen Schüler:innen zu analysieren. Zur Darstellung der Veränderungen der Leistungswerte zwischen den Studienzyklen 2007, 2019 und 2023 werden erneut ausschließlich diejenigen Staaten und Regionen berücksichtigt, die in allen drei Zyklen vertreten waren. Die Anordnung der Daten erfolgt gemäß den in Abbildung 3.5 angegebenen Durchschnittswerten auf der Gesamtskala Mathematik.

In einem ersten Schritt erfolgt die Betrachtung im Hinblick auf die leistungsschwächsten 5 Prozent Viertklässler:innen, bevor eine Analyse in Bezug auf die leistungsstärksten 5 Prozent Viertklässler:innen durchgeführt wird.

Die Tabelle 3.11 verdeutlicht, dass die leistungsschwächsten 5 Prozent der Schüler:innen in Deutschland in TIMSS 2023 maximal 393 Punkte erreichten. Die durch ein grünes Quadrat markierten Werte der teilnehmenden Staaten signalisieren erneut keine statistisch signifikanten Abweichungen zu den Ergebnissen Deutschlands. Ein nach oben oder unten gerichtetes Dreieck hingegen kennzeichnet statistisch signifikante Unterschiede. So weichen die zum 5. Perzentil gehörenden Werte von Tschechien (403), Dänemark (403), Schweden (399) und England (393) nicht signifikant von den Ergebnissen Deutschlands ab.

Demgegenüber erzielten die leistungsschwächsten Viertklässler:innen aus sechs Staaten – Taiwan (491), Japan (472), Singapur (450), Hongkong (440), Litauen (426) und den Niederlanden (426) – signifikant höhere Punktwerte im Vergleich zu Deutschland. Umgekehrt lassen sich in acht Staaten – Italien (375), Australien (366), der Slowakei (363), Ungarn (355), den USA (348), Georgien (358), Neuseeland (338) und Iran (250) – signifikant niedrigere Leistungswerte im 5. Perzentil feststellen.

Die leistungsstärksten 5 Prozent der Schüler:innen in Deutschland erzielten Punktwerte von mindestens 643. Diese Ergebnisse unterscheiden sich nicht signifikant von den Werten für Schweden (650), Tschechien (646), die Niederlande (638), Dänemark (638), Italien (638) sowie Neuseeland (635). Dagegen erreichten die leistungsstärksten 5 Prozent der Schüler:innen in neun Staaten – darunter nicht nur die führenden asiatischen Staaten, sondern auch England (695), Litauen (679), Australien (668), die USA (667) und Ungarn (656) – signifikant höhere Werte. Signifikant niedrigere Punktwerte wurden hingegen für die Slowakei (632), Georgien (621) und den Iran (569) festgestellt.

Die Tabelle 3.11 zeigt nicht nur die Mathematikleistungen des 5. und 95. Perzentils, sondern auch die des 25. und 75. Perzentils. Veränderungen der Mathematikleistungen in den Perzentilen können durch die Analyse der ebenfalls notierten Werte aus TIMSS 2007 und 2019 nachgezeichnet werden. Statistisch signifikante Veränderungen sind durch ein Pluszeichen für signifikante Leistungssteigerungen und ein Minuszeichen für signifikante Leistungsrückgänge gekennzeichnet.

Tabelle 3.11: Mathematikleistungen nach Perzentilen im Vergleich: TIMSS 2007, TIMSS 2019 und TIMSS 2023 – Gesamtskala Mathematik

Fußnoten 2007	2019	2023	T	Teilnehmer	TIMSS 2007 Perzentile										TIMSS 2019 Perzentile										TIMSS 2023 Perzentile									
					5	(SE)		25	(SE)		75	(SE)		95	(SE)		5	(SE)		25	(SE)		75	(SE)		95	(SE)		5	(SE)		25	(SE)		75	(SE)		95	(SE)	
	2		8	Singapur	447	(6.5)		548	(5.1)	+	659	(4.0)	+	725	(4.1)	+	481	(8.0)	–	578	(4.9)	–	682	(3.8)		741	(3.8)		450	(6.1)	▲	565	(3.9)	▲	675	(2.4)	▲	743	(3.7)	▲
				Taiwan	457	(4.1)	+	532	(2.1)	+	623	(2.2)	+	686	(2.1)	+	483	(5.6)		557	(2.7)	+	645	(3.2)	+	701	(2.6)	+	491	(2.9)	▲	565	(1.7)	▲	652	(1.5)	▲	712	(3.0)	▲
2 3	3	3	8	Hongkong	493	(9.1)	–	564	(4.2)	–	653	(4.0)		712	(5.3)		479	(5.2)	–	558	(4.1)		649	(3.7)		710	(4.3)		440	(9.0)	▲	547	(5.2)	▲	651	(4.8)	▲	718	(7.3)	▲
	3		8	Japan	438	(2.6)	+	520	(2.1)	+	620	(2.1)	+	688	(3.8)	+	474	(3.7)		548	(2.5)		641	(2.3)		703	(3.1)		472	(4.9)	▲	545	(3.1)	▲	639	(3.3)	▲	704	(4.4)	▲
2	2	2	8	Litauen	396	(3.7)	+	482	(3.5)	+	583	(2.9)	+	645	(5.6)	+	410	(5.7)	+	493	(3.4)	+	596	(2.8)	+	660	(4.9)	+	426	(3.7)	▲	512	(3.6)	▲	614	(2.3)	▲	679	(3.6)	▲
1 3	1 2	1 2 3	8	England	392	(4.2)		487	(3.5)		600	(3.6)	+	676	(4.3)	+	411	(5.5)	–	499	(3.5)		615	(4.1)		693	(6.5)		393	(5.9)	■	493	(4.1)	▲	616	(2.4)	▲	695	(5.1)	▲
3	3	3	8	Niederlande	429	(7.0)		495	(2.9)		577	(2.8)		632	(2.3)		432	(4.3)		495	(2.7)		581	(2.5)		634	(3.7)		426	(5.2)	▲	494	(3.1)	▲	583	(2.1)	▲	638	(3.3)	■
		2		Tschechien	361	(6.6)	+	440	(4.9)	+	536	(2.9)	+	597	(2.9)	+	406	(5.8)		484	(3.5)		585	(2.8)		650	(3.9)		403	(3.9)	■	482	(2.6)	▲	581	(2.5)	■	646	(3.1)	■
	2	2		Schweden	388	(4.5)		459	(3.3)	+	548	(3.0)	+	608	(2.7)	+	396	(5.4)		474	(4.0)		571	(3.4)	+	640	(4.4)		399	(6.6)	■	480	(4.0)	■	583	(2.2)	▲	650	(4.3)	■
				Australien	373	(8.2)		463	(4.1)		573	(4.2)	+	647	(3.9)	+	364	(6.0)		460	(3.6)		575	(3.4)	+	654	(5.8)	+	366	(4.3)	▼	465	(3.0)	▼	588	(3.0)	▲	668	(3.0)	▲
				Deutschland	409	(10.3)		483	(2.6)	–	572	(2.2)		629	(2.6)	+	405	(4.5)		474	(2.9)		570	(2.8)		630	(3.2)	+	393	(5.0)	■	475	(2.8)	■	576	(1.9)	■	643	(3.6)	■
3	3	2 3	8	Dänemark	403	(9.9)		478	(3.9)		571	(2.4)		634	(4.8)		402	(5.5)		474	(2.7)		577	(2.4)		644	(4.2)		403	(4.1)	■	476	(2.7)	■	573	(2.2)	■	638	(3.1)	■
		2		Ungarn	347	(12.4)		452	(6.6)		574	(3.7)	+	647	(4.6)		388	(5.2)	–	472	(4.1)		578	(2.8)	+	645	(3.3)	+	355	(11.8)	▼	462	(5.8)	▼	585	(2.8)	▲	656	(4.2)	▲
2 3	2 3	2 3		USA	401	(3.8)	–	479	(2.3)	–	581	(3.0)		650	(5.2)	+	383	(5.9)	–	480	(3.2)	–	594	(3.0)		663	(4.0)		348	(3.8)	▼	451	(3.4)	▼	587	(2.7)	▲	667	(3.9)	▲
	2			Slowakei	350	(9.8)		446	(4.2)	+	553	(3.8)	+	623	(5.2)		368	(10.9)		466	(4.6)		561	(3.2)	+	626	(5.0)		363	(8.6)	▼	469	(4.3)	■	572	(2.9)	■	632	(3.8)	▼
2		2	8	Italien	374	(6.1)		457	(3.7)		558	(3.3)	+	629	(12.2)		403	(4.8)	–	470	(3.2)		560	(2.8)	+	619	(4.4)	+	375	(5.3)	▼	461	(3.8)	▼	570	(2.6)	■	638	(3.3)	■
2	2	2		Georgien	289	(6.3)	+	378	(6.8)	+	501	(5.8)	+	582	(6.4)	+	340	(6.2)	+	429	(5.3)	+	538	(3.9)	+	608	(4.4)		358	(5.7)	▼	446	(3.6)	▼	553	(3.1)	▼	621	(5.3)	▼
1 2	1 2	1 2 3	8	Neuseeland	341	(7.3)		436	(2.6)		553	(2.5)		626	(3.6)	+	338	(4.7)		425	(3.4)		549	(3.3)		634	(4.2)		338	(5.1)	▼	427	(3.9)	▼	555	(3.4)	▼	635	(2.7)	■
		4	8	Iran	260	(5.6)		346	(4.9)		461	(4.2)	+	534	(6.4)	+	278	(8.6)	–	379	(5.7)	–	512	(3.9)	–	592	(4.7)	–	250	(6.5)	▼	353	(5.0)	▼	491	(3.5)	▼	569	(3.7)	▼
				Benchmark-Teilnehmer																																		
2 3 6	2		8	Dubai, VAE	406	(2.8)	–	473	(4.1)	+	566	(4.5)	+	627	(4.5)	+	400	(4.5)		491	(2.7)	+	601	(2.6)	+	671	(4.2)	+	390	(3.9)	■	502	(2.4)	▲	618	(1.4)	▲	695	(3.4)	▲
2	3	2 3	8	Québec, Kanada	293	(9.5)	+	384	(3.6)	+	506	(2.3)	+	589	(6.2)	+	421	(5.3)	–	487	(3.8)	–	578	(3.2)	–	638	(3.6)		390	(4.1)	■	465	(3.0)	▼	566	(2.3)	▼	632	(2.8)	▼
2		2	8	Ontario, Kanada	395	(4.5)	–	468	(6.2)	–	558	(5.2)		621	(3.6)	+	383	(5.7)	–	458	(4.5)		567	(4.2)		637	(6.1)		367	(4.2)	▼	448	(3.4)	▼	559	(3.0)	▼	634	(4.0)	■

▲ = Punktwerte in den Perzentilen im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Punktwerte in den Perzentilen im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu Punktwerten in den Perzentilen im Vergleich zu Deutschland nicht statistisch signifikant ($p > .05$)
\+ = Mittelwert in 2023 statistisch signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 statistisch signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007

© TIMSS 2023

3.5.2 Kompetenzstufen

Das für diese Untersuchung genutzte TIMSS-Kompetenzstufenmodell, einschließlich der Merkmale der fünf Kompetenzstufen, wurde in Abschnitt 3.4.3 dieses Kapitels dargelegt. Nachfolgend werden die Ergebnisse von TIMSS 2023 hinsichtlich der Verteilung der Viertklässler:innen in Deutschland auf die Kompetenzstufen vorgestellt. Ferner werden die Ergebnisse international eingeordnet und im Trend betrachtet.

Verteilung der Schüler:innen auf die Kompetenzstufen: Die Verteilung der Viertklässler:innen der Teilnehmer in TIMSS 2023 auf die Kompetenzstufen ist in Abbildung 3.7 dargestellt. Die rechte Seite der Abbildung zeigt den prozentualen Anteil der Schüler:innen, die die jeweilige Kompetenzstufe erreicht haben. Die linke Seite gibt die kumulierten Anteile wieder. Dadurch lässt sich ablesen, wie viel Prozent der an TIMSS teilnehmenden Viertklässler:innen eine bestimmte Kompetenzstufe mindestens erreichen konnten. In Bezug auf Deutschland lässt sich festhalten, dass sich 94 Prozent der Kinder mindestens auf Kompetenzstufe II, 75 Prozent mindestens auf Kompetenzstufe III, 38 Prozent mindestens auf Kompetenzstufe IV und 8 Prozent auf der höchsten Kompetenzstufe V befinden.

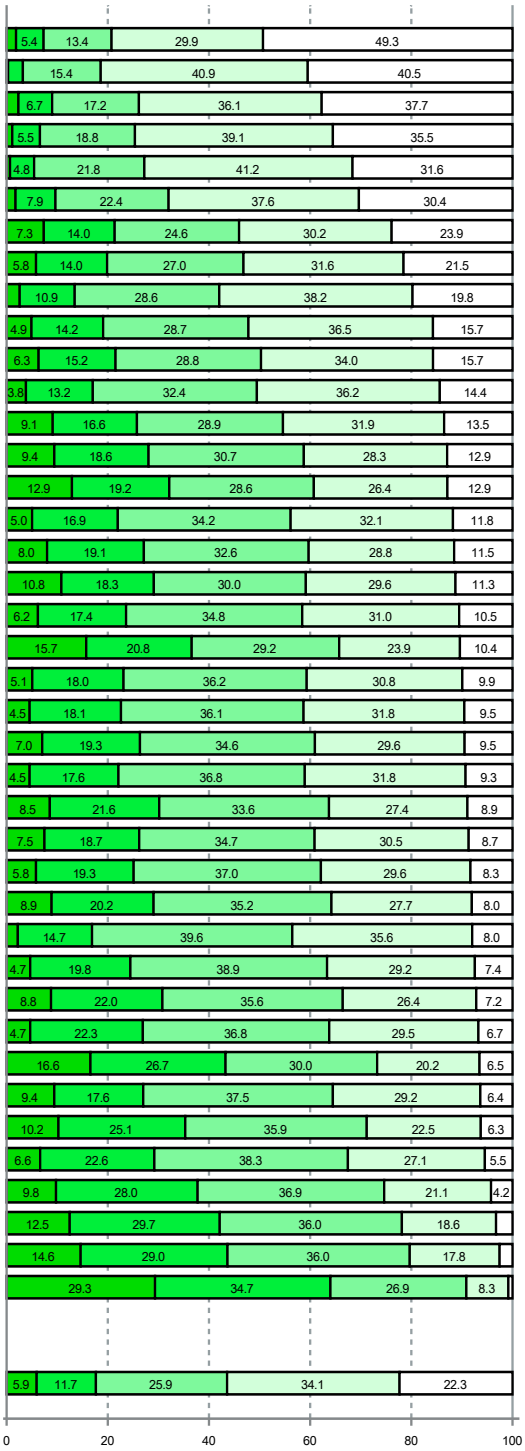
Im Vergleich zu den neun Staaten der *deutschen Leistungsgruppe* ist zu konstatieren, dass in Bulgarien (14 %), Australien (13 %), Ungarn (11 %), Finnland (11 %), Schweden (10 %), Portugal (9 %) und Serbien (9 %) anteilig mehr Schüler:innen der höchsten Kompetenzstufe zuzuordnen sind. Hingegen gilt dies für Dänemark (7 %) und die Flämische Gemeinschaft in Belgien (7 %) für anteilig weniger Schüler:innen. In den Vergleichsgruppen VG_{OECD} und VG_{EU} sind die Anteile mit 12 Prozent bzw. 10 Prozent im Vergleich zu Deutschland größer. Die größten Anteile von Viertklässler:innen auf der höchsten Kompetenzstufe V sind für Singapur (49 %), Taiwan (41 %) und Hongkong (38 %) zu verzeichnen. Dementsprechend ist im internationalen Vergleich ein deutlicher Unterschied zwischen den Staaten mit den höchsten Leistungsmittelwerten und Deutschland erkennbar. Die höchsten Werte in Europa erreichen England (22 %), Litauen (20 %), Irland (16 %) und Rumänien (16 %).

Der Anteil an Schüler:innen in Deutschland auf der untersten Kompetenzstufe I liegt bei 6 Prozent. Im Vergleich zu den Staaten der *deutschen Leistungsgruppe* zeigt sich, dass in Ungarn (11 %), Bulgarien (9 %), Australien (9 %), Portugal (9 %) und Serbien (8 %) anteilig mehr Schüler:innen der niedrigsten Kompetenzstufe zuzuordnen sind, während Finnland (6 %), Dänemark (5 %), Schweden (5 %) sowie die Flämische Gemeinschaft in Belgien (5 %) in etwa einen gleich großen Anteil von Lernenden auf dieser Kompetenzstufe aufweisen. Mit Ausnahme der Türkei (7 %) weisen alle Teilnehmer mit signifikant besseren Leistungsmittelwerten auf der Gesamtskala als Deutschland einen geringeren oder gleich großen Anteil auf. Die Werte der Vergleichsgruppen VG_{OECD} (8 %) und VG_{EU} (7 %) liegen oberhalb des Wertes für Deutschland.

Der Anteil der Viertklässler:innen in Deutschland mit mindestens durchschnittlichen mathematischen Kompetenzen – was in TIMSS das Erreichen von mindestens Kompetenzstufe III bedeutet – beträgt 75 Prozent. In einem Vergleich mit den Staaten der *deutschen Leistungsgruppe* zeigt sich, dass dieser Wert in sechs teilnehmenden Staaten etwas geringer ist. Dazu zählen Bulgarien (74 %), Serbien (74 %), die Flämische Gemeinschaft in Belgien (73 %), Australien (72 %), Ungarn (71 %) und Portugal (70 %). In drei teilnehmenden Staaten hin-

Abbildung 3.7: Prozentuale Verteilung der Schüler:innen auf die fünf Kompetenzstufen (Mathematik) im internationalen Vergleich

	Anteil (%) derer, die mindestens folgende Kompetenzstufe erreichen							
	II		III		IV		V	
Teilnehmer	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
2 *Singapur*	98.1	(0.2)	92.7	(0.5)	79.3	(0.9)	49.3	(1.1)
Taiwan	99.7	(0.1)	96.8	(0.2)	81.4	(0.7)	40.5	(1.0)
3 *Hongkong*	97.7	(0.4)	91.0	(0.9)	73.9	(1.7)	37.7	(2.3)
Republik Korea (Südkorea)	98.9	(0.4)	93.4	(0.8)	74.6	(1.1)	35.5	(1.3)
Japan	99.3	(0.2)	94.6	(0.6)	72.8	(1.3)	31.6	(1.4)
Macau	98.2	(0.2)	90.3	(0.4)	68.0	(0.5)	30.4	(0.6)
1 2 3 *Türkei (5. Jgst.)*	92.7	(0.8)	78.6	(1.5)	54.1	(1.9)	23.9	(1.4)
1 2 3 *England*	94.2	(0.7)	80.1	(1.0)	53.2	(1.2)	21.5	(1.0)
2 *Litauen*	97.4	(0.4)	86.5	(1.0)	58.0	(1.2)	19.8	(1.1)
Irland	95.1	(0.7)	80.9	(1.2)	52.2	(1.7)	15.7	(0.9)
2 3 *Rumänien*	93.7	(1.1)	78.5	(2.1)	49.7	(2.5)	15.7	(1.6)
2 *Polen*	96.2	(0.5)	83.0	(0.8)	50.5	(1.2)	14.4	(0.8)
Bulgarien	90.9	(1.3)	74.2	(1.7)	45.4	(1.3)	13.5	(0.9)
Australien	90.6	(0.6)	71.9	(0.9)	41.3	(1.1)	12.9	(0.7)
2 3 *USA*	87.1	(0.6)	67.8	(1.0)	39.3	(1.2)	12.9	(0.7)
Lettland	95.0	(0.6)	78.0	(1.4)	43.9	(1.5)	11.8	(0.7)
VG OECD	**92.0**	**(0.1)**	**72.9**	**(0.2)**	**40.3**	**(0.2)**	**11.5**	**(0.1)**
2 *Ungarn*	89.2	(1.3)	70.9	(1.5)	40.9	(1.5)	11.3	(0.8)
Finnland	93.8	(0.6)	76.4	(1.0)	41.6	(1.2)	10.5	(0.7)
Internationaler Mittelwert	**84.3**	**(0.1)**	**63.4**	**(0.2)**	**34.3**	**(0.2)**	**10.4**	**(0.1)**
2 *Schweden*	94.9	(0.8)	76.9	(1.5)	40.7	(1.4)	9.9	(0.8)
1 2 *Norwegen (5. Jgst.)*	95.5	(0.4)	77.4	(1.1)	41.3	(1.1)	9.5	(0.6)
VG EU	**93.0**	**(0.2)**	**73.6**	**(0.3)**	**39.1**	**(0.3)**	**9.5**	**(0.2)**
2 *Tschechien*	95.5	(0.4)	77.9	(0.9)	41.1	(1.2)	9.3	(0.7)
Portugal	91.5	(0.8)	69.8	(1.3)	36.3	(1.3)	8.9	(0.7)
2 *Serbien*	92.5	(0.9)	73.8	(1.7)	39.1	(1.7)	8.7	(0.8)
Deutschland	94.2	(0.6)	74.9	(1.1)	37.9	(1.0)	8.3	(0.5)
2 *Zypern*	91.1	(0.7)	71.0	(1.2)	35.8	(1.1)	8.0	(0.5)
3 *Niederlande*	97.8	(0.4)	83.1	(1.1)	43.5	(1.3)	8.0	(0.7)
2 3 *Dänemark*	95.3	(0.5)	75.5	(1.0)	36.7	(1.1)	7.4	(0.6)
2 *Italien*	91.2	(0.8)	69.2	(1.5)	33.6	(1.3)	7.2	(0.7)
2 3 *Belgien (Fläm. Gem.)*	95.3	(0.7)	73.1	(1.4)	36.2	(1.3)	6.7	(0.6)
1 2 3 *Neuseeland*	83.4	(0.9)	56.7	(1.2)	26.8	(1.2)	6.5	(0.4)
Slowakei	90.6	(1.0)	73.0	(1.5)	35.5	(1.3)	6.4	(0.6)
2 3 *Kanada*	89.8	(0.5)	64.7	(0.8)	28.8	(0.7)	6.3	(0.4)
2 *Slowenien*	93.4	(0.5)	70.8	(1.0)	32.5	(0.9)	5.5	(0.6)
2 *Spanien*	90.2	(0.7)	62.2	(1.0)	25.4	(0.9)	4.2	(0.4)
2 *Belgien (Frz. Gem.)*	87.5	(0.9)	57.9	(1.3)	21.9	(0.9)	3.3	(0.3)
2 *Frankreich*	85.4	(1.1)	56.3	(1.5)	20.3	(1.3)	2.6	(0.5)
2 3 *Chile*	70.7	(1.6)	36.0	(1.4)	9.1	(0.6)	0.8	(0.2)
Benchmark-Teilnehmer								
Dubai, VAE	94.1	(0.4)	82.4	(0.5)	56.4	(0.7)	22.3	(0.7)



% der Schüler:innen, die Kompetenzstufe I erreichen
% der Schüler:innen, die Kompetenzstufe II erreichen
% der Schüler:innen, die Kompetenzstufe III erreichen
% der Schüler:innen, die Kompetenzstufe IV erreichen
% der Schüler:innen, die Kompetenzstufe V erreichen

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

© TIMSS 2023

gegen ist dieser Wert etwas höher. Hierbei handelt es sich um Schweden (77 %), Finnland (76 %) und Dänemark (76 %).

Zusammenfassend ist für Deutschland somit festzustellen, dass ein Viertel der Schüler:innen über unterdurchschnittliche mathematische Kompetenzen verfügt und das Weiterlernen dieser Kinder in der Sekundarstufe I im Fach Mathematik mit erheblichen Schwierigkeiten verbunden sein kann. Auf der anderen Seite des Leistungsspektrums sind es lediglich 8 Prozent der Schüler:innen, die die höchste Kompetenzstufe erreichen. Dieser Wert liegt deutlich unter den Werten, die für die allesamt aus Asien stammenden Staaten der Spitzengruppe festzustellen sind, aber auch unter den Werten europäischer Länder wie England (22 %), Litauen (20 %), Irland (16 %) oder Rumänien (16 %).

Vergleich der Kompetenzstufen in 2007, 2019 und 2023: Tabelle 3.12 stellt die Verteilung der Viertklässler:innen auf die Kompetenzstufen im Vergleich der TIMSS-Erhebungen 2007, 2019 und 2023 dar. Hierbei werden erneut die Staaten und Regionen in den Fokus genommen, die an allen drei Erhebungsrunden teilgenommen haben.

Der Vergleich der Erhebungsrunden von 2007 und 2023 zeigt, dass für zwei Staaten sowohl für die Kompetenzstufe I als auch für die Kompetenzstufe V eine signifikante Steigerung der Anteile der Lernenden zu verzeichnen ist. Hierbei handelt es sich um die USA und um Deutschland. Während nur bei Hongkong ebenfalls ein signifikanter Zuwachs des Anteils von Lernenden auf Kompetenzstufe I zu beobachten ist, sind es analog für die Kompetenzstufe V 14 Staaten inklusive Deutschland. Sechs teilnehmenden Ländern gelingt sowohl eine signifikante Reduktion der Anteile der Lernenden auf den Kompetenzstufen I und II als auch eine signifikante Steigerung des Anteils der Lernenden auf Kompetenzstufe V. In drei weiteren Staaten ist eine signifikante Reduktion im Hinblick auf die Kompetenzstufe II zu beobachten. Deutschland gehört nicht dazu.

In der vergleichenden Analyse der beiden jüngsten Studienzyklen aus den Jahren 2019 und 2023 lassen sich folgende Erkenntnisse festhalten: Deutschland gehört zu den elf Staaten, deren Anteil an Schüler:innen auf Kompetenzstufe I sich nicht signifikant verändert. In sieben Staaten steigt der Anteil der Lernenden auf Kompetenzstufe I, darunter beispielsweise England, Ungarn, Italien und die USA, während es Georgien als einzigem Staat gelingt, diesen Anteil trotz der Pandemie zu senken. Auch bei den Kompetenzstufen II bis IV befindet sich Deutschland jeweils in der größeren Gruppe von 13 bis 15 Staaten, deren Anteil an Lernenden auf dieser Kompetenzstufe sich nicht signifikant ändert. Dagegen gehört Deutschland zu den sieben Staaten, bei denen sich der Anteil der Schüler:innen der höchsten Kompetenzstufe V signifikant steigert. Auffällig sind beim Vergleich der jüngsten Studienzyklen von 2019 und 2023 die Staaten Litauen und Australien, bei denen sich die Anteile auf den unteren Kompetenzstufen nicht signifikant ändern, während auf den beiden höchsten Kompetenzstufen eine signifikante Steigerung zu verzeichnen ist.

Tabelle 3.12: Verteilung der Schüler:innen auf die Kompetenzstufen in TIMSS 2007, 2019 und 2023 (Angaben in Prozent)

Fußnoten				Teilnehmer	TIMSS 2007 Kompetenzstufen										TIMSS 2019 Kompetenzstufen										TIMSS 2023 Kompetenzstufen									
2007	2019	2023	T		I		II		III		IV		V		I		II		III		IV		V		I		II		III		IV		V	
					%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
	2	2	8	Singapur	1.8	(0.3)	6.4	(0.8)	17.3	(1.1)	– 33.7	(1.3)	– 40.7	(2.1) +	0.8	(0.3) +	3.7	(0.5) +	11.8	(1.1)	29.5	(1.5)	54.2	(2.2) –	1.9	(0.2)	5.4	(0.4)	13.4	(0.6)	29.9	(0.8)	49.3	(1.1)
				Taiwan	1.1	(0.2) –	6.5	(0.5) –	26.4	(1.2) –	41.9	(1.4)	24.1	(1.2) +	0.3	(0.2)	3.8	(0.5)	17.5	(0.9)	41.6	(1.0)	36.7	(1.3) +	0.3	(0.1)	2.9	(0.2)	15.4	(0.7)	40.9	(0.8)	40.5	(1.0)
2 3	3	3	8	Hongkong	0.4	(0.1) +	2.8	(0.4) +	16.1	(1.3)	40.3	(1.3) –	40.4	(2.2)	0.3	(0.2) +	4.1	(0.6) +	17.1	(1.3)	40.6	(1.1) –	37.9	(1.9)	2.3	(0.4)	6.7	(0.8)	17.2	(1.3)	36.1	(1.6)	37.7	(2.3)
	3		8	Japan	2.2	(0.4) –	8.7	(0.6) –	27.7	(0.9) –	38.7	(1.3)	22.8	(1.2) +	0.6	(0.2)	4.5	(0.4)	20.6	(0.9)	41.1	(1.2)	33.1	(1.3)	0.7	(0.2)	4.8	(0.5)	21.8	(1.2)	41.2	(1.3)	31.6	(1.4)
2	2	2	8	Litauen	5.6	(0.7) –	17.1	(1.3) –	35.4	(1.1) –	32.2	(1.2) +	9.7	(0.7) +	3.8	(0.6)	14.8	(1.0) –	33.7	(1.2) –	34.3	(1.2) +	13.4	(1.1) +	2.6	(0.4)	10.9	(0.9)	28.6	(0.9)	38.2	(1.0)	19.8	(1.1)
1 3	1 2	1 2 3	8	England	5.8	(0.7)	15.3	(0.9)	31.0	(1.1) –	31.7	(0.9)	16.2	(1.2) +	3.9	(0.5) +	13.4	(1.0)	29.4	(1.5)	32.0	(1.1)	21.3	(1.4)	5.8	(0.7)	14.0	(0.8)	27.0	(1.0)	31.6	(0.9)	21.5	(1.0)
3	3	3	8	Niederlande	2.0	(0.4)	14.2	(1.2)	42.1	(1.2)	35.1	(1.5)	6.6	(0.7)	1.7	(0.4)	14.2	(0.9)	40.3	(1.6)	36.7	(1.5)	7.2	(0.9)	2.2	(0.4)	14.7	(0.9)	39.6	(1.0)	35.6	(1.2)	8.0	(0.7)
		2		Tschechien	11.8	(1.1) –	29.6	(1.2) –	39.9	(1.3)	16.9	(1.5) +	1.9	(0.4) +	4.3	(0.6)	17.3	(1.0)	36.2	(1.1)	32.1	(1.1)	10.2	(1.0)	4.5	(0.4)	17.6	(0.6)	36.8	(0.9)	31.8	(0.9)	9.3	(0.7)
	2	2		Schweden	6.8	(0.7)	25.6	(1.0) –	43.7	(1.0) –	21.2	(1.3) +	2.7	(0.3) +	5.6	(0.7)	20.0	(1.0)	38.8	(1.2)	27.9	(1.4)	7.6	(0.8) +	5.1	(0.8)	18.0	(1.1)	36.2	(1.2)	30.8	(1.4)	9.9	(0.8)
				Australien	8.5	(1.0)	21.0	(1.2)	35.6	(1.2) –	26.1	(1.7)	8.8	(0.8) +	9.8	(1.0)	20.6	(1.0)	34.0	(1.1)	25.5	(1.0) +	10.1	(0.9) +	9.4	(0.6)	18.6	(0.9)	30.7	(0.9)	28.3	(0.7)	12.9	(0.7)
				Deutschland	4.1	(0.5) +	17.5	(0.9)	41.3	(0.9) –	31.6	(1.1)	5.6	(0.5) +	4.4	(0.6)	20.9	(1.1)	39.1	(1.0)	29.5	(1.4)	6.0	(0.6) +	5.8	(0.6)	19.3	(0.9)	37.0	(1.2)	29.6	(0.9)	8.3	(0.5)
3	3	2 3	8	Dänemark	4.7	(0.8)	19.0	(0.9)	40.3	(1.4)	29.3	(1.3)	6.6	(0.7)	4.7	(0.5)	20.5	(0.8)	37.5	(1.4)	28.8	(0.9)	8.4	(0.9)	4.7	(0.5)	19.8	(0.9)	38.9	(1.0)	29.2	(1.1)	7.4	(0.6)
		2		Ungarn	11.8	(1.2)	21.1	(1.1) –	31.9	(1.2)	26.2	(1.2) +	9.0	(0.8) +	6.8	(0.8) +	19.6	(1.0)	35.0	(1.1) –	29.6	(1.2)	9.0	(0.8) +	10.8	(1.3)	18.3	(0.8)	30.0	(1.1)	29.6	(1.1)	11.3	(0.8)
2 3	2 3	2 3		USA	4.8	(0.5) +	18.4	(0.9)	36.9	(0.9) –	29.8	(0.9) –	10.0	(0.8) +	6.8	(0.6) +	16.6	(0.8) +	30.6	(0.7) –	32.1	(1.1) –	13.9	(0.8)	12.9	(0.6)	19.2	(0.6)	28.6	(0.7)	26.4	(0.8)	12.9	(0.7)
	2			Slowakei	11.9	(1.5)	24.9	(1.1) –	36.8	(1.0)	21.6	(1.2) +	4.8	(0.7)	8.8	(1.2)	20.0	(1.2)	40.3	(1.3)	25.8	(1.5)	5.1	(0.7)	9.4	(1.0)	17.6	(1.0)	37.5	(1.2)	29.2	(1.1)	6.4	(0.6)
2		2	8	Italien	8.9	(1.0)	23.9	(1.1)	38.2	(1.6)	23.5	(1.3)	5.5	(0.7)	4.5	(0.5) +	22.5	(1.2)	42.7	(1.0) –	26.0	(1.3)	4.3	(0.5) +	8.8	(0.8)	22.0	(1.1)	35.6	(1.1)	26.4	(1.1)	7.2	(0.7)
2	2	2		Georgien	33.2	(2.0) –	31.6	(1.2) –	25.4	(1.2) +	8.4	(0.9) +	1.3	(0.4) +	16.2	(1.4) –	28.0	(1.2)	35.7	(1.3)	17.3	(1.3) +	2.8	(0.4)	11.8	(0.9)	25.2	(1.1)	36.7	(1.3)	22.0	(1.0)	4.4	(0.7)
1 2	1 2	1 2 3	8	Neuseeland	14.7	(1.0)	24.8	(0.9)	34.1	(0.9) –	21.3	(1.0)	5.1	(0.5) +	17.3	(0.9)	26.9	(0.9)	31.2	(0.9)	18.5	(1.0)	6.2	(0.5)	16.6	(0.9)	26.7	(0.9)	30.0	(0.9)	20.2	(1.0)	6.5	(0.4)
			4 8	Iran	47.3	(2.0) –	32.7	(1.4) –	16.9	(1.2) +	2.9	(0.5) +	0.2	(0.1) +	31.9	(1.5) +	29.2	(0.9)	25.7	(1.1)	11.3	(0.8) –	1.9	(0.3) –	40.6	(1.6)	28.2	(0.8)	22.9	(1.0)	7.4	(0.6)	0.8	(0.1)
				Benchmark-Teilnehmer																														
2 3 6	2		8	Dubai, VAE	30.8	(1.3) –	31.8	(1.7) –	25.6	(1.1)	10.0	(0.6) +	1.8	(0.3) +	5.1	(0.5)	14.7	(0.7) –	30.6	(1.0) –	33.7	(1.2)	16.0	(0.9) +	5.9	(0.4)	11.7	(0.4)	25.9	(0.6)	34.1	(0.7)	22.3	(0.7)
2	3	2 3	8	Québec, Kanada	4.2	(0.6) +	21.6	(1.3)	40.6	(1.8)	28.3	(1.8)	5.4	(0.7)	2.7	(0.5) +	17.3	(1.1) +	39.4	(1.1)	32.8	(1.4) –	7.8	(0.8)	6.4	(0.6)	22.8	(0.9)	38.0	(0.9)	26.5	(1.0)	6.3	(0.6)
2		2	8	Ontario, Kanada	5.9	(1.1) +	22.9	(1.3)	42.1	(1.5) –	24.8	(1.5)	4.3	(0.6) +	7.7	(0.9) +	24.6	(1.3)	35.4	(1.4)	25.0	(1.6)	7.2	(1.0)	10.7	(0.8)	25.2	(0.9)	35.6	(1.0)	22.1	(0.9)	6.4	(0.5)

\+ = Prozentwert in 2023 statistisch signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Prozentwert in 2023 statistisch signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

3.5.3 Inhaltsbereiche

Durch die Analyse der TIMSS-Ergebnisse ist es möglich, relative Stärken und Schwächen der Schüler:innen in Bezug auf die verschiedenen Inhaltsbereiche *Arithmetik*, *Messen und Geometrie* sowie *Daten* zu identifizieren.

Kompetenzen in den Inhaltsbereichen im internationalen Vergleich: Die Abbildung 3.8 zeigt die Leistungsmittelwerte in den drei Inhaltsbereichen für die EU- und OECD-Staaten sowie für weitere Teilnehmer, die auf der Gesamtskala Mathematik signifikant bessere Ergebnisse erzielt haben oder sich von Deutschland nicht signifikant unterscheiden. Dabei handelt es sich insgesamt um 37 Teilnehmerstaaten sowie eine Benchmark-Region. Darüber hinaus sind auch der internationale Mittelwert aller Staaten sowie die Mittelwerte VG_{OECD} und VG_{EU} vermerkt. Im linken Teil der Abbildung wird neben den Leistungsmittelwerten (*M*) und den Standardfehlern (*SE*) für jeden der drei Inhaltsbereiche auch die Abweichung zum Mittelwert der Gesamtskala angegeben. Erneut werden signifikant höhere bzw. niedrigere Leistungswerte anderer Staaten durch ein Plus- bzw. ein Minuszeichen gekennzeichnet. Die Tabelle wurde auf dieser Grundlage dreigeteilt. Die erste bzw. die dritte Gruppe bilden diejenigen Teilnehmer, die in allen drei Inhaltsbereichen signifikant höhere bzw. niedrigere Werte als Deutschland erzielten, während im Vergleich zu den Teilnehmern der zweiten Gruppe keine signifikanten Unterschiede zu beobachten sind. Darüber hinaus sind im rechten Teil der Abbildung die Abweichungen der Mittelwerte der Inhaltsbereiche zur Gesamtskala visualisiert. Auch hier signalisieren grüne Balken signifikante Abweichungen.

Zu den Ergebnissen: Die Viertklässler:innen in Deutschland erzielten im Bereich *Arithmetik* 524 Punkte, im Bereich *Messen und Geometrie* 527 Punkte und im Bereich *Daten* 520 Punkte. Der Mittelwert im Inhaltsbereich *Daten* fällt signifikant niedriger aus als der Mittelwert Deutschlands für die Gesamtskala (524), während der Mittelwert im Inhaltsbereich *Messen und Geometrie* in dieser Hinsicht signifikant höher ausfällt. Die Differenz zwischen den Leistungswerten in *Messen und Geometrie* einerseits und *Daten* andererseits beträgt 7 Punkte. Der Mittelwert im Bereich *Arithmetik* entspricht dem der Gesamtskala.

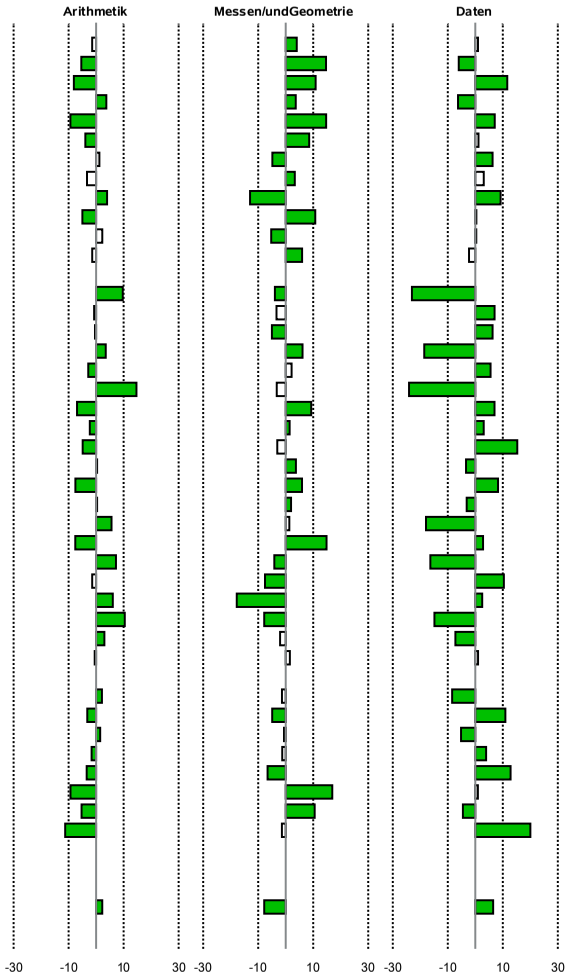
Der Vergleich mit den Staaten der *deutschen Leistungsgruppe* zeigt, dass für alle anderen Staaten, die auf der Gesamtskala keine zu Deutschland signifikanten Unterschiede der Testergebnisse aufweisen, eine größere Heterogenität zu erkennen ist. So liegen für Bulgarien (39), Serbien (24), Ungarn und die flämische Gemeinschaft in Belgien (jeweils 23), Portugal (18), Finnland (17), Dänemark (16) und Schweden (8) größere Differenzen zwischen den Leistungswerten des jeweils stärksten und schwächsten Inhaltsbereiches vor.

Im internationalen Vergleich zeigt sich darüber hinaus, dass zwölf teilnehmende Staaten in allen drei Inhaltsbereichen signifikant bessere Leistungsmittelwerte als Deutschland erreichen, während sieben Teilnehmer in allen drei Inhaltsbereichen signifikant schwächer abschneiden.

Hinsichtlich des Vergleichs der Leistungsmittelwerte Deutschlands in den drei Inhaltsbereichen mit denen der VG_{EU} und VG_{OECD} ergeben sich folgende Erkenntnisse: In fünf von sechs Fällen liegen keine signifikanten Unterschiede vor. Lediglich im Vergleich zur VG_{OECD} (528) sind die Leistungswerte in Deutschland im Bereich *Daten* mit 520 Punkten signifikant niedriger.

Abbildung 3.8: Testleistung der Schüler:innen im internationalen Vergleich in den Inhaltsbereichen *Arithmetik*, *Messen und Geometrie* sowie *Daten* im Vergleich zur Gesamtskala

	Gesamtskala				Arithmetik^						Messen/Geometrie^						Daten^					
Teilnehmer	M_{GS}	*(SE)*	*SD*	*(SE)*	M_A	*(SE)*		*SD*	*(SE)*	$M_A - M_{GS}$	*(SE)*	M_M	*(SE)*		*SD*	*(SE)*	$M_M - M_{GS}$	*(SE)*	M_D	*(SE)*		*SD*	*(SE)*	$M_D - M_{GS}$	*(SE)*
2 *Singapur*	615	(2.3)	88	(1.4)	613	(2.3)	+	87	(1.4)	-1	(0.8)	619	(2.5)	+	97	(1.4)	4	(0.7)	616	(2.9)	+	93	(2.2)	1	(1.9)
Taiwan	607	(1.4)	67	(0.7)	602	(1.3)	+	68	(1.2)	-5	(0.6)	622	(1.5)	+	78	(0.9)	15	(1.0)	601	(1.7)	+	65	(1.0)	-6	(1.2)
Republik Korea (Südkorea)	594	(2.6)	75	(2.1)	586	(2.7)	+	77	(2.4)	-8	(1.4)	605	(2.7)	+	82	(2.1)	11	(0.9)	606	(3.0)	+	77	(2.6)	12	(2.3)
3 *Hongkong*	594	(4.0)	84	(2.2)	598	(4.2)	+	89	(2.7)	4	(1.1)	598	(4.5)	+	92	(3.2)	4	(1.7)	588	(4.0)	+	86	(2.2)	-6	(1.3)
Japan	591	(2.3)	71	(1.1)	581	(2.3)	+	73	(1.3)	-9	(1.0)	605	(2.7)	+	83	(1.7)	15	(1.2)	598	(2.5)	+	70	(1.9)	7	(1.5)
Macau	582	(0.9)	79	(0.9)	578	(0.8)	+	80	(0.7)	-4	(0.8)	591	(1.0)	+	87	(1.3)	9	(0.9)	583	(1.0)	+	81	(1.3)	1	(0.7)
2 *Litauen*	561	(2.4)	76	(1.4)	562	(2.3)	+	80	(1.2)	1	(0.9)	556	(2.5)	+	73	(1.3)	-5	(1.4)	567	(2.6)	+	85	(1.5)	6	(1.0)
123 *Türkei (5. Jgst.)*	553	(4.1)	99	(2.0)	550	(4.5)	+	104	(2.1)	-3	(1.8)	557	(4.4)	+	107	(2.4)	3	(1.3)	556	(3.9)	+	96	(2.4)	3	(1.8)
123 *England*	552	(2.7)	92	(1.9)	556	(2.9)	+	96	(2.0)	4	(1.6)	539	(3.2)	+	99	(2.0)	-13	(1.6)	561	(3.5)	+	98	(2.5)	9	(2.6)
2 *Polen*	546	(2.0)	76	(1.1)	541	(2.2)	+	77	(1.0)	-5	(1.0)	557	(2.4)	+	77	(1.2)	11	(1.2)	546	(2.3)	+	87	(1.7)	0	(1.7)
Irland	546	(2.9)	82	(1.6)	548	(3.0)	+	82	(2.1)	2	(1.3)	540	(3.2)	+	90	(1.7)	-5	(1.9)	546	(3.2)	+	86	(1.7)	0	(1.7)
Lettland	534	(2.6)	77	(1.3)	533	(2.8)	+	80	(1.2)	-1	(1.8)	540	(3.0)	+	77	(1.5)	6	(1.4)	532	(2.9)	+	84	(1.5)	-2	(1.7)
23 *Rumänien*	542	(4.8)	85	(2.2)	552	(5.0)	+	88	(2.1)	10	(1.4)	538	(5.2)		93	(2.4)	-4	(1.5)	519	(5.4)		94	(2.9)	-23	(1.9)
3 *Niederlande*	537	(2.0)	65	(1.3)	536	(2.2)	+	69	(1.7)	-1	(1.2)	534	(2.8)		61	(1.2)	-3	(2.2)	544	(2.4)	+	75	(1.8)	7	(1.4)
12 *Norwegen (5. Jgst.)*	531	(2.0)	74	(0.9)	530	(2.2)	+	77	(1.1)	0	(1.2)	526	(2.5)		77	(1.3)	-5	(2.0)	537	(2.1)	+	80	(1.8)	6	(1.1)
2 *Tschechien*	530	(2.0)	74	(1.0)	534	(2.0)	+	74	(1.5)	4	(0.9)	537	(1.9)	+	76	(1.7)	6	(1.1)	512	(2.7)	–	85	(1.1)	-19	(2.1)
2 *Schweden*	530	(2.6)	76	(1.6)	527	(2.5)		80	(1.8)	-3	(1.0)	532	(2.6)		76	(1.7)	2	(1.6)	535	(2.9)	+	84	(2.2)	6	(1.4)
Bulgarien	530	(3.6)	91	(3.1)	545	(3.3)	+	87	(3.0)	15	(0.9)	527	(3.9)		91	(3.3)	-3	(2.0)	506	(4.8)	–	117	(4.3)	-24	(2.2)
Finnland	529	(2.3)	80	(1.2)	522	(2.4)		80	(1.2)	-7	(0.9)	539	(2.6)	+	82	(1.7)	9	(1.5)	536	(2.8)	+	93	(1.8)	7	(0.9)
VG OECD	**525**	**(0.5)**	**80**	**(0.3)**	**523**	**(0.5)**		**82**	**(0.3)**	**-2**	**(0.2)**	**527**	**(0.5)**		**83**	**(0.3)**	**1**	**(0.3)**	**528**	**(0.5)**	+	**87**	**(0.3)**	**3**	**(0.3)**
Australien	525	(2.0)	91	(1.3)	520	(2.1)		97	(1.1)	-5	(1.1)	522	(2.5)		95	(1.4)	-3	(1.7)	540	(2.2)	+	94	(1.2)	15	(1.3)
Deutschland	524	(1.9)	75	(1.4)	524	(2.2)		77	(1.6)	0	(1.3)	527	(2.1)		75	(1.5)	4	(1.2)	520	(2.3)		87	(1.7)	-3	(1.1)
23 *Dänemark*	524	(2.1)	72	(1.1)	516	(2.3)	–	76	(1.4)	-8	(0.8)	530	(2.1)		73	(1.2)	6	(1.2)	532	(2.3)	+	76	(1.3)	8	(1.3)
VG EU	**524**	**(0.6)**	**78**	**(0.3)**	**524**	**(0.6)**		**80**	**(0.4)**	**0**	**(0.2)**	**526**	**(0.6)**		**80**	**(0.4)**	**2**	**(0.3)**	**521**	**(0.7)**		**88**	**(0.4)**	**-3**	**(0.4)**
2 *Serbien*	523	(3.3)	80	(1.9)	529	(3.3)		80	(1.9)	6	(1.0)	524	(3.7)		82	(1.9)	1	(2.1)	505	(4.0)	–	90	(2.7)	-18	(2.5)
23 *Belgien (Fläm. Gem.)*	521	(2.4)	71	(1.4)	513	(2.3)	–	72	(1.6)	-8	(0.9)	536	(3.0)	+	78	(1.9)	15	(1.3)	524	(2.7)		78	(1.7)	3	(1.2)
2 *Ungarn*	520	(3.5)	91	(2.5)	527	(3.4)		92	(2.7)	7	(1.4)	516	(3.7)	–	92	(2.3)	-4	(2.1)	504	(3.8)	–	105	(2.6)	-16	(1.5)
Portugal	517	(2.8)	82	(1.4)	516	(2.7)	–	84	(1.5)	-1	(1.1)	510	(3.0)	–	91	(1.9)	-8	(1.0)	528	(2.9)		87	(1.6)	10	(1.9)
23 *USA*	517	(2.6)	97	(0.9)	523	(2.6)		100	(1.2)	6	(0.9)	499	(2.9)	–	106	(1.5)	-18	(1.1)	519	(2.6)		98	(1.4)	3	(0.9)
2 *Zypern*	516	(2.4)	81	(1.3)	527	(2.2)		83	(1.4)	10	(0.8)	508	(2.5)	–	84	(1.4)	-8	(0.9)	501	(3.1)	–	95	(1.5)	-15	(1.5)
Slowakei	515	(3.0)	81	(2.1)	518	(2.8)		82	(2.4)	3	(1.3)	513	(3.0)	–	82	(2.3)	-2	(1.5)	508	(3.6)	–	95	(2.6)	-7	(2.5)
2 *Slowenien*	514	(1.8)	73	(0.9)	513	(1.9)	–	72	(1.0)	-1	(1.6)	515	(2.0)	–	73	(1.1)	2	(1.6)	515	(2.2)		86	(1.7)	1	(1.5)
2 *Italien*	513	(2.8)	80	(1.3)	515	(2.6)	–	80	(1.4)	2	(0.9)	512	(3.3)	–	89	(1.6)	-1	(2.2)	505	(3.4)	–	87	(1.6)	-8	(1.6)
23 *Kanada*	504	(1.5)	81	(0.9)	501	(1.6)	–	82	(1.0)	-3	(0.6)	499	(1.7)	–	87	(1.1)	-5	(0.6)	515	(1.8)	–	90	(1.0)	11	(0.6)
Internationaler Mittelwert	**503**	**(0.4)**	**84**	**(0.2)**	**507**	**(0.4)**	–	**86**	**(0.2)**	**1**	**(0.2)**	**504**	**(0.4)**	–	**89**	**(0.2)**	**-1**	**(0.2)**	**500**	**(0.4)**	–	**93**	**(0.3)**	**-5**	**(0.2)**
2 *Spanien*	498	(1.8)	75	(1.2)	497	(2.1)	–	77	(1.3)	-2	(0.6)	497	(2.2)	–	78	(1.4)	-1	(1.4)	502	(1.9)	–	82	(1.0)	4	(1.1)
123 *Neuseeland*	490	(2.6)	91	(1.2)	487	(2.4)	–	94	(1.3)	-3	(1.1)	484	(2.5)	–	93	(1.4)	-7	(1.6)	503	(2.4)	–	98	(2.1)	13	(1.7)
2 *Belgien (Frz. Gem.)*	489	(2.2)	76	(1.2)	480	(2.2)	–	77	(1.1)	-9	(1.0)	506	(2.2)	–	83	(1.6)	17	(1.0)	490	(2.9)	–	82	(1.0)	1	(2.3)
2 *Frankreich*	484	(2.8)	77	(1.2)	479	(2.9)	–	79	(1.2)	-5	(1.6)	495	(3.1)	–	77	(1.3)	11	(1.4)	480	(3.0)	–	85	(1.6)	-5	(2.3)
23 *Chile*	444	(2.7)	80	(1.3)	432	(2.4)	–	83	(1.4)	-11	(1.9)	442	(2.3)	–	83	(1.3)	-1	(1.4)	464	(2.9)	–	89	(2.0)	20	(2.8)
Benchmark-Teilnehmer																									
Dubai, VAE	557	(1.4)	91	(0.9)	559	(1.4)	+	94	(0.9)	2	(0.6)	549	(1.4)	+	99	(1.0)	-8	(1.5)	563	(1.4)	+	97	(1.1)	6	(1.0)



Keine statistisch signifikanten Unterschiede zur Gesamtskala ($p > .05$)
Statistisch signifikante Unterschiede zur Gesamtskala ($p \leq .05$)
\+ = Mittelwert statistisch signifikant höher als der Mittelwert von Deutschland ($p \leq .05$)
– = Mittelwert statistisch signifikant niedriger als der Mittelwert von Deutschland ($p \leq .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Vergleich der Kompetenzen in den Inhaltsbereichen 2007, 2019 und 2023: Die Abbildung 3.9 zeigt die nach Inhaltsbereichen getrennten mathematischen Kompetenzen für die Untersuchungen der Jahre 2007, 2019 und 2023. Im oberen Teil dieser Abbildung sind die Differenzwerte zwischen den Zyklen numerisch dargestellt. Liegen positive Werte vor, wurden höhere Punktwerte im Vergleich zum jeweils früheren Zyklus erzielt. Negative Werte deuten entsprechend auf das Gegenteil hin. Im unteren Teil der Abbildung folgt eine grafische Darstellung der Differenzen der Ergebnisse von 2019 und 2023 mittels Differenzbalken. Weiße Balken deuten auf statistisch nicht signifikante und grüne auf signifikante Veränderungen zwischen den Erhebungszyklen hin. Wie in Abbildung 3.6 und der Tabelle 3.12 zur Beschreibung von Trendvergleichen werden auch in Abbildung 3.9 die Staaten und Regionen gelistet, die an den drei angegebenen TIMSS-Zyklen teilgenommen haben. Die Reihung der Staaten entspricht den Veränderungen der Leistungsmittelwerte auf der Gesamtskala (siehe Abbildung 3.6).

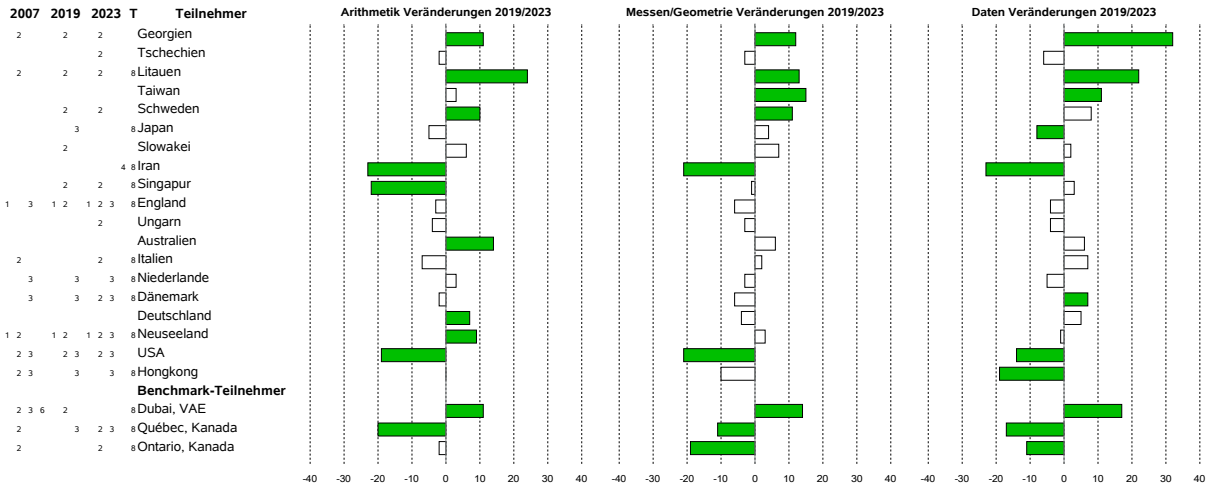
Aus Abbildung 3.9 geht für TIMSS 2023 hervor, dass in Deutschland im Bereich *Arithmetik* der gleiche Testwert wie in 2007 und signifikant höhere Testleistungen als in 2019 (+7) erzielt wurden. Im Bereich *Messen und Geometrie* haben sich im Vergleich zu den Ergebnissen von TIMSS 2007 und 2019 keine statistisch bedeutsamen Veränderungen ergeben. Hingegen sind im Bereich *Daten* im Vergleich zu TIMSS 2007 signifikant geringere Punktwerte (-12) zu verzeichnen, während es zum vorangegangenen Zyklus keinen signifikanten Unterschied gibt.

Im internationalen Vergleich der Daten von 2023 und 2007 zeigt sich folgendes Bild: In *Arithmetik* ist Deutschland einer von acht Staaten, für die in 2023 keine signifikante Veränderung vorliegt. Hingegen sind für die anderen elf Staaten und Regionen signifikant höhere Werte als 2007 festzustellen. In *Messen und Geometrie* zeigen sich für zehn Staaten signifikant höhere Leistungswerte als 2007, während für sechs Staaten signifikant niedrigere Werte vorliegen. Für drei Staaten – neben Deutschland auch Italien und Ungarn – lassen sich keine signifikanten Unterschiede feststellen. Im Inhaltsbereich *Daten* sind in 2023 für neun Staaten signifikant höhere, für drei Staaten (Deutschland, Hongkong, USA) signifikant niedrigere und für sieben Staaten keine Unterschiede bei den Leistungswerten festzustellen. Letztlich ist festzuhalten, dass für sechs Staaten in allen Inhaltsbereichen (darunter die Slowakei und Litauen) und für drei weitere Staaten in zwei Inhaltsbereichen (darunter England und Schweden) sowie für drei Staaten in einem Inhaltsbereich signifikant höhere Leistungswerte in 2023 als noch in 2007 vorliegen. Deutschland gehört nicht dazu.

Der Vergleich der Daten der Zyklen von 2023 und 2019 zeigt, dass sechs Teilnehmerstaaten – darunter Deutschland – in 2023 signifikant besser und drei signifikant schwächer in *Arithmetik* abschneiden. Für *Messen und Geometrie* gilt dies für vier bzw. drei Teilnehmerstaaten und für *Daten* für jeweils vier Staaten. Jeweils zwei Staaten schneiden in allen drei Inhaltsbereichen (Georgien und Litauen) bzw. in zwei Inhaltsbereichen (Schweden und Taiwan) in 2023 besser als in 2019 ab. Für vier Teilnehmerstaaten – neben Deutschland auch Dänemark, Australien und Neuseeland – ist in einem Inhaltsbereich ein signifikant höherer Leistungswert in 2023 im Vergleich zu 2019 zu verzeichnen.

Abbildung 3.9: Vergleich der Testleistungen zwischen TIMSS 2007, 2019 und 2023 in den Inhaltsbereichen *Arithmetik*, *Messen/Geometrie* sowie *Daten*

Fußnoten 2007	2019	2023	T	Teilnehmer	Arithmetik Veränderungen^ M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		Messen/Geometrie Veränderungen^ M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		Daten Veränderungen^ M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)	
2	2	2		Georgien	512	(2.9)	42	(4.7)	+	11	(4.6)	+	482	(3.9)	87	(7.1)	+	12	(5.7)	+	476	(3.2)	86	(6.3)	+	32	(5.6)	+
		2		Tschechien	534	(2.0)	48	(3.5)	+	-2	(3.1)		537	(1.9)	50	(3.8)	+	-3	(3.5)		512	(2.7)	30	(5.0)	+	-6	(4.0)	
2	2	2	8	Litauen	562	(2.3)	26	(3.2)	+	24	(3.6)	+	556	(2.5)	38	(3.9)	+	13	(3.9)	+	567	(2.6)	38	(4.4)	+	22	(4.0)	+
				Taiwan	602	(1.3)	19	(2.2)	+	3	(2.1)		622	(1.5)	56	(3.1)	+	15	(2.3)	+	601	(1.7)	25	(2.9)	+	11	(2.9)	+
	2	2		Schweden	527	(2.5)	32	(3.5)	+	10	(3.8)	+	532	(2.6)	29	(3.9)	+	11	(4.3)	+	535	(2.9)	8	(4.3)		8	(4.5)	
	3		8	Japan	581	(2.3)	17	(3.1)	+	-5	(2.9)		605	(2.7)	30	(3.7)	+	4	(3.8)		598	(2.5)	10	(4.3)	+	-8	(3.3)	–
	2			Slowakei	518	(2.8)	18	(4.8)	+	6	(4.6)		513	(3.0)	19	(6.1)	+	7	(4.8)		508	(3.6)	26	(6.5)	+	2	(5.5)	
			4 8	Iran	423	(3.5)	16	(4.9)	+	-23	(5.3)	–	424	(3.6)	16	(5.3)	+	-21	(5.1)	–	401	(3.8)	27	(6.3)	+	-23	(5.4)	–
	2	2	8	Singapur	613	(2.3)	2	(4.7)		-22	(4.6)	–	619	(2.5)	35	(5.1)	+	-1	(4.6)		616	(2.9)	19	(4.9)	+	3	(4.8)	
1 3	1 2	1 2 3	8	England	556	(2.9)	21	(4.2)	+	-3	(4.4)		539	(3.2)	-13	(4.6)	–	-6	(4.6)		561	(3.5)	10	(4.7)	+	-4	(4.7)	
		2		Ungarn	527	(3.4)	12	(4.9)	+	-4	(4.3)		516	(3.7)	9	(5.4)		-3	(5.0)		504	(3.8)	7	(5.7)		-4	(5.0)	
				Australien	520	(2.1)	17	(4.1)	+	14	(3.7)	+	522	(2.5)	-14	(4.5)	–	6	(4.1)		540	(2.2)	8	(4.7)		6	(4.0)	
2		2	8	Italien	515	(2.6)	5	(4.0)		-7	(3.6)		512	(3.3)	5	(4.9)		2	(4.6)		505	(3.4)	6	(5.3)		7	(4.5)	
3	3	3	8	Niederlande	536	(2.2)	-3	(3.1)		3	(3.1)		534	(2.8)	12	(3.9)	+	-3	(3.6)		544	(2.4)	-1	(3.7)		-5	(3.8)	
3	3	2 3	8	Dänemark	516	(2.3)	3	(3.7)		-2	(3.1)		530	(2.1)	-16	(3.8)	–	-6	(3.2)		532	(2.3)	5	(4.8)		7	(3.3)	+
				Deutschland	524	(2.2)	0	(3.1)		7	(3.0)	+	527	(2.1)	0	(3.2)		-4	(3.3)		520	(2.3)	-12	(4.4)	–	5	(3.9)	
1 2	1 2	1 2 3	8	Neuseeland	487	(2.4)	2	(3.5)		9	(3.8)	+	484	(2.5)	-11	(3.6)	–	3	(3.7)		503	(2.4)	-3	(3.8)		-1	(3.9)	
2 3	2 3	2 3		USA	523	(2.6)	-6	(3.7)		-19	(3.7)	–	499	(2.9)	-23	(4.2)	–	-21	(3.9)	–	519	(2.6)	-27	(3.9)	–	-14	(4.0)	–
2 3	3	3	8	Hongkong	598	(4.2)	-10	(5.6)		0	(5.5)		598	(4.5)	-15	(5.8)	–	-10	(5.5)		588	(4.0)	-12	(5.2)	–	-19	(5.4)	–
				Benchmark-Teilnehmer																								
2 3 6	2		8	Dubai, VAE	559	(1.4)	107	(2.4)	+	11	(2.2)	+	549	(1.4)	125	(3.6)	+	14	(2.5)	+	563	(1.4)	119	(3.4)	+	17	(2.4)	+
2	3	2 3	8	Québec, Kanada	510	(2.3)	-5	(3.8)		-20	(3.3)	–	521	(2.6)	-3	(4.7)		-11	(3.7)	–	518	(2.7)	-5	(5.2)		-17	(4.1)	–
2		2	8	Ontario, Kanada	499	(2.3)	4	(4.2)		-2	(4.3)		497	(2.6)	-33	(4.4)	–	-19	(4.1)	–	516	(2.9)	-29	(4.9)	–	-11	(4.9)	–



Keine statistisch signifikanten Veränderungen zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Veränderungen zwischen 2019 und 2023 ($p \leq .05$)
\+ = Mittelwert in 2023 statistisch signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 statistisch signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

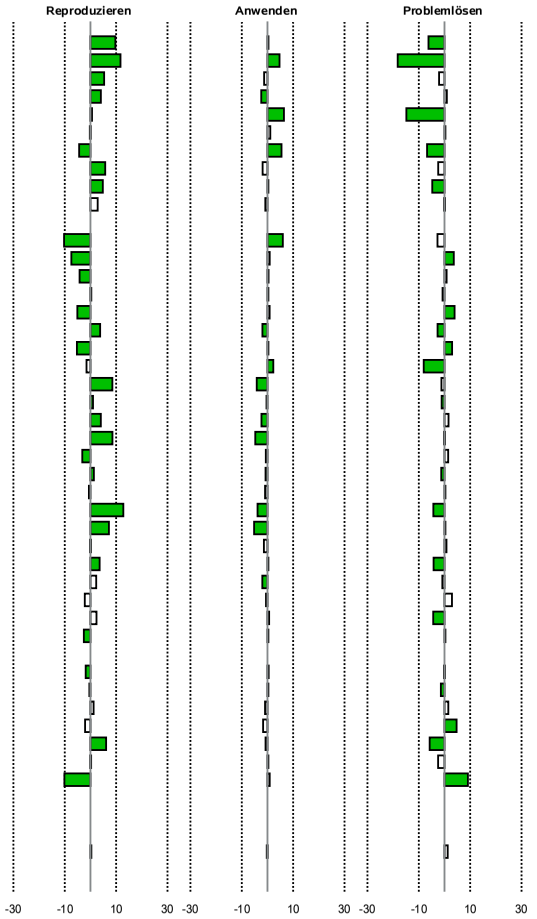
3.5.4 Kognitive Anforderungsbereiche

In diesem Abschnitt werden die Testergebnisse von TIMSS 2023 in Bezug auf die kognitiven Anforderungsbereiche dargestellt und anschließend im Vergleich zu den vorangegangenen Zyklen 2007 und 2019 eingeordnet.

Kompetenzen in den Anforderungsbereichen im internationalen Vergleich: Abbildung 3.10 zeigt die nach *Reproduzieren*, *Anwenden* und *Problemlösen* gegliederten Leistungsmittelwerte der Teilnehmerstaaten von TIMSS 2023. Der Aufbau der Abbildung orientiert sich an dem von Abbildung 3.8.

Abbildung 3.10: Testleistung der Schüler:innen in den kognitiven Anforderungsbereichen *Reproduzieren*, *Anwenden* und *Problemlösen* im Vergleich zur Gesamtskala

	Gesamtskala				Reproduzieren^						Anwenden^						Problemlösen^					
Teilnehmer	M_{GS}	(SE)	SD	(SE)	M_R	(SE)		SD	(SE)	M_R - M_{GS}	(SE)	M_A	(SE)		SD	(SE)	M_A - M_{GS}	(SE)	M_P	(SE)		SD	(SE)	M_P - M_{GS}	(SE)
2 *Singapur*	615	(2.3)	88	(1.4)	624	(2.3)	+	88	(1.5)	10	(1.0)	615	(2.3)	+	91	(1.5)	0	(0.7)	609	(2.4)	+	93	(1.8)	-6	(0.7)
Taiwan	607	(1.4)	67	(0.7)	619	(1.4)	+	66	(1.0)	12	(1.3)	612	(1.5)	+	70	(1.2)	5	(1.2)	589	(1.4)	+	73	(0.9)	-18	(0.6)
Republik Korea (Südkorea)	594	(2.6)	75	(2.1)	600	(3.0)	+	77	(2.0)	5	(1.5)	593	(2.6)	+	76	(1.9)	-1	(1.2)	592	(2.8)	+	85	(2.2)	-2	(1.5)
3*Hongkong*	594	(4.0)	84	(2.2)	598	(4.0)	+	86	(2.5)	4	(1.2)	592	(4.2)	+	86	(2.7)	-2	(1.1)	595	(4.4)	+	91	(2.5)	1	(2.5)
Japan	591	(2.3)	71	(1.1)	591	(2.5)	+	70	(1.3)	1	(1.7)	597	(2.5)	+	75	(1.2)	6	(1.2)	576	(2.5)	+	75	(1.6)	-15	(0.8)
Macau	582	(0.9)	79	(0.9)	582	(1.4)	+	77	(1.5)	0	(0.9)	583	(1.1)	+	82	(0.8)	1	(1.2)	582	(1.2)	+	83	(1.5)	0	(1.3)
2 *Litauen*	561	(2.4)	76	(1.4)	556	(2.3)	+	73	(1.2)	-4	(1.4)	566	(2.3)	+	78	(1.3)	5	(1.1)	554	(2.5)	+	87	(1.5)	-7	(0.8)
123*England*	552	(2.7)	92	(1.9)	558	(2.9)	+	98	(2.4)	6	(1.1)	550	(2.7)	+	93	(2.0)	-2	(1.1)	550	(3.3)	+	93	(2.1)	-2	(2.0)
Irland	546	(2.9)	82	(1.6)	551	(3.3)	+	87	(2.0)	5	(1.4)	546	(3.0)	+	83	(1.7)	0	(1.0)	541	(2.9)	+	87	(1.9)	-5	(1.4)
3*Niederlande*	537	(2.0)	65	(1.3)	540	(2.6)	+	67	(1.2)	3	(1.7)	536	(2.2)	+	64	(1.3)	-1	(0.9)	537	(2.3)	+	75	(1.9)	0	(1.5)
123*Türkei (5. Jgst.)*	553	(4.1)	99	(2.0)	543	(4.9)		98	(1.9)	-10	(2.6)	559	(4.4)	+	102	(2.0)	6	(1.2)	551	(4.5)	+	109	(2.1)	-3	(1.4)
2 *Polen*	546	(2.0)	76	(1.1)	539	(2.2)		77	(1.5)	-7	(1.4)	547	(2.3)	+	75	(1.2)	1	(1.3)	550	(2.5)	+	86	(1.5)	4	(1.3)
23*Rumänien*	542	(4.8)	85	(2.2)	538	(4.7)		86	(2.1)	-4	(1.6)	542	(5.0)	+	88	(2.3)	0	(1.3)	543	(5.2)	+	91	(2.3)	1	(2.1)
Lettland	534	(2.6)	77	(1.3)	534	(2.5)		78	(1.8)	0	(1.5)	534	(2.7)	+	78	(1.3)	0	(1.3)	534	(2.8)	+	87	(1.5)	-1	(1.8)
12 *Norwegen (5. Jgst.)*	531	(2.0)	74	(0.9)	525	(2.4)	–	77	(1.4)	-5	(1.6)	531	(2.0)	+	73	(1.1)	1	(1.1)	534	(2.2)	+	79	(1.0)	4	(1.3)
2 *Tschechien*	530	(2.0)	74	(1.0)	534	(2.5)		72	(1.2)	4	(1.1)	528	(2.0)	+	76	(1.1)	-2	(0.9)	528	(2.4)		81	(1.5)	-3	(0.8)
2 *Schweden*	530	(2.6)	76	(1.6)	525	(2.4)	–	79	(1.6)	-5	(1.3)	530	(2.5)	+	75	(1.7)	0	(1.2)	533	(2.6)	+	83	(2.0)	3	(1.1)
Bulgarien	530	(3.6)	91	(3.1)	528	(3.6)		85	(3.1)	-2	(1.5)	532	(3.7)	+	94	(3.4)	2	(1.1)	522	(4.6)		111	(4.7)	-8	(1.6)
Finnland	529	(2.3)	80	(1.2)	538	(2.6)		84	(1.4)	9	(1.2)	525	(2.4)		80	(1.6)	-4	(1.0)	528	(2.9)		86	(1.5)	-1	(1.5)
VG OECD	**525**	**(0.5)**	**80**	**(0.3)**	**526**	**(0.5)**	–	**82**	**(0.3)**	**1**	**(0.3)**	**525**	**(0.5)**	+	**81**	**(0.3)**	**0**	**(0.2)**	**524**	**(0.5)**		**86**	**(0.3)**	**-1**	**(0.3)**
Australien	525	(2.0)	91	(1.3)	529	(2.5)		96	(1.4)	4	(1.2)	523	(2.0)		91	(1.1)	-2	(0.5)	526	(2.4)		96	(1.9)	2	(1.4)
Deutschland	524	(1.9)	75	(1.4)	532	(2.5)		78	(1.4)	9	(1.7)	519	(2.2)		77	(1.4)	-5	(1.3)	524	(2.2)		79	(1.9)	0	(1.2)
23*Dänemark*	524	(2.1)	72	(1.1)	521	(2.2)	–	77	(1.2)	-3	(0.7)	523	(2.5)		71	(1.2)	-1	(1.6)	525	(2.5)		75	(1.5)	1	(1.3)
VG EU	**524**	**(0.6)**	**78**	**(0.3)**	**525**	**(0.6)**	–	**80**	**(0.4)**	**1**	**(0.3)**	**523**	**(0.6)**		**79**	**(0.4)**	**-1**	**(0.3)**	**522**	**(0.6)**		**84**	**(0.4)**	**-1**	**(0.3)**
2 *Serbien*	523	(3.3)	80	(1.9)	522	(3.3)	–	76	(2.1)	-1	(1.6)	522	(3.3)		82	(2.0)	-1	(1.5)	523	(3.4)		87	(2.1)	0	(1.0)
23*Belgien (Fläm. Gem.)*	521	(2.4)	71	(1.4)	534	(2.9)		78	(1.6)	13	(1.4)	517	(2.4)		73	(1.5)	-4	(0.8)	516	(2.8)	–	76	(2.0)	-4	(1.3)
2 *Ungarn*	520	(3.5)	91	(2.5)	527	(3.5)		94	(2.4)	7	(1.5)	515	(3.4)		95	(2.4)	-5	(0.9)	521	(3.3)		96	(2.1)	0	(1.0)
Portugal	517	(2.8)	82	(1.4)	517	(2.8)	–	83	(1.3)	0	(1.1)	516	(2.9)		85	(1.4)	-1	(0.8)	518	(2.7)		85	(1.6)	1	(1.0)
23*USA*	517	(2.6)	97	(0.9)	521	(2.9)	–	100	(1.4)	4	(1.4)	517	(2.7)		97	(0.9)	0	(0.7)	513	(2.8)	–	102	(1.3)	-4	(1.2)
2 *Zypern*	516	(2.4)	81	(1.3)	519	(2.3)	–	84	(1.5)	2	(1.2)	514	(2.6)		82	(1.6)	-2	(1.0)	515	(2.4)	–	86	(1.9)	-1	(1.1)
Slowakei	515	(3.0)	81	(2.1)	513	(3.2)	–	85	(2.7)	-2	(2.0)	514	(3.1)		81	(2.1)	-1	(1.8)	518	(3.0)		87	(1.9)	3	(2.2)
2 *Slowenien*	514	(1.8)	73	(0.9)	516	(2.0)	–	71	(0.9)	2	(1.5)	514	(1.8)		76	(1.1)	1	(1.3)	509	(2.0)	–	80	(1.2)	-4	(1.4)
2 *Italien*	513	(2.8)	80	(1.3)	511	(2.9)	–	83	(1.2)	-3	(0.9)	513	(2.9)		82	(1.5)	0	(1.4)	513	(2.7)	–	79	(1.5)	0	(1.3)
23*Kanada*	504	(1.5)	81	(0.9)	502	(1.6)	–	85	(1.1)	-2	(0.5)	504	(1.6)	–	80	(0.9)	0	(0.6)	504	(1.8)	–	86	(1.4)	0	(0.9)
Internationaler Mittelwert	**503**	**(0.4)**	**84**	**(0.2)**	**505**	**(0.4)**	–	**87**	**(0.2)**	**0**	**(0.2)**	**505**	**(0.4)**	–	**86**	**(0.2)**	**0**	**(0.2)**	**504**	**(0.4)**	–	**89**	**(0.3)**	**-1**	**(0.2)**
2 *Spanien*	498	(1.8)	75	(1.2)	500	(2.5)	–	80	(0.8)	1	(1.5)	497	(1.7)	–	75	(0.7)	-1	(0.8)	500	(1.9)	–	78	(1.1)	1	(1.1)
123*Neuseeland*	490	(2.6)	91	(1.2)	488	(2.7)	–	96	(2.0)	-2	(1.5)	489	(2.4)	–	90	(1.5)	-2	(1.3)	495	(2.3)	–	97	(1.4)	5	(1.9)
2 *Belgien (Frz. Gem.)*	489	(2.2)	76	(1.2)	496	(2.4)	–	83	(1.2)	6	(1.1)	489	(2.6)	–	77	(1.3)	-1	(1.5)	484	(2.4)	–	82	(1.3)	-6	(1.1)
2 *Frankreich*	484	(2.8)	77	(1.2)	484	(3.4)	–	80	(1.7)	0	(2.6)	484	(3.0)	–	78	(1.2)	0	(1.3)	482	(2.8)	–	81	(1.5)	-2	(1.4)
23*Chile*	444	(2.7)	80	(1.3)	433	(2.7)	–	80	(2.1)	-10	(0.9)	444	(2.3)	–	82	(1.1)	1	(1.3)	453	(3.0)	–	87	(1.5)	9	(1.7)
Benchmark-Teilnehmer																									
Dubai, VAE	557	(1.4)	91	(0.9)	557	(1.2)	+	93	(1.1)	0	(0.7)	556	(1.4)	+	93	(1.0)	0	(0.7)	558	(1.7)	+	90	(1.0)	1	(1.8)



☐ Keine statistisch signifikanten Unterschiede zur Gesamtskala ($p > .05$)
■ Statistisch signifikante Unterschiede zur Gesamtskala ($p \leq .05$)
\+ = Mittelwert statistisch signifikant höher als der Mittelwert von Deutschland ($p \leq .05$)
– = Mittelwert statistisch signifikant niedriger als der Mittelwert von Deutschland ($p \leq .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Die Analyse zeigt, dass Schüler:innen in Deutschland am Ende ihrer Grundschulzeit im Bereich *Reproduzieren* 532 Punkte, im Bereich *Anwenden* 519 Punkte und im Bereich *Problemlösen* 524 Punkte erzielen. Die Testleistungen liegen im Bereich *Reproduzieren* signifikant über dem Mittelwert Deutschlands für die Gesamtskala Mathematik (524), während die Testleistungen aus dem Bereich *Anwenden* signifikant unterhalb dieses Wertes einzuordnen sind. Der Unterschied zwischen *Reproduzieren* und *Anwenden* beträgt 13 Punkte. Der Mittelwert im Bereich *Problemlösen* entspricht dem der Gesamtskala.

Im Vergleich zu den Teilnehmerstaaten, die auf der Gesamtskala Mathematik in einer Leistungsgruppe mit Deutschland liegen, ist der Leistungsunterschied zwischen stärkstem und schwächstem Bereich in nahezu allen Staaten geringer. Lediglich in Finnland ist er mit 13 Punkten gleich groß und in der flämischen Gemeinschaft in Belgien mit 18 Punkten größer.

Ferner zeigt sich, dass zehn Staaten in allen drei kognitiven Anforderungsbereichen signifikant höhere Leistungswerte als Deutschland erzielen, während für sechs Staaten in allen Bereichen signifikant niedrigere Leistungswerte festzustellen sind.

Eine separate Analyse der einzelnen kognitiven Anforderungsbereiche zeigt überdies, dass beim *Reproduzieren* für zehn, beim *Anwenden* für 19 und beim *Problemlösen* für 16 Staaten signifikant bessere Testleistungen als für Deutschland festzustellen sind. Hingegen sind signifikant schwächere Testleistungen bei 16 teilnehmenden Staaten im Bereich *Reproduzieren*, bei sechs Staaten im Bereich *Anwenden* und elf Staaten im Bereich *Problemlösen* zu verzeichnen.

Im Vergleich zu den jeweiligen Leistungsmittelwerten der VG_{EU} und VG_{OECD} in den kognitiven Anforderungsbereichen zeigt sich, dass beide Gruppen im *Reproduzieren* signifikant niedrigere Werte aufweisen. Hingegen sind die Leistungswerte im *Anwenden* für Deutschland signifikant niedriger als in der Vergleichsgruppe VG_{OECD}.

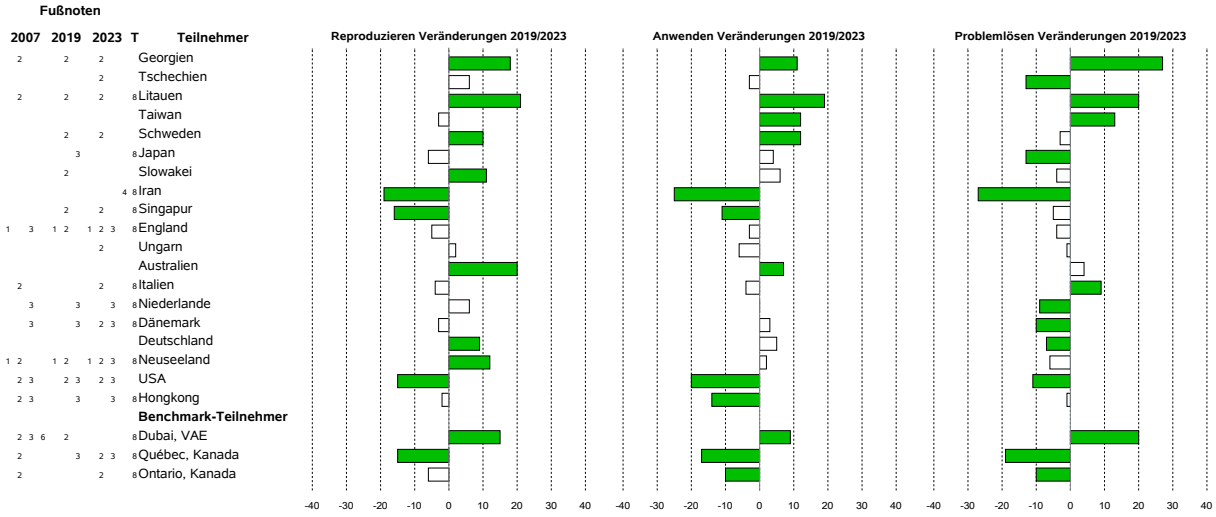
Vergleich der Leistungswerte in den kognitiven Anforderungsbereichen 2007, 2019 und 2023: In Abbildung 3.11 werden die Testergebnisse aus den drei ausgewählten Zyklen in den Bereichen *Reproduzieren*, *Anwenden* und *Problemlösen* berichtet. Für die Auswahl der teilnehmenden Staaten und den Aufbau der Abbildung werden die für die Abbildung 3.9 beschriebenen Kriterien herangezogen.

Abbildung 3.11 ist zu entnehmen, dass in Deutschland im Jahr 2023 im Bereich *Reproduzieren* signifikant bessere Testergebnisse erzielt wurden als in den Jahren 2007 und 2019. Im Bereich *Anwenden* wurden in 2023 signifikant schwächere Ergebnisse erzielt als im Jahr 2007, im Vergleich zu 2019 ist kein bedeutsamer Unterschied festzustellen. Für den Bereich *Problemlösen* ist im Vergleich der Resultate von 2023 und 2007 kein statistisch signifikanter Unterschied auszumachen, wobei im Vergleich zu 2019 signifikant geringere Werte erreicht werden.

Im internationalen Vergleich der Daten von 2023 und 2007 zeigt sich folgendes Bild: 13 Staaten, darunter Deutschland, schneiden im Jahr 2023 im Bereich *Reproduzieren* signifikant besser und zwei signifikant schlechter ab. Für *Anwenden* gilt dies für zehn bzw. zwei Staaten (Deutschland und Hongkong) und für *Problemlösen* für zehn Staaten bzw. einen Staat. Jeweils sechs Staaten schneiden in 2023 in allen drei bzw. in zwei kognitiven Anforderungsbereichen signifikant besser ab als noch in 2007, drei Staaten – Deutschland, Italien und die Niederlande – in einem Bereich. Überdies sind für lediglich drei Staaten in TIMSS

Abbildung 3.11: Vergleich der Testleistungen zwischen TIMSS 2007, 2019 und 2023 in den kognitiven Anforderungsbereichen *Reproduzieren*, *Anwenden* und *Problemlösen*

Fußnoten 2007	2019	2023	T	Teilnehmer	Reproduzieren Veränderungen^ M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		Anwenden Veränderungen^ M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		Problemlösen Veränderungen^ M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)	
2	2	2		Georgien	491	(3.3)	46	(5.3)	+	18	(5.1)	+	501	(3.1)	71	(5.6)	+	11	(4.8)	+	496	(2.9)	63	(5.4)	+	27	(5.4)	+
		2		Tschechien	534	(2.5)	62	(3.5)	+	6	(3.9)		528	(2.0)	35	(3.5)	+	-3	(3.3)		528	(2.4)	37	(4.4)	+	-13	(3.7)	–
2	2	2	8	Litauen	556	(2.3)	36	(3.6)	+	21	(3.6)	+	566	(2.3)	26	(3.5)	+	19	(3.5)	+	554	(2.5)	25	(3.8)	+	20	(4.1)	+
				Taiwan	619	(1.4)	33	(2.4)	+	-3	(2.4)		612	(1.5)	38	(2.4)	+	12	(2.1)	+	589	(1.4)	18	(2.4)	+	13	(2.3)	+
	2	2		Schweden	525	(2.4)	42	(3.5)	+	10	(3.9)	+	530	(2.5)	24	(3.4)	+	12	(3.8)	+	533	(2.6)	14	(3.8)	+	-3	(3.9)	
	3		8	Japan	591	(2.5)	24	(3.5)	+	-6	(3.2)		597	(2.5)	27	(3.3)	+	4	(3.2)		576	(2.5)	7	(3.3)	+	-13	(3.3)	–
	2			Slowakei	513	(3.2)	22	(5.4)	+	11	(4.6)	+	514	(3.1)	18	(5.4)	+	6	(4.6)		518	(3.0)	19	(5.4)	+	-4	(4.6)	
4	8			Iran	417	(3.2)	13	(5.0)	+	-19	(5.0)	–	425	(3.6)	28	(5.3)	+	-25	(5.4)	–	399	(4.1)	-2	(5.9)		-27	(5.9)	–
	2	2	8	Singapur	624	(2.3)	-1	(4.9)		-16	(4.5)	–	615	(2.3)	18	(4.7)	+	-11	(4.5)	–	609	(2.4)	25	(4.8)	+	-5	(4.7)	
1 3	1 2	1 2 3	8	England	558	(2.9)	12	(4.7)	+	-5	(4.4)		550	(2.7)	8	(4.3)		-3	(4.3)		550	(3.3)	11	(4.7)	+	-4	(4.7)	
		2		Ungarn	527	(3.5)	16	(5.0)	+	2	(4.4)		515	(3.4)	9	(5.1)		-6	(4.4)		521	(3.3)	11	(5.3)	+	-1	(4.5)	
				Australien	529	(2.5)	18	(5.1)	+	20	(4.1)	+	523	(2.0)	1	(4.3)		7	(3.5)	+	526	(2.4)	10	(4.4)	+	4	(3.8)	
2		2	8	Italien	511	(2.9)	-1	(4.5)		-4	(4.2)		513	(2.9)	14	(4.2)	+	-4	(3.9)		513	(2.7)	2	(4.3)		9	(4.0)	+
3	3	3	8	Niederlande	540	(2.6)	12	(3.5)	+	6	(3.3)		536	(2.2)	-4	(3.1)		0	(3.1)		537	(2.3)	0	(3.4)		-9	(3.7)	–
3	3	2 3	8	Dänemark	521	(2.2)	7	(3.6)	+	-3	(3.1)		523	(2.5)	-4	(3.8)		3	(3.4)		525	(2.5)	0	(3.3)		-10	(3.3)	–
				Deutschland	532	(2.5)	17	(3.3)	+	9	(3.4)	+	519	(2.2)	-11	(3.3)	–	5	(3.3)		524	(2.2)	-6	(3.6)		-7	(3.6)	–
1 2	1 2	1 2 3	8	Neuseeland	488	(2.7)	4	(3.8)		12	(3.8)	+	489	(2.4)	-4	(3.5)		2	(3.4)		495	(2.3)	-7	(3.6)		-6	(3.5)	
2 3	2 3	2 3		USA	521	(2.9)	-20	(4.0)	–	-15	(3.9)	–	517	(2.7)	-7	(3.9)		-20	(3.7)	–	513	(2.8)	-12	(3.7)	–	-11	(3.8)	–
2 3	3	3	8	Hongkong	598	(4.0)	-24	(5.4)	–	-2	(5.0)		592	(4.2)	-14	(5.7)	–	-14	(5.3)	–	595	(4.4)	-1	(5.8)		-1	(6.1)	
				Benchmark-Teilnehmer																								
2 3 6	2		8	Dubai, VAE	557	(1.2)	103	(2.7)	+	15	(2.2)	+	556	(1.4)	120	(2.3)	+	9	(2.1)	+	558	(1.7)	117	(3.4)	+	20	(2.5)	+
2	3	2 3	8	Québec, Kanada	520	(2.3)	1	(3.9)		-15	(3.5)	–	516	(2.2)	0	(3.6)		-17	(3.2)	–	505	(2.5)	-18	(4.1)	–	-19	(3.8)	–
2		2	8	Ontario, Kanada	498	(2.7)	0	(4.3)		-6	(4.6)		504	(2.5)	-9	(4.1)	–	-10	(4.2)	–	506	(2.6)	-20	(3.9)	–	-10	(4.4)	–



Keine statistisch signifikanten Veränderungen zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Veränderungen zwischen 2019 und 2023 ($p \leq .05$)
\+ = Mittelwert in 2023 statistisch signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 statistisch signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

2023 in einem oder in zwei Bereichen signifikant niedrigere Werte als in 2007 festzustellen, darunter auch für Deutschland.

Der Vergleich der Daten von 2019 und 2023 zeigt, dass im Bereich *Reproduzieren* sieben Staaten, darunter auch Deutschland, im Jahr 2023 signifikant besser und drei signifikant schwächer abschneiden. Für *Anwenden* gilt dies für jeweils vier Staaten und für *Problemlösen* für vier Staaten (besser) bzw. sieben Staaten (inkl. Deutschland, schlechter). Jeweils zwei Staaten schneiden in allen drei kognitiven Anforderungsbereichen in 2023 besser (Georgien und Litauen) bzw. schwächer (Iran und USA) ab als in 2019. Ein Staat (Singapur) schneidet in zwei Bereichen signifikant schwächer ab.

3.5.5 Einstellung und Selbstkonzept

Über die Erfassung mathematischer Kompetenzen am Ende der Grundschulzeit hinaus wurden in TIMSS mittels einer Fragebogenerhebung Daten zur Einstellung der Viertklässler:innen zur Mathematik sowie zu ihrem mathematikbezogenem Selbstkonzept erhoben. Die nationalen Ergebnisse der aktuellen Erhebungsrunde werden nachfolgend dargestellt und mit den Daten der TIMSS-Zyklen 2007 und 2019 verglichen.

Wie auch in den bisherigen Berichtsbänden wird auf eine international vergleichende Analyse, bezogen auf die Einstellung und das mathematikbezogene Selbstkonzept, verzichtet. Beide Variablen sind anfällig für Referenzgruppeneffekte, worunter Auswirkungen verstanden werden, die durch den Vergleich eigener Fähigkeiten mit denen einer Bezugsgruppe entstehen können. Hinsichtlich schulischer Kontexte können etwaige Referenzgruppeneffekte entstehen, da die Einstellung zur Mathematik sowie die Einschätzung eigener mathematischer Leistungen durchaus unterschiedlich ausfallen kann, je nachdem, ob für ein Kind eine eher leistungsstarke oder eine eher leistungsschwache Referenzgruppe (Klasse) vorliegt. Aufgrund der Ausrichtung der TIMS-Studie auf internationale Vergleichsanalysen von Leistungsdaten sind die Ergebnisse zu Einstellung und Selbstkonzept somit von Referenzgruppeneffekten betroffen, weswegen ausschließlich nationale Daten berichtet werden.

Einstellung zur Mathematik: Zur Erhebung der Einstellung gaben die teilnehmenden Viertklässler:innen in einem Fragebogen an, inwiefern sie den folgenden Aussagen *völlig zustimmen, eher zustimmen, eher nicht zustimmen* oder überhaupt nicht zustimmen.

1. *Ich lerne gern Mathematik.*
2. *Mathematik ist langweilig.*
3. *Ich mag Mathematik.*

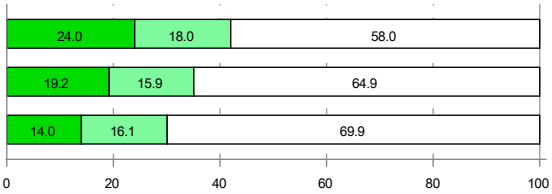
Die Ergebnisse werden auf Basis der Angaben der Schüler:innen bei Aussage 1 und 3 mit einem Punktesystem von 1 (*stimme überhaupt nicht zu*) bis 4 (*stimme völlig zu*) dargestellt. Im Gegensatz dazu erfolgt die Kodierung bei Aussage 2 in umgekehrter Reihenfolge. Im Anschluss erfolgt die Berechnung des Mittelwertes der drei Aussagen für jedes einzelne Kind. Der daraus resultierende Skalenwert gibt schließlich Aufschluss über die jeweilige positive Einstellung der Schüler:innen gegenüber Mathematik. Die interne Konsistenz der Skala (Cronbachs Alpha) liegt bei $\alpha = .87$.

Nachfolgend werden Gruppen von Schüler:innen mit niedriger, mittlerer und hoher positiver Einstellung zur Mathematik gebildet. Die positive Einstellung wird bei einem Mittelwert (M) von $1 \leq M < 2$ als *niedrig*, bei einem Mittelwert von $2 \leq M < 3$ als *mittel* und bei $3 \leq M \leq 4$ als *hoch* klassifiziert. Abbildung 3.12 stellt die Einstellung der Schüler:innen in Deutschland zum Fach Mathematik in den Erhebungsrunden von TIMSS 2007, 2019 und 2023 dar. Ferner ist die Verteilung der Kinder auf die Gruppen mit *niedriger*, *mittlerer* und *hoher* positiver Einstellung angegeben.

Für die aktuelle Erhebungsrunde liegt der mittlere Skalenwert für die Einstellung zur Mathematik bei 2.89. Dieser Wert lag in 2007 noch bei 3.15 bzw. in 2019 bei 3.04 und ist signifikant geringer im Vergleich zu den vorherigen Erhebungsrunden (kleiner Effekt von 2019 zu 2023 mit $d = .16$). Dementsprechend

Abbildung 3.12: Mittlere positive Einstellung von Schüler:innen zur Mathematik sowie prozentuale Verteilungen nach *niedriger*, *mittlerer* und *hoher* positiver Einstellung – TIMSS 2007, 2019 und 2023 im Vergleich

	Mittlerer Skalenwert		Positive Einstellung zur Mathematik niedrig		mittel		hoch	
	M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
2023	2.89	(0.02) ■	24.0	(0.7) ■	18.0	(0.5) ■	58.0	(0.9) ■
2019	3.04	(0.02) ▲	19.2	(1.0) ▼	15.9	(0.8) ▼	64.9	(1.2) ▲
2007	3.15	(0.02) ▲	14.0	(0.7) ▼	16.1	(0.6) ▼	69.9	(0.9) ▲



% der Schüler:innen mit niedriger positiver Einstellung zur Mathematik
% der Schüler:innen mit mittlerer positiver Einstellung zur Mathematik
% der Schüler:innen mit hoher positiver Einstellung zur Mathematik
▲ = Wert im Vergleich zu 2023 statistisch signifikant höher ($p \leq .05$)
▼ = Wert im Vergleich zu 2023 statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu 2023 nicht statistisch signifikant ($p > .05$)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

liegt für TIMSS 2023 im Mittel eine deutlich geringere positive Einstellung von Schüler:innen zur Mathematik vor als noch in 2007 und in 2019. Zudem dokumentiert der mittlere Skalenwert für Schüler:innen in Deutschland erstmals seit TIMSS 2007 keine *hohe* positive Einstellung zur Mathematik mehr.

Ein Vergleich der in Abbildung 3.12 angegebenen Verteilungen auf die drei Gruppen zeigt darüber hinaus, dass sich im Vergleich von 2023 zu 2007 und 2019 signifikante Verschiebungen erkennen lassen. Die Anteile der Schüler:innen mit *niedriger* (24 %) sowie *mittlerer* (18 %) positiver Einstellung sind in 2023 signifikant größer als in 2007 (14 % bzw. 16 %) sowie 2019 (19 % bzw. 16 %). Demgegenüber ist der Anteil von Schüler:innen mit einer *hohen* positiven Einstellung zur Mathematik in 2023 signifikant niedriger als in 2007 und 2019. Während dieser Anteil in 2007 noch bei 70 Prozent und in 2019 bei 65 Prozent lag, ist für die aktuelle Erhebungsrunde ein Anteil von 58 Prozent zu verzeichnen.

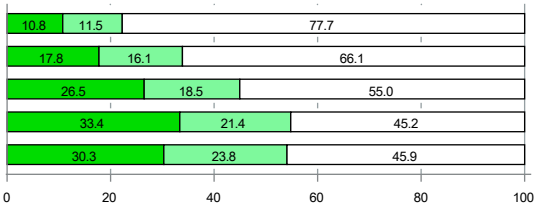
Obwohl der Anteil der Schüler:innen mit einer *hohen* positiven Einstellung zur Mathematik im Vergleich zu den Gruppen mit *niedriger* bzw. *mittlerer* positiver Einstellung nach wie vor deutlich am größten ist, wird die Einstellung der Grundschüler:innen in Deutschland zur Mathematik im Trend negativer.

Neben der Analyse der mittleren Skalenwerte im Trend ist ebenfalls von Interesse, welche Zusammenhänge zwischen der Einstellung zur Mathematik und erreichter Kompetenzstufe verzeichnet werden können. Abbildung 3.13 zeigt, dass Schüler:innen, die die Kompetenzstufe V (3.30) oder IV (3.07) erreichen, im Mittel eine *hohe* positive Einstellung zur Mathematik aufweisen. Eine *mittlere* positive Einstellung zur Mathematik ist demgegenüber für Schüler:innen der Kompetenzstufen III (2.81), II (2.62) sowie I (2.64) festzustellen.

Den Daten kann überdies entnommen werden, dass der Anteil der Schüler:innen mit *hoher* positiver Einstellung zur Mathematik bei den höheren Kompetenzstufen größer ist. Nichtsdestotrotz ist bei etwa der Hälfte der Schüler:innen mit geringen mathematischen Kompetenzen (Kompetenzstufen I und II) eine hohe positive Einstellung zur Mathematik vorhanden. Hingegen ist der Anteil an Schüler:innen mit *niedriger* positiver Einstellung geringer, je höher die Kompetenzstufe ist. Ausgenommen ist hierbei der Vergleich bezogen auf die Kompetenzstufen I und II, für die sich die Anteile an Schüler:innen mit *niedriger*, *mittlerer* und *hoher* positiver Einstellung nicht nennenswert voneinander unterscheiden.

Abbildung 3.13: Prozentuale Verteilung von Schüler:innen nach *niedriger*, *mittlerer* und *hoher* positiver Einstellung zur Mathematik differenziert nach den fünf Kompetenzstufen

Kompetenzstufe	Mittlerer Skalenwert		Positive Einstellung zur Mathematik: niedrig		mittel		hoch	
	M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
V	3.30	(0.04)	10.8	(2.0)	11.5	(2.3)	77.7	(2.4)
IV	3.07	(0.03)	17.8	(1.2)	16.1	(1.2)	66.1	(1.7)
III	2.81	(0.02)	26.5	(1.1)	18.5	(1.0)	55.0	(1.4)
II	2.62	(0.03)	33.4	(2.1)	21.4	(1.3)	45.2	(1.9)
I	2.64	(0.06)	30.3	(4.1)	23.8	(3.2)	45.9	(3.9)



% der Schüler:innen mit niedriger positiver Einstellung zur Mathematik
% der Schüler:innen mit mittlerer positiver Einstellung zur Mathematik
% der Schüler:innen mit hoher positiver Einstellung zur Mathematik

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Mathematikbezogenes Selbstkonzept: Im Rahmen der TIMS-Studie wurde neben der Einstellung auch das mathematikbezogene Selbstkonzept der Schüler:innen erfasst. In Anlehnung an die zuvor beschriebene Methodik zur Ermittlung der Einstellung wurden den teilnehmenden Kindern ebenfalls Aussagen[4] präsentiert, zu denen sie ihre Zustimmung oder Ablehnung äußern sollten. Auch hier wurde eine vierstufige Antwortskala verwendet, welche die Möglichkeiten *völlig zustimmen*, *eher zustimmen*, *eher nicht zustimmen* und *gar nicht zustimmen* umfasste.

1. *Normalerweise bin ich gut in Mathematik.*
2. *Mathematik fällt mir schwerer als vielen meiner Mitschülerinnen und Mitschüler.*
3. *Ich bin einfach nicht gut in Mathematik.*
4. *Mathematik fällt mir leicht.*
5. *Ich finde Mathematik schwerer als alle anderen Fächer.*

Die Antworten der Kinder werden hinsichtlich der Aussagen 1 und 4 mit einem Punktwert von 1 (*stimme überhaupt nicht zu*) bis 4 (*stimme völlig zu*) versehen, während für die Aussagen 2, 3 und 5 eine umgekehrte Reihenfolge der Punktwerte gilt. Im Anschluss erfolgt die Bildung des Mittelwertes für alle Schüler:innen, woraufhin – wie bereits dargelegt – die Einteilung in Gruppen mit *niedrigem* ($1 \leq M < 2$), *mittlerem* ($2 \leq M < 3$) und *hohem* ($3 \leq M \leq 4$) Selbstkonzept erfolgt. Die interne Konsistenz der Skala liegt bei einem Cronbachs Alpha von $\alpha = .82$.

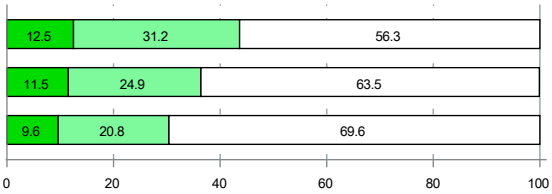
Die Ergebnisse zum mathematikbezogenen Selbstkonzept der Viertklässler:innen in Deutschland sind in Abbildung 3.14 dargestellt. Es sind die mittleren Skalenwerte für die Erhebungsrunden von 2007, 2019 und 2023 angegeben. Ferner ist die Verteilung der Kinder auf die Gruppen mit *niedrigem*, *mittlerem* und *hohem* mathematikbezogenen Selbstkonzept vermerkt.

Für TIMSS 2023 liegt der mittlere Skalenwert bei 2.99. Damit liegt – unter der erwähnten Einschränkung – ein signifikant geringerer Wert im Vergleich zu 2007 (3.18) und 2019 (3.11) vor. Somit ist das mathematikbezogene Selbstkon-

4 In den Erhebungsrunden vor TIMSS 2023 wurde statt der Aussagen 4 und 5 die Aussage *Ich lerne schnell in Mathematik.* genutzt. Durch Veränderungen im internationalen Erhebungsinstrument wurde diese Aussage nicht mehr in TIMSS 2023 erhoben. Stattdessen wurden die Aussagen 4 und 5 ergänzt. Aufgrund der veränderten Zusammenstellung der Aussagen sind die Daten von TIMSS 2023 im Trend lediglich eingeschränkt vergleichbar.

Abbildung 3.14: Mittleres mathematikbezogenes Selbstkonzept von Schüler:innen sowie prozentuale Verteilungen nach *niedrigem*, *mittlerem* und *hohem* Selbstkonzept – TIMSS 2007, 2019 und 2023 im Vergleich

	Mittlerer Skalenwert		Mathematikbezogenes Selbstkonzept niedrig		mittel		hoch	
	M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
2023[A]	2.99	(0.01) ■	12.5	(0.5) ■	31.2	(0.7) ■	56.3	(0.8) ■
2019	3.11	(0.02) ▲	11.5	(0.7) ■	24.9	(0.8) ▼	63.5	(1.1) ▲
2007	3.18	(0.01) ▲	9.6	(0.5) ▼	20.8	(0.7) ▼	69.6	(0.9) ▲



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen mit niedrigem mathematikbezogenen Selbstkonzept
% der Schüler:innen mit mittlerem mathematikbezogenen Selbstkonzept
% der Schüler:innen mit hohem mathematikbezogenen Selbstkonzept

A = Aufgrund veränderter Operationalisierung sind die Ergebnisse nur eingeschränkt mit vorherigen Zyklen vergleichbar.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

zept der Schüler:innen in Deutschland in der aktuellen Erhebungsrunde geringer als noch im ersten und im vergangenen Studienzyklus (kleiner Effekt von 2019 zu 2023 mit $d = .07$). Analog zu den Analysen zur Einstellung der Schüler:innen zur Mathematik ist im Durchschnitt nicht mehr ein *hohes*, sondern erstmals ein *mittleres* mathematikbezogenes Selbstkonzept zu konstatieren – wobei der mittlere Skalenwert mit 2.99 unmittelbar an der Schwelle zwischen *mittlerem* und *hohem* Selbstkonzept zu verorten ist.

Der Vergleich der in Abbildung 3.14 dargestellten Gruppen zeigt zudem, dass der Anteil an Schüler:innen mit *hohem* mathematikbezogenen Selbstkonzept in TIMSS 2023 bei 56 Prozent liegt und damit signifikant geringer ist als in 2019 (64 %) und 2007 (70 %). Ferner zeigen 31 Prozent der Viertklässler:innen ein *mittleres* mathematikbezogenes Selbstkonzept. Auch hier ist der Wert signifikant höher im Vergleich zu 2007 (21 %) und 2019 (25 %). Die Gruppe der Schüler:innen mit *niedrigem* mathematikbezogenen Selbstkonzept liegt in TIMSS 2023 bei 13 Prozent und unterscheidet sich damit kaum vom Wert für 2019 (12 %), ist jedoch signifikant größer als noch in 2007 (10 %).

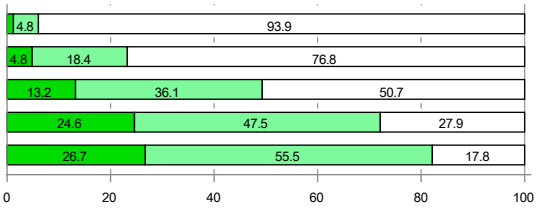
Insgesamt zeigen sich im Trend niedrigere mittlere Skalenwerte, die auf geringere mathematikbezogene Selbstkonzepte der Schüler:innen in Deutschland im Verlauf der Erhebungsrunden hinweisen. Gleichzeitig wird der Anteil der Gruppe mit *mittlerem* mathematikbezogenen Selbstkonzept größer, mit *hohem* hingegen kleiner. Allerdings ist die Gruppe der Lernenden mit einem *hohem* mathematikbezogenen Selbstkonzept nach wie vor deutlich am größten. Mehr als die Hälfte der Schüler:innen verfügt demnach über ein *hohes* mathematikbezogenes Selbstkonzept.

Auch hinsichtlich des mathematikbezogenen Selbstkonzepts wird analysiert, welche Zusammenhänge mit den erreichten Kompetenzstufen festzuhalten sind. Abbildung 3.15 bildet die entsprechenden Daten ab. Hier zeigt sich für Schüler:innen in Deutschland, welche die Kompetenzstufen V (3.64) oder IV (3.32) erreichen, im Mittel ein *hohes* mathematikbezogenes Selbstkonzept, wohingegen für die Kompetenzstufen III (2.89), II (2.52) sowie I (2.37) lediglich ein *mittleres* mathematikbezogenes Selbstkonzept festgestellt werden kann.

Darüber hinaus lassen sich in den Gruppen der Lernenden, die höhere Kompetenzstufen erreichen, deutlich mehr Schüler:innen mit hohem und deutlich weniger mit niedrigem mathematikbezogenen Selbstkonzept feststellen. So wei-

Abbildung 3.15: Prozentuale Verteilung von Schüler:innen nach *niedrigem*, *mittlerem* und *hohem* mathematikbezogenen Selbstkonzept differenziert nach den fünf Kompetenzstufen

Kompetenzstufe	Mittlerer Skalenwert		Mathematikbezogenes Selbstkonzept niedrig		mittel		hoch	
	M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
V	3.64	(0.02)	1.2	(0.6)	4.8	(1.2)	93.9	(1.3)
IV	3.32	(0.03)	4.8	(0.7)	18.4	(1.4)	76.8	(1.8)
III	2.89	(0.02)	13.2	(0.9)	36.1	(1.3)	50.7	(1.6)
II	2.52	(0.03)	24.6	(1.6)	47.5	(1.8)	27.9	(2.1)
I	2.37	(0.03)	26.7	(3.0)	55.5	(4.0)	17.8	(2.9)



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen mit niedrigem mathematikbezogenen Selbstkonzept
% der Schüler:innen mit mittlerem mathematikbezogenen Selbstkonzept
% der Schüler:innen mit hohem mathematikbezogenen Selbstkonzept

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

sen beispielsweise 77 Prozent der Schüler:innen, die Kompetenzstufe IV erreichten, ein *hohes* mathematikbezogenes Selbstkonzept auf, 18 Prozent ein *mittleres* und nur 5 Prozent ein *niedriges*. Auf der anderen Seite zeigen etwa für die Kompetenzstufe II nur 28 Prozent der Lernenden ein *hohes*, 48 Prozent ein *mittleres* und 25 Prozent ein *niedriges* mathematikbezogenes Selbstkonzept.

3.6 Zusammenfassung

TIMSS 2023 hat das Ziel, die mathematischen und naturwissenschaftlichen Kompetenzen von Schüler:innen am Ende ihrer Grundschulzeit im internationalen Vergleich zu erfassen. In diesem Abschnitt werden abschließend zentrale Befunde für das Fach Mathematik zusammengefasst.

Leistungsmittelwerte: Viertklässler:innen in Deutschland erzielen im internationalen Vergleich einen Leistungsmittelwert von 524 Punkten, womit sie sich im oberen Mittelfeld der 58 teilnehmenden Staaten befinden. Dieser Wert liegt über dem Skalenmittelwert von 500 und signifikant über dem internationalen Durchschnittswert von 503. Vom Mittelwert der an der Studie teilnehmenden EU-Staaten (VG_{EU}, 524) sowie der OECD-Staaten (VG_{OECD}, 525) unterscheidet sich der Mittelwert für Deutschland nicht signifikant. Deutschland liegt zusammen mit Schweden (530), Bulgarien (530), Finnland (529), Australien (525), Dänemark (524), Serbien (523), der Flämischen Gemeinschaft in Belgien (521), Ungarn (520) und Portugal (517) in einer Staatengruppe, deren Testergebnisse sich statistisch nicht signifikant voneinander unterscheiden.

Zu den asiatischen Staaten am oberen Ende der Rangreihe besteht ein sehr großer Abstand. Die in TIMSS 2023 leistungsstärksten drei Staaten sind Singapur (615), Taiwan (607) und Südkorea (594). Weitere 13 Staaten erzielen signifikant bessere Resultate, darunter auch England (552), Irland (546) und die Niederlande (537).

Dadurch, dass Deutschland im Jahr 2023 zum fünften Mal seit 2007 an TIMSS teilgenommen hat, können über einen Zeitraum von 16 Jahren Trendaussagen zur mathematischen Kompetenz von Viertklässler:innen gemacht werden. Dabei zeigt sich, dass Deutschland einer von fünf Teilnehmerstaaten ist, für die keine signifikanten Leistungsunterschiede zwischen 2007 und 2023 zu verzeich-

nen sind. Schüler:innen aus zwölf Staaten schnitten in TIMSS 2023 signifikant besser und zwei signifikant schwächer ab als 2007.

Im Vergleich zu TIMSS 2019 lässt sich festhalten, dass fünf Staaten in 2023 signifikant besser abschneiden als in 2019. Bei elf Staaten – zu denen auch Deutschland gehört – sind keine signifikanten Unterschiede der Differenzwerte festzustellen, während drei Staaten signifikant niedrigere Werte erzielen.

Leistungsstreuung: In den vorangegangenen TIMSS-Zyklen konnte für Deutschland eine im Vergleich zu vielen anderen Staaten eher geringe Leistungsstreuung verzeichnet werden. Dieser Befund setzt sich auch in TIMSS 2023 fort. Dafür sprechen sowohl die Größe der Standardabweichung (75) als auch die Leistungsspanne zwischen den jeweils 5 Prozent schwächsten und 5 Prozent stärksten Viertklässler:innen. Es gibt wenige Staaten, die eine vergleichbare oder eine geringere Leistungsheterogenität aufweisen.

Kompetenzstufen: Zur inhaltlichen Beschreibung der Mathematikleistung der Schüler:innen am Ende der Grundschulzeit wurden die Leistungswerte der Lernenden jeweils einer von fünf Kompetenzstufen zugeordnet. In TIMSS 2023 erreichen 94 Prozent der Schüler:innen in Deutschland mindestens die Stufe II, 75 Prozent mindestens die Stufe III, 38 Prozent mindestens die Stufe IV und 8 Prozent die Stufe V. Dementsprechend ist festzuhalten, dass ein Viertel der Schüler:innen in Deutschland am Ende ihrer Grundschulzeit nur die erste oder zweite Kompetenzstufe erreicht und damit allenfalls über einfache mathematische Fertigkeiten und Fähigkeiten verfügt. Besonders bedenklich ist dabei die Tatsache, dass 6 Prozent der Schüler:innen der untersten Kompetenzstufe zuzuordnen sind und nur über rudimentäres mathematisches Wissen verfügen. Von den Staaten der *deutschen Leistungsgruppe* zeigen fünf Staaten anteilig etwas mehr Schüler:innen auf der niedrigsten Kompetenzstufe als Deutschland, vier in etwa einen gleich großen Anteil. Mit Ausnahme der Türkei (8 %) weisen dagegen alle teilnehmenden Staaten mit signifikant besseren Leistungsmittelwerten auf der Gesamtskala als Deutschland einen geringeren oder gleich großen Anteil auf. Die Werte der Vergleichsgruppen VG_{OECD} (8 %) und VG_{EU} (7 %) liegen oberhalb des Wertes für Deutschland.

Insgesamt 75 Prozent der Grundschüler:innen aus Deutschland erreichen die Kompetenzstufe III. Sie zeigen damit mathematische Kenntnisse in einfachen Situationen und stellen Zusammenhänge zwischen Repräsentationen her. Die 8 Prozent der Schüler:innen auf der höchsten Kompetenzstufe V können ihre mathematischen Fertigkeiten und Fähigkeiten verständig beim Lösen verhältnismäßig komplexer Probleme nutzen und ihr Vorgehen erläutern. Es zeigt sich, dass dieser Anteil nicht nur bei den Staaten der Leistungsspitze auf der Gesamtskala, sondern auch bei Staaten ohne signifikante Leistungsunterschiede zu Deutschland wie Bulgarien (14 %), Ungarn (11 %) oder Finnland (11 %), höher ist. In den Vergleichsgruppen VG_{OECD} und VG_{EU} sind die Anteile mit 12 Prozent bzw. 10 Prozent im Vergleich zu Deutschland ebenfalls größer. Die größten Anteile von Viertklässler:innen auf der höchsten Kompetenzstufe V sind für Singapur (49 %), Taiwan (41 %) und Hongkong (38 %) zu verzeichnen. Von Werten dieser Größenordnung ist Deutschland – nach wie vor – weit entfernt.

Angesichts dieser Verteilung der Schüler:innen in Deutschland auf die Kompetenzstufen wird Handlungsbedarf deutlich. Für die 6 Prozent der Schüler:innen auf der untersten Kompetenzstufe wird anschlussfähiges Lernen in der Sekundarstufe I aufgrund der Anforderungen extrem schwierig sein. Dass in Deutschland

nur vergleichsweise wenige Grundschüler:innen (8 %) die höchste Kompetenzstufe erreichen, deutet darauf hin, dass Potenziale immer noch nicht hinreichend ausgeschöpft werden.

Der Vergleich der Erhebungsrunden von 2007 und 2023 zeigt allerdings auch, dass für 14 Staaten inklusive Deutschland für die Kompetenzstufe V eine signifikante Steigerung der Anteile der Lernenden zu verzeichnen ist. Für drei Staaten ist dieses auch für die Kompetenzstufe I der Fall (Hongkong, USA und Deutschland). Sechs Staaten gelingt eine signifikante Reduktion der Anteile der Lernenden auf den Kompetenzstufen I und II, in drei weiteren Staaten ist eine signifikante Reduktion im Hinblick auf die Kompetenzstufe II zu beobachten. Deutschland gehört nicht dazu.

Bezüglich der Veränderungen im Zeitraum von TIMSS 2019 zu TIMSS 2023, in dem die COVID-19-Pandemie die Bildungssysteme weltweit vor große Herausforderungen stellte, gehört Deutschland zu den elf Staaten, deren Anteil an Schüler:innen auf Kompetenzstufe I sich nicht signifikant verändert, während in sieben Staaten der Anteil der Lernenden auf Kompetenzstufe I steigt. Auch bei den Kompetenzstufen II bis IV befindet sich Deutschland jeweils in der größeren Gruppe von 13 bis 15 Staaten, deren Anteil an Lernenden auf dieser Kompetenzstufe sich nicht signifikant ändert. Dagegen gehört Deutschland zu den sieben Staaten, bei denen sich der Anteil der Schüler:innen der höchsten Kompetenzstufe V signifikant steigert.

Mathematische Inhaltsbereiche: Die Viertklässler:innen in Deutschland erzielten im Bereich *Arithmetik* 524 Punkte, im Bereich *Messen und Geometrie* 527 Punkte und im Bereich *Daten* 520 Punkte. Die Mittelwerte fallen im Inhaltsbereich *Daten* signifikant niedriger und im Inhaltsbereich *Messen und Geometrie* signifikant höher aus als der Mittelwert Deutschlands für die Gesamtskala (524). Die Differenz zwischen den Leistungswerten in *Messen und Geometrie* einerseits und *Daten* andererseits beträgt 7 Punkte. Der Mittelwert im Bereich *Arithmetik* entspricht dem der Gesamtskala. Im internationalen Vergleich zeigt sich darüber hinaus, dass zwölf Teilnehmerstaaten in allen drei Inhaltsbereichen signifikant bessere Leistungsmittelwerte als Deutschland erreichen, während sieben Staaten in allen drei Inhaltsbereichen signifikant schwächer abschneiden. Hinsichtlich des Vergleichs der Leistungsmittelwerte Deutschlands in den drei Inhaltsbereichen mit denen der VG_{EU} und VG_{OECD} zeigen sich keine signifikanten Abweichungen bis auf den Bereich *Daten*, wo Deutschland mit 520 Punkten im Vergleich zur VG_{OECD} (528 Punkte) signifikant niedriger liegt.

Der Vergleich der Untersuchungsergebnisse von TIMSS 2023 mit denen der Erhebungsrunden von 2007 und 2019 zeigt darüber hinaus, dass in Deutschland in 2023 im Bereich *Arithmetik* der gleiche Testwert wie in 2007 und signifikant höhere Testleistungen als in 2019 (+7) erzielt werden. Im Bereich *Messen und Geometrie* ergeben sich im Vergleich zu den Ergebnissen von TIMSS 2007 und 2019 keine statistisch bedeutsamen Veränderungen. Hingegen sind im Bereich *Daten* im Vergleich zu TIMSS 2007 signifikant geringere Punktwerte (-12) zu verzeichnen, während es zum vorangegangenen Zyklus von 2019 keinen signifikanten Unterschied gibt.

Kognitive Anforderungsbereiche: Die Ergebnisse zeigen, dass Schüler:innen in Deutschland am Ende ihrer Grundschulzeit im Bereich *Reproduzieren* 532 Punkte, im Bereich *Anwenden* 519 Punkte und im Bereich *Problemlösen* 524 Punkte erzielen. Die Testleistungen liegen im Bereich *Reproduzieren*

signifikant über dem Mittelwert Deutschlands für die Gesamtskala Mathematik (524), während die Testleistungen aus dem Bereich *Anwenden* signifikant unterhalb dieses Wertes einzuordnen sind. Der Unterschied zwischen *Reproduzieren* und *Anwenden* beträgt 13 Punkte. Der Mittelwert im Bereich *Problemlösen* entspricht dem der Gesamtskala. Ferner zeigt sich, dass zehn Staaten signifikant höhere Leistungswerte in allen drei kognitiven Anforderungsbereichen erzielen. Hingegen sind für sechs Staaten in allen Bereichen signifikant niedrigere Leistungswerte festzustellen.

Im Vergleich der Testergebnisse von TIMSS 2023 zu den Erhebungsrunden von 2007 und 2019 zeigt sich zudem, dass in Deutschland im Jahr 2023 beim *Reproduzieren* signifikant bessere Testergebnisse erzielt werden als in den Zyklen von 2007 und 2019. Im Bereich *Anwenden* werden in 2023 signifikant schwächere Ergebnisse erzielt als in TIMSS 2007; im Vergleich zu 2019 ist kein statistisch bedeutsamer Unterschied festzustellen. Für den Bereich *Problemlösen* ist im Vergleich zu 2007 kein statistisch signifikanter Unterschied auszumachen, wobei im Vergleich zu 2019 signifikant geringere Werte erreicht werden.

Einstellung: Der Vergleich der mittleren Skalenwerte der hier herangezogenen drei TIMSS-Erhebungsrunden zeigt auf einer Skala von 1 bis 4, dass für 2023 (2.89) im Vergleich zu 2007 (3.15) und zu 2019 (3.04) signifikant niedrigere mittlere Skalenwerte zu verzeichnen sind. Erstmalig seit 2007 dokumentieren diese keine hohe, sondern nur noch eine mittlere positive Einstellung zur Mathematik. Die Anteile der Schüler:innen mit *niedriger* (24 %) sowie *mittlerer* (18 %) positiver Einstellung sind in 2023 signifikant größer als in 2007 (14 % bzw. 16 %) sowie 2019 (19 % bzw. 16 %). Demgegenüber ist der Anteil von Schüler:innen mit einer *hohen* positiven Einstellung zur Mathematik in 2023 signifikant niedriger als in 2007 und 2019.

Obwohl der Anteil der Schüler:innen mit einer *hohen* positiven Einstellung zur Mathematik im Vergleich zu den Gruppen mit *niedriger* bzw. *mittlerer* positiver Einstellung nach wie vor deutlich am größten ist, verstärkt sich der negative Trend im Hinblick auf die Einstellung zur Mathematik.

Selbstkonzept: Vergleicht man den mittleren Skalenwert für 2023 (2.99) mit denen von 2007 (3.18) und 2019 (3.11), lässt sich für 2023 ein signifikant niedrigerer Wert in Bezug auf das mathematische Selbstkonzept feststellen. Wie bei den Analysen zur Einstellung der Schüler:innen zur Mathematik ist im Durchschnitt nicht mehr ein *hohes*, sondern erstmals ein *mittleres* mathematikbezogenes Selbstkonzept zu konstatieren. Allerdings sollte beachtet werden, dass der mittlere Skalenwert mit 2.99 unmittelbar an der Schwelle zwischen *mittlerem* und *hohem* Selbstkonzept zu verorten ist.

Außerdem zeigt sich, dass der Anteil an Schüler:innen mit *hohem* mathematikbezogenen Selbstkonzept in TIMSS 2023 bei 56 Prozent liegt und damit signifikant geringer ist als in 2019 (64 %) und 2007 (70 %). Des Weiteren zeigen 31 Prozent der Viertklässler:innen ein *mittleres* mathematikbezogenes Selbstkonzept. Dieser Wert ist signifikant höher im Vergleich zu 2007 (21 %) und 2019 (25 %). Die Gruppe der Schüler:innen mit *niedrigem* mathematikbezogenen Selbstkonzept liegt in TIMSS 2023 bei 13 Prozent und unterscheidet sich damit kaum vom Wert für 2019 (12 %), ist jedoch signifikant größer als noch in 2007 (10 %).

Insgesamt zeigen sich im Trendvergleich gesunkene mittlere Skalenwerte, die auf geringere mathematikbezogene Selbstkonzepte der Schüler:innen in Deutsch-

land hindeuten. Gleichzeitig wird der Anteil der Gruppe mit *mittlerem* mathematikbezogenen Selbstkonzept größer, mit *hohem* hingegen kleiner. Allerdings ist die Gruppe der Lernenden mit einem *hohem* mathematikbezogenen Selbstkonzept, die mehr als die Hälfte ausmacht, nach wie vor deutlich am größten.

Abschließend bleibt festzuhalten, dass der Leistungsmittelwert der Schüler:innen in Deutschland am Ende ihrer Grundschulzeit als vergleichsweise gut zu bezeichnen ist und signifikant oberhalb des internationalen Mittelwertes aller Staaten liegt. Die Leistungen der Schüler:innen sind im oberen Mittelfeld der internationalen Rangreihe anzusiedeln. Dabei zeigen die Viertklässler:innen relative Stärke im Bereich *Messen und Geometrie* sowie im Bereich *Reproduzieren*. Für den Inhaltsbereich *Daten* sowie den Anforderungsbereich *Anwenden* zeigen sich hingegen relative Schwächen.

Während im Vergleich der Erhebungen von 2019 und 2023 in fünf Staaten signifikant höhere und in drei Staaten signifikant niedrigere Differenzwerte festzustellen sind, können in elf teilnehmenden Staaten – zu denen auch Deutschland gehört – keine signifikanten Unterschiede konstatiert werden. Angesichts der gestiegenen Anforderungen, die an Grundschule gestellt wurden und werden, sowie der pandemiebedingten Unterrichtsausfälle, ist dieses Resultat zunächst einmal als positiv zu beurteilen, wird beachtet, dass die in TIMSS 2023 getesteten Grundschüler:innen zu Beginn der Schulschließungen in ihrem ersten Schuljahr waren. Ein Abwärtstrend in den Leistungen der Schüler:innen, wie er im IQB-Bildungstrend 2021 (Stanat et al., 2022) für die Grundschule berichtet wurde, ist nicht erkennbar, wenngleich es bedenkenswert ist, dass für die Einstellung zur Mathematik und für das mathematikbezogene Selbstkonzept signifikant gesunkene mittlere Skalenwerte zu berichten sind.

Zudem zeigt sich, dass Deutschland einer von fünf Staaten ist, für die keine signifikanten Leistungsunterschiede zwischen 2007 und 2023 zu verzeichnen sind. Schüler:innen aus zwölf Staaten schneiden in TIMSS 2023 signifikant besser ab als 2007. Zwei weitere Staaten (Hongkong und USA) verschlechtern sich, mit Hongkong dabei ein Land auf hohem Niveau. Im Vergleich zu den meisten Staaten, die an den Erhebungen in 2007, 2019 und 2023 teilgenommen haben, findet also in Deutschland vergleichsweise wenig Weiterentwicklung statt. Ein Aufwärtstrend ist ebenfalls nicht erkennbar.

Darüber hinaus ist auch festzuhalten, dass sich weiterhin etwa ein Viertel der Schüler:innen auf den Kompetenzstufen I und II befindet und nur über unterdurchschnittliche mathematische Kompetenzen verfügt, sodass das Weiterlernen dieser Kinder in der Sekundarstufe I mit erheblichen Schwierigkeiten verbunden sein kann. Hier gelingt es im Trendvergleich vielen Staaten besser, den Anteil der Lernenden auf den Kompetenzstufen I und II zu reduzieren. Im Hinblick auf die verstärkte Sicherung der Basiskompetenzen ist Handlungsbedarf erkennbar, um mehr Schüler:innen das Erreichen der Kompetenzstufe III sowie ein erfolgreiches Weiterlernen im Mathematikunterricht zu ermöglichen. Neben den arithmetischen sind zukünftig auch verstärkt die nicht-arithmetischen Basiskompetenzen in den Blick zu nehmen, um die Anteile der Lernenden auf den Kompetenzstufen I und II zu reduzieren.

Auf der anderen Seite des Leistungsspektrums sind es lediglich 8 Prozent der Schüler:innen in Deutschland, die die höchste Kompetenzstufe erreichen. Wenngleich es hier im Vergleich zum TIMSS 2019 zu einer signifikanten Steigerung kam, handelt es sich im internationalen Vergleich um einen immer noch vergleichsweise niedrigen Wert. Im Hinblick auf die Förderung der leistungsstarken

Lernenden ist Handlungsbedarf erkennbar, damit diese ihre Potenziale im Mathematikunterricht besser entfalten können. Während es diverse Programme zur Förderung der (arithmetischen) Basiskompetenzen gibt, liegen vergleichbare, direkt umsetzbare Programme zur Potenzialförderung im Mathematikunterricht gegenwärtig nicht vor.

Die Befunde von TIMSS 2023 untermauern die Notwendigkeit, Maßnahmen zur Weiterentwicklung des Mathematikunterrichts in der Primarstufe zu intensivieren. Diese bedürfen unbedingt der deutlichen Verbesserung der Rahmenbedingungen, unter denen Unterricht in der Grundschule gegenwärtig stattfindet.

Die Forschungs- und Entwicklungsarbeit zu den Themen Sicherung der arithmetischen sowie der nicht-arithmetischen Basiskompetenzen und Unterrichtsintegrierte Förderung von mathematikbezogenen Potenzialen sollte zudem intensiviert werden und in Konzeptionen und Materialien resultieren, die für die Lehrkräfte direkt handlungswirksam werden können.

Zentral wird es außerdem sein, die Lehrkräfte durch ein wirksames System unterrichtsrelevanter Fortbildung und fachbezogener, durch Multiplikator:innen begleiteter Netzwerkarbeit beständig zu unterstützen, wie es etwa in den Programmen *QuaMath – Unterrichts- und Fortbildungs-Qualität in Mathematik entwickeln* und *Startchancen* umgesetzt wird.

Literatur

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2005). *Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4). Beschluss vom 15.10.2004*. Wolters Kluwer.

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2022). *Bildungsstandards für das Fach Mathematik. Primarbereich*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2022/2022_06_23-Bista-Primarbereich-Mathe.pdf

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen. (Hrsg.). (2008). *Richtlinien und Lehrpläne für die Grundschule in Nordrhein-Westfalen*. Ritterbach.

Mullis, I. V. S., Martin, M. O., & von Davier, M. (2021). *TIMSS 2023 Assessment Frameworks*. TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2023/frameworks/pdf/T23_Frameworks.pdf

Philpot, R., Lindquist, M., Mullis, I. V. S., & Aldrich, C. E. A. (2021). TIMSS 2023 Mathematics Framework. In I. V. S. Mullis, M. O. Martin & M. von Davier (Hrsg.), *TIMSS 2023 Assessment Frameworks* (S. 5–18). TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2023/frameworks/pdf/T23_Frameworks.pdf

Selter, C., Walter, D., Heinze, A., Brandt, J., & Jentsch, A. (2020). Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern im internationalen Vergleich* (S. 57–113). Waxmann. https://doi.org/10.31244/9783830993193

Selter, C., Walter, D., Walther, G., & Wendt, H. (2016). Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 79–136). Waxmann. https://doi.org/10.25656/01:14022

Selter, C., Walther, G., Wessel, J., & Wendt, H. (2012). Mathematische Kompetenz im internationalen Vergleich. Testkonzeption und Ergebnisse. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche*

Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 69–122). Waxmann. https://elibrary.utb.de/doi/epdf/10.31244/9783830978145

Stanat, P., Schipolowski, S., Schneider, R., Sachse, K. A., Weirich, S., & Henschel, S. (Hrsg.). (2022). *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich*. Waxmann. https://www.waxmann.com/index.php?eID=download&buchnr=4606

von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U., & Yin, L. (2024). *TIMSS 2023 International Results in Mathematics and Science.* TIMSS & PIRLS. International Study Center.

Walther, G., Selter, C., Bonsen, M., & Bos, W. (2008). Mathematische Kompetenz im internationalen Vergleich. Testkonzeption und Ergebnisse. In W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther (Hrsg.), *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 49–85). Waxmann.

Yin, L., & Foy, P. (2021). TIMSS 2023 Assessment Design. In I. V. S. Mullis, M. O. Martin & M. von Davier (Hrsg.), *TIMSS 2023 Assessment Frameworks* (S. 71–85). TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2023/frameworks/pdf/T23_Frameworks.pdf

Kapitel 4
Naturwissenschaftliche Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse

Mirjam Steffensky, Luise A. Scholz und Olaf Köller

4.1 Einleitung

Die Naturwissenschaften gehören in der Grundschule in Deutschland zu dem vielperspektivischen Fach Sachunterricht[1]. Der Sachunterricht umfasst neben der naturwissenschaftlichen Perspektive, zu der die Bereiche Biologie (belebte Natur) und Physik/Chemie (unbelebte Natur) gehören, die sozialwissenschaftliche (Politik – Wirtschaft – Soziales), die technische (Technik – Arbeit), die geografische (Räume – Naturgrundlagen – Lebenssituationen) und die historische Perspektive (Zeit – Wandel). Er versteht sich als Perspektiven-integrierendes Fach, in dem naturwissenschaftlich-technische und sozialwissenschaftliche Perspektiven verbunden werden, um so Schüler:innen zu unterstützen, sich ihre natürliche, kulturelle und soziale Umwelt zu erschließen und darin zu handeln (GDSU, 2013). Damit unterscheidet sich die Organisation des naturwissenschaftlichen Unterrichts in der Grundschule von der in den meisten anderen Staaten.

Trotz dieser Spezifika nimmt Deutschland nun zum fünften Mal an der international vergleichenden Studie *Trends in International Mathematics and Science Study* (TIMSS) im Primarbereich teil. Dadurch liegen repräsentative und belastbare Befunde über die Studienzyklen 2007 (Bos, Bonsen et al., 2008), 2011 (Bos, Wendt et al., 2012), 2015 (Wendt et al., 2016) und 2019 (Schwippert et al., 2020) vor, die es ermöglichen, die naturwissenschaftlichen Kompetenzen von Viertklässler:innen in Deutschland international und zwischen den Erhebungsrunden zu vergleichen. Im Folgenden werden zunächst wesentliche Befunde der ersten vier Studienzyklen zusammengefasst (Kleickmann et al., 2012; Steffensky et al., 2016; Steffensky et al., 2020; Wittwer et al., 2008).

1 In Bayern und Thüringen wird das Fach als Heimat- und Sachunterricht bzw. Heimat- und Sachkunde bezeichnet. Der Einfachheit halber wird in diesem Band von dem Fach Sachunterricht gesprochen.

1. *Naturwissenschaftliche Kompetenzen*: Zwischen 2007 und 2015 bestand eine hohe Stabilität in den naturwissenschaftlichen Kompetenzständen der Schüler:innen in Deutschland. In TIMSS 2019 lag der Mittelwert mit 518 Punkten und einer Differenz von 10 Punkten signifikant unter dem Mittelwert (M = 528 Punkte) der vorherigen Studienzyklen. Gleichwohl war der deutsche Mittelwert in 2019 wie in den Jahren davor höher als der Mittelwert aller teilnehmenden Staaten und zudem vergleichbar mit den Werten aus anderen EU-Staaten wie den Niederlanden, Nordirland und Bulgarien. Allerdings schnitt Deutschland 2019, anders als in den Zyklen davor, im Mittel etwas schlechter ab als die teilnehmenden OECD-Staaten. Zwischen den Inhaltsbereichen (*Biologie, Physik/Chemie* und *Geografie*) zeigten sich 2019 zudem signifikante, aber eher kleine Unterschiede in den relativen Leistungen zugunsten der *Biologie* und zuungunsten der *Geografie,* die vorher nicht aufgetreten waren. In den drei kognitiven Anforderungsbereichen *Reproduzieren, Anwenden* und *Problemlösen* zeigten sich dagegen keine relativen Stärken oder Schwächen der Schüler:innen in Deutschland.
2. *Leistungsstarke und -schwache Schüler:innen*: Zwischen 2007 und 2019 waren die Anteile der leistungsstarken und -schwachen Schüler:innen jeweils recht ähnlich. So befand sich knapp ein Viertel der Schüler:innen auf den beiden unteren Kompetenzstufen. Diese Schüler:innen haben weniger gute Voraussetzungen für eine erfolgreiche Teilnahme am naturwissenschaftlichen Unterricht der weiterführenden Schulen, weil ihnen wichtige Grundlagen fehlen. Dieses ungünstige Muster zeigte sich in 2019 etwas stärker als in den Jahren davor. Gleichzeitig war der Anteil der Schüler:innen, die die höchste Kompetenzstufe erreichten, in allen Studienzyklen eher gering (7–10 %).
3. *Einstellungen zum Sachunterricht*: Der überwiegende Teil der Viertklässler:innen verfügte in TIMSS 2019 über ausgesprochen positive Einstellungen zum Sachunterricht und ein hohes sachunterrichtsbezogenes Selbstkonzept eigener Fähigkeiten. Der Vergleich zwischen 2019 und den vorherigen Studienzyklen zeigte eine leichte Abnahme der Mittelwerte. Zudem zeigte sich, dass insbesondere bei Schüler:innen im unteren Leistungsspektrum beide Merkmale niedriger ausgeprägt waren.

Zusammengefasst ergibt sich ein Bild relativ stabil bleibender naturwissenschaftlicher Leistungen und Einstellungen über die Zyklen von 2007 bis 2019. Die leichte, aber signifikante Abnahme der Leistungen in 2019 steht im Einklang mit den Ergebnissen des IQB-Bildungstrends 2016, der ebenfalls leicht zurückgehende Leistungen in den Fächern Deutsch und Mathematik ergab (Stanat et al., 2017).

Durch die Teilnahme von Deutschland an TIMSS 2023 lassen sich Trends in Bezug auf die naturwissenschaftlichen Kompetenzen von Viertklässler:innen über den Zeitraum von 16 Jahren untersuchen. Gleichzeitig geben die aktuellen Daten Hinweise auf mögliche Effekte der COVID-19-Pandemie in einem Nicht-Kernfach. Die in TIMSS 2023 getesteten Kinder waren von den Schulschließungen 2020 und 2021 in den Jahrgangsstufen 1 und 2 betroffen (siehe Kapitel 2 in diesem Band). Für die Fächer Deutsch und Mathematik weist der IQB-Bildungstrend 2021 darauf hin, dass es negative Effekte der Schulschließungen auf die Leistungen am Ende der 4. Jahrgangsstufe gab (Stanat et al., 2022). Allerdings gilt für die in 2021 getestete Kohorte des Bildungstrends, dass Maßnahmen des Corona-Aufholprogramms, das Bund und Länder im Nachgang zur COVID-

19-Pandemie etabliert hatten, noch nicht greifen konnten. Dies ist in TIMSS 2023 anders, da das Programm in den Jahren 2021 und 2022 umgesetzt wurde.

In den folgenden Abschnitten werden grundlegende Aspekte naturwissenschaftlicher Grundbildung in Deutschland (Abschnitt 4.2), die TIMSS-Rahmenkonzeption zur Erfassung naturwissenschaftlicher Kompetenzen (Abschnitt 4.3), sowie der TIMSS-Test zur Erfassung naturwissenschaftlicher Kompetenzen (Abschnitt 4.4) skizziert. Danach werden die Ergebnisse des aktuellen Studienzyklus dargestellt (Abschnitt 4.5) und abschließend zusammengefasst (Abschnitt 4.6). Die Darstellung orientiert sich an den folgenden Fragestellungen unter Einbezug des Vergleichs mit den Ergebnissen aus TIMSS 2007 und 2019:

1. Welches naturwissenschaftliche Kompetenzniveau zeigen Schüler:innen in Deutschland am Ende der vierten Jahrgangsstufe im internationalen Vergleich? (Abschnitt 4.5.1)
2. Wie verteilen sich die Leistungskennwerte in Deutschland im internationalen Vergleich auf die Kompetenzstufen? Wie groß sind die Gruppen der auffällig leistungsstarken und leistungsschwachen Kinder? (Abschnitt 4.5.2)
3. Welche Ergebnisse erzielen Viertklässler:innen in Deutschland in den drei naturwissenschaftlichen Inhaltsbereichen (*Biologie*, *Physik/Chemie* und *Geografie*)? (Abschnitt 4.5.3)
4. Welche Ergebnisse erzielen Viertklässler:innen in Deutschland in den drei kognitiven Anforderungsbereichen *Reproduzieren*, *Anwenden* und *Problemlösen*? (Abschnitt 4.5.4)
5. Welches Niveau zeigen Schüler:innen im Bereich der Nachhaltigkeitskompetenzen am Ende der vierten Jahrgangsstufe im internationalen Vergleich? (Abschnitt 4.5.5)
6. Wie sind die Einstellungen zum Sachunterricht und das sachunterrichtsbezogene Selbstkonzept eigener Fähigkeiten der Schüler:innen in Deutschland ausgeprägt? (Abschnitt 4.5.6)
7. Wie sind die nachhaltigkeitsbezogenen Einstellungen der Schüler:innen in Deutschland ausgeprägt? (Abschnitt 4.5.7)

Auf Ergebnisse zu Unterschieden in der naturwissenschaftlichen Kompetenz in Abhängigkeit von *Geschlecht*, *sozialer Herkunft* und *Migrationshintergrund* wird in den Kapiteln 6 bis 8 dieses Bandes eingegangen.

4.2 Naturwissenschaftliche Grundbildung in der Primarstufe in Deutschland

Wie viele naturwissenschaftliche Lerngelegenheiten Schüler:innen im Sachunterricht tatsächlich erhalten, lässt sich schwer abschätzen. Dies liegt daran, dass die Naturwissenschaften lediglich einen Teil des vielperspektivischen Sachunterrichts ausmachen und zudem die Kontingentstundentafeln den Sachunterricht in vielen Ländern der Bundesrepublik Deutschland mit anderen Fächern im Verbund ausweisen. Unabhängig davon sind die Naturwissenschaften in den Bildungsplänen aller 16 Länder fest verankert. Auch die bisherigen Befragungen der Lehrkräfte im Rahmen von TIMSS 2015 und 2019 weisen darauf hin, dass die Naturwissenschaften eine wichtige Perspektive des Sachunterrichts darstellen (Steffensky et al., 2016; Steffensky et al., 2020). Anzumerken ist allerdings, dass es sich bei diesen Befragungen um möglicherweise positiv verzerrte Selbstberichte handelt.

Die bisherigen Ergebnisse zeigten allerdings auch eine große Streuung zwischen den Angaben.

Die entsprechenden Befragungen im aktuellen Studienzyklus bestätigen die bisherigen Befunde. So geben die Lehrkräfte an, dass rund 40 Prozent der behandelten Themen der naturwissenschaftlichen Perspektive zugeordnet werden können, wobei die Werte sehr stark schwanken (Tabelle 4.1). Die Lehrkräfte geben dabei an, die Themen der belebten und unbelebten Natur im Verhältnis von drei zu zwei zu unterrichten (M = 24.2 bzw. M = 16.7). Neben den vergleichbaren Einschätzungen ist auch die Rangfolge der fünf Perspektiven identisch mit denen aus den vorangegangen Studienzyklen. Berücksichtigt man zudem unterschiedliche Ausbildungshintergründe der Sachunterrichtslehrkräfte und nimmt unterschiedliche Interessen an, sind diese Unterschiede zwischen den Angaben, Ländern, Schulen und individuellen Lehrkräften erwartungskonform.

Tabelle 4.1: Einschätzung der Lehrkräfte zum prozentualen Anteil der Unterrichtsstunden in Klasse drei und vier in den fünf Perspektiven des Perspektivrahmens

Perspektiven	*M*	(*SE*)	*SD*	Min - Max
Sozialwissenschaftliche Perspektive Politik - Wirtschaft - Soziales	16.0	(0.6)	9.3	0 - 50
Naturwissenschaftliche Perspektive Belebte und unbelebte Natur	40.7	(0.9)	14.5	0 - 100
Geografische Perspektive Räume - Naturgrundlagen - Lebenssituationen	20.0	(0.5)	8.3	0 - 60
Historische Perspektive Zeit - Wandel	13.1	(0.4)	6.9	0 - 50
Technische Perspektive Technik - Arbeit	10.5	(0.4)	6.1	0 - 30

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Im Folgenden wird ausschließlich auf die naturwissenschaftlichen Anteile des Sachunterrichts eingegangen. Bei der Interpretation der Daten und dem internationalen Vergleich ist die Besonderheit des naturwissenschaftlichen Unterrichts in Grundschulen in Deutschland allerdings zu berücksichtigen.

Der naturwissenschaftliche (Grundschul-)Unterricht verfolgt multikriteriale Ziele. Diese finden sich z. B. in Beschreibungen des internationalen Bildungskonzepts *Scientific Literacy* (Bybee & McCrae, 2011), auf das sich viele Bildungspläne sowie der Perspektivrahmen[2] der *Gesellschaft für die Didaktik des Sachunterrichts* (GDSU) beziehen (GDSU, 2013). Zentrales Ziel des naturwissenschaftlichen Unterrichts ist die Entwicklung von Kompetenzen, die sich einer-

2 Der Perspektivrahmen hat keinen bindenden Charakter, stellt aber einen wichtigen Orientierungsrahmen für die Weiterentwicklung von Bildungsplänen und didaktischen Materialien für die Grundschule dar, in der es bislang keine Bildungsstandards für die Naturwissenschaften gibt.

seits auf die Inhalte, also z. B. Konzepte, Theorien, Gesetze, und andererseits auf die Denk- und Arbeitsweisen (auch Denk-, Arbeits- und Handlungsweisen), mit denen naturwissenschaftliche Inhalte entwickelt werden, beziehen. Neben diesen inhalts- und prozessbezogenen Kompetenzen sind positiv getönte motivationale Orientierungen und Einstellungen gegenüber den Naturwissenschaften wichtige Ziele des naturwissenschaftlichen Unterrichts. Diese verschiedenen Komponenten stellen im Zusammenspiel die Grundlage für weiteres Lernen und prospektiv für die aktive Mitgestaltung und die informierte Meinungsbildung in einer Lebenswelt dar, die stark von Naturwissenschaften und Technik geprägt ist.

Auch wenn es keine Bildungsstandards für die Naturwissenschaften in der Primarstufe gibt, werden in den verschiedenen Bildungsplänen der Länder der Bundesrepublik Deutschland recht ähnliche Kompetenzerwartungen, in der Regel für das Ende der 2. und 4. Klasse, formuliert, z. B. für das Ende der Klasse 4 „Die Schüler:innen beschreiben das Prinzip der Angepasstheit von Tier- und Pflanzenarten an ihren Lebensraum (u. a. Wald)“ (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2012, S. 188) oder „Schüler:innen wissen um die Angepasstheit von Pflanzen- und Tierarten an ihren Lebensraum Wald“ (Ministerium für Bildung, Wissenschaft und Kultur Mecklenburg-Vorpommern, 2020, S. 21). Dabei sind die inhalts- und prozessbezogenen Kompetenzen in der Regel integriert dargestellt. Dies entspricht auch der regulären Umsetzung im Unterricht. Im Perspektivrahmen des Sachunterrichts sind ebenfalls entsprechende Kompetenzen formuliert, hier wird allerdings zwischen den inhalts- und den prozessbezogenen Kompetenzen unterschieden, z. B.: Die Schüler:innen können

- morphologische Merkmale von Pflanzen (Teile der Pflanze) und Tieren (Körperbau) untersuchen, benennen, beschreiben und vergleichen (verschiedene Pflanzen- und Tierarten),
- Rosten und Verbrennung als Umwandlung von Stoffen beschreiben,
- einfache Ursache-Wirkungs-Zusammenhänge erkennen (z. B. Verdrängung von Wasser durch Luft) und angemessen sprachlich darstellen (GDSU, 2013).

Die eher heuristische Trennung zwischen inhalts- und prozessbezogenem Wissen verdeutlicht, dass die naturwissenschaftlichen Denk- und Arbeitsweisen selbst auch ein Lerngegenstand im naturwissenschaftlichen Unterricht sind. Erste Studien sprechen allerdings dafür, dass Denk- und Arbeitsweisen zwar genutzt, aber selten reflektiert werden (Böschl et al., 2023), was als Voraussetzung für ein Lernen über das Vorgehen in den Naturwissenschaften angenommen wird (Schmerse et al., 2024). Das bedeutet, dass möglicherweise viele Schüler:innen nur sehr wenig Lerngelegenheiten für die Auseinandersetzung mit den naturwissenschaftlichen Denk- und Arbeitsweisen haben.

Die inhaltsbezogenen Kompetenzen beziehen sich auf die Bereiche der belebten Natur (*Biologie*) und der unbelebten Natur (*Physik/Chemie*). Dabei findet man häufig eine Orientierung an zentralen Konzepten wie Materie oder Ökosysteme. International und so auch im Rahmen der TIMSS-Konzeption umfasst *Science* auch Inhalte der *Geografie*, wobei es dabei um naturgeografische, z. B. Wetter, Jahreszeiten oder den Mondlauf, und nicht um humangeografische Inhalte geht.

Im Zusammenhang mit naturwissenschaftlichen Kompetenzen stehen Kompetenzen, die sich auf durch menschliche Aktivitäten verursachte globale Veränderungen der Umwelt (z. B. Klimawandel oder Verlust von Biodiversität), deren Auswirkungen (z. B. Abschmelzen der Pole) sowie Maßnahmen, um diesen Pro-

blemen entgegenzuwirken (z. B. alternative Energien), beziehen. Allerdings sind diese Nachhaltigkeitskompetenzen deutlich breiter als rein naturwissenschaftliche Kompetenzen, weil sie zudem politische, soziale und auch kulturelle Kompetenzen und deren Zusammenspiel umfassen (Joint Research Centre, 2022). Sie beziehen demnach mehrere Disziplinen mit ein und sind gleichzeitig interdisziplinär angelegt. Nachhaltigkeitskompetenzen werden verstanden als Kompetenzen, die zu einer nachhaltigen Entwicklung beitragen können (*Sustainable Development Goals*; UN DESA, 2021). Ähnliche Ansätze sind z. B. *Global Competencies* (Costa et al., 2024; OECD, 2019) oder *Climate Literacy* (Johnston, 2020; NOAA, 2014; U.S. Global Change Research Program, 2009). Thematisch verortet sind diese Kompetenzen im Querschnittsthema *Bildung für Nachhaltige Entwicklung*, das für den Sachunterricht, aber auch für viele andere Fächer wichtig ist. Dem naturwissenschaftlichen Unterricht kommt dabei eine besondere Rolle zu, da für das Verständnis vieler globaler Herausforderungen naturwissenschaftliche Kompetenzen grundlegend sind. Dies spiegelt sich auch in manchen Weiterentwicklungen des Konzepts *Scientific Literacy* wider, die Nachhaltigkeitsaspekte einschließen (Owens & Sadler, 2024).

4.3 TIMSS-Rahmenkonzeption zur Erfassung der naturwissenschaftlichen Kompetenzen

Ziel der TIMS-Studie ist es, naturwissenschaftliche Kompetenzen einerseits breit zu erfassen, andererseits aber Analysen von Teilkompetenzen zu ermöglichen, die ein differenziertes Bild eben dieser zulassen. Grundlegend für die Einordnung der Testaufgaben ist die TIMSS-Rahmenkonzeption (Mullis et al., 2021). Diese umfasst genau wie in den letzten TIMSS-Zyklen sowohl die für die Studie relevanten naturwissenschaftlichen Inhaltsbereiche *Biologie, Physik/Chemie* und *Geografie* als auch die zentralen Denkprozesse, die dem Lösen der Aufgaben zugrunde liegen. Letztere werden als kognitive Anforderungsbereiche bezeichnet, wobei zwischen dem *Reproduzieren* von Wissen, dem *Anwenden* von Wissen sowie dem *Problemlösen* differenziert wird. Die prozessbezogenen Kompetenzen (naturwissenschaftliche Denk- und Arbeitsweisen) werden bei TIMSS nicht separat erhoben, sondern sind in die Inhaltsbereiche integriert.

Da in TIMSS das Wissen über curriculare Inhalte im Vordergrund steht, erfordert die Rahmenkonzeption zur Erfassung naturwissenschaftlicher Kompetenzen die Abdeckung zentraler Bereiche der nationalen Curricula aller teilnehmenden Staaten. Um längerfristige Trends in den Kompetenzständen im nationalen und internationalen Vergleich untersuchen zu können, ist zudem die Vergleichbarkeit der Studienzyklen notwendig. Entsprechend gibt es große Überschneidungen zu den zentralen Bereichen der vorangegangenen Rahmenkonzeption. Gleichzeitig erfordern Änderungen in Curricula immer wieder kleinere Anpassungen in der Rahmenkonzeption, um eine möglichst hohe Passung der Testaufgaben zu den Curricula herzustellen. Zudem ermöglicht die computerbasierte Testung in TIMSS 2019 und 2023 die stärkere Berücksichtigung von Aufgaben, mit denen das Verständnis und die Anwendung naturwissenschaftlicher Denk- und Arbeitsweisen erhoben werden können, die in vielen internationalen wie nationalen Curricula eine zunehmende Bedeutung erfahren. Die computerbasierte Testung eignet sich hierfür, weil sie aufeinander aufbauende Testaufgaben ermöglicht, in denen verschiedene Schritte von naturwissenschaftlichen Untersuchungen simuliert werden können. Das bedeutet, dass die Rahmenkonzeption und Aufgabenentwick-

lung einerseits anschlussfähig an die vorangegangenen Rahmenkonzeptionen sein und andererseits aber auch neuere Entwicklungen abbilden müssen. Im Folgenden wird ein detaillierter Überblick über die Rahmenkonzeption gegeben.

4.3.1 Naturwissenschaftliche Inhaltsbereiche

Die drei Inhaltsbereiche *Biologie, Physik/Chemie* und *Geografie* sind bei TIMSS 2023 durch fünf bzw. drei Themengebiete beschrieben, die nochmal einzelne Themen umfassen (Tabelle 4.2). Zur Konkretisierung der Erwartungen gibt es zu allen Themen Kompetenzerwartungen, die hier aus Platzgründen nur exemplarisch dargestellt sind. Ausführliche Informationen finden sich bei Mullis et al. (2021).

Der Inhaltsbereich *Biologie* umfasst fünf Themengebiete: 1. Kennzeichen des Lebendigen und Lebensvorgänge von Organismen, 2. Lebenszyklen, Fortpflanzung und Vererbung, 3. Organismen, Umwelt und ihre Interaktion, 4. Ökosysteme und 5. Gesundheit. Von den Schüler:innen wird grundlegendes Wissen über Lebewesen, ihre Merkmale und ihre Entwicklung sowie ihre Interaktionen mit anderen Lebewesen und ihrer Umgebung erwartet. Das Themengebiet Gesundheit fokussiert auf alltagspraktisches Wissen über die Verbreitung von und den Schutz vor ansteckenden Krankheiten sowie auf gesunde Ernährungs- und Verhaltensweisen.

Der Inhaltsbereich *Physik/Chemie* umfasst die drei Themengebiete: 1. Stoffe (Klassifizierung von Stoffen, Stoffeigenschaften und Veränderungen von Stoffen), 2. Energieformen und Energieumwandlung sowie 3. Kräfte und Bewegungen. Von den Schüler:innen wird grundlegendes Wissen über Stoffklassen, die drei Aggregatzustände und einfache physikalische und chemische Veränderungen von Stoffen erwartet. Außerdem sollen sie gängige Energieformen und -quellen, z. B. Sonne, Kohle, Wasserkraft, und ihre praktische Anwendung kennen. Zudem sollen sie ein erstes phänomenbasiertes Wissen über Licht, Schall, Elektrizität und Magnetismus haben. Auch das dritte Themengebiet Kräfte und Bewegungen zielt ‚lediglich' auf ein phänomenbasiertes Verständnis von Kräften und Bewegungen ab, z. B. Änderungen von Bewegungen durch Ziehen oder Schieben zu erklären. Für diesen Bereich sind auch erste Kenntnisse über einfache Maschinen, mit denen sich Kraft sparen lässt, relevant.

Der Inhaltsbereich *Geografie* umfasst die drei Themengebiete: 1. Charakteristische Landschaftsmerkmale, Ressourcen und Erdgeschichte, 2. Wetter und Klima sowie 3. die Erde im Sonnensystem. Von den Schüler:innen wird grundlegendes Wissen über Merkmale der Erde und langfristige Veränderungen von Landschaften erwartet. Außerdem sollen sie Wettererscheinungen über Tage, Jahreszeiten und verschiedene Orte hinweg erläutern und die Entstehung von Tag und Nacht oder Jahreszeiten mit der Erdbewegung erklären können.

Zusammenfassend zeigt sich eine große Überlappung zwischen den Inhalten und dazugehörigen Themenbereichen aus TIMSS 2019 und den Empfehlungen bzw. Vorgaben aus dem Perspektivrahmen der GDSU sowie den Bildungsplänen für den Sachunterricht.

Tabelle 4.2: Themengebiete der einzelnen naturwissenschaftlichen Inhaltsbereiche

Inhaltsbereiche	Teilgebiete
Biologie	**1. Kennzeichen des Lebendigen und Lebensvorgänge von Organismen** - Unterschiede zwischen Lebewesen und unbelebten Dingen und was Lebewesen benötigen, um zu leben (z.B. Luft, Nahrung, Wasser, Lebensraum) - Baumerkmale und Verhalten von Tier- und Pflanzengruppen (z.B. Insekten, Vögel, Säugetiere, Reptilien, Blütenpflanzen) - Zusammenhang zwischen Struktur und Funktion (z.B. Knochen, Wurzeln) **2. Lebenszyklen, Fortpflanzung und Vererbung** - Lebenszyklen bekannter Tiere und Pflanzen (z.B. Schmetterlinge) - Vererbung und Fortpflanzungsstrategien (z.B. Samenpflanzen, Brutpflege) **3. Organismen, Umwelt und ihre Interaktion** - Angepasstheit von Lebewesen an ihren Lebensraum durch Körperbau und Verhalten (z.B. tiefe Wurzeln, Färbung) - Reaktionen von Lebewesen auf Bedingungen ihrer Umwelt (z.B. Licht, Temperatur, Gefahr) - Einfluss des Menschen auf die Umwelt (z.B. Verschmutzung von Wasser und Luft) **4. Ökosysteme** - Diversität in den Lebensräumen und bekannte Ökosysteme (z.B. Wald, Teich) - Einfache Nahrungsketten - Wettbewerb in Ökosystemen **5. Gesundheit** - Symptome von Krankheiten, Übertragung ansteckender Krankheiten und die Möglichkeiten, sich vor Ansteckung zu schützen - Maßnahmen der Gesunderhaltung (z.B. Ernährungsweisen, Hygiene)
Physik/Chemie	**1. Stoffe (Klassifizierung von Stoffen, Stoffeigenschaften und Veränderungen von Stoffen)** - Aggregatzustände und ihre spezifischen Merkmale (z.B. Form, Volumen) - Physikalische Eigenschaften von Stoffen und Stoffklassen (z.B. Leitfähigkeit) - Magnetische Anziehung und Abstoßung - Physikalische Veränderungen in Alltagssituationen (z.B. beim Schmelzen, Verdunsten, Lösen) - Chemische Veränderungen in Alltagssituationen (z.B. beim Rosten, Verbrennen, Kochen) **2. Energieformen und Energieumwandlung** - Energiequellen (z.B. Sonne, Wind, Kohle) und deren Nutzung (z.B. für Wärme, Licht) - Licht und Schall in Alltagssituationen (z.B. Lichtquellen, Schatten) - Wärmeübertragung - Elektrizität und einfache Stromkreise **3. Kräfte und Bewegungen** - Bekannte Kräfte (z.B. Erdanziehung, Reibung, Ziehen, Schieben) und deren Einfluss auf die Bewegung von Objekten - Einfache Maschinen, die den Kraftaufwand verringern (z.B. Hebel oder Rampen)
Geografie	**1. Charakteristische Landschaftsmerkmale, Ressourcen und Erdgeschichte** - Charakteristische Landschaftsmerkmale (z.B. Vorkommen von Salz- und Süßwasser) - Natürliche Ressourcen (z.B. Boden, Wasser, Wind) - Erdgeschichte (z.B. Landschaftsbildung, Fossilien) **2. Wetter und Klima** - Wetter und Klima auf der Erde (z.B. Wasserphänomene, Wasserkreislauf, Wettererscheinungen im Laufe eines Tages, Jahres und an verschiedenen Orten) **3. Die Erde im Sonnensystem** - Planeten des Sonnensystems und ihre Bewegungen (z.B. Mondphasen) - Auswirkungen der Bewegungen der Erde (z.B. Tag und Nacht, Jahreszeiten)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

4.3.2 Kognitive Anforderungsbereiche

Den drei kognitiven Anforderungsbereichen (*Reproduzieren* von Wissen, *Anwenden* von Wissen sowie *Problemlösen*) lassen sich verschiedene Denkprozesse zuordnen, die in Tabelle 4.3 dargestellt sind.

Tabelle 4.3: Denkprozesse der kognitiven Anforderungsbereiche

Reproduzieren	Anwenden	Problemlösen
- Erinnern/Wiedererkennen - Beschreiben - Angemessene Beispiele finden	- Vergleichen, Unterschiede feststellen und Klassifizieren - Zusammenhänge herstellen - Anwenden von Modellen - Informationen interpretieren - Erklären	- Fragen formulieren, Hypothesen aufstellen und Vorhersagen treffen - Experimente planen und Modelle entwickeln - Einschätzen von Erklärungen, Modellen - Schlussfolgerungen ziehen - Synthetisieren - Generalisieren - Begründen

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Im Anforderungsbereich *Reproduzieren* geht es darum, Fakten zu naturwissenschaftlichen Sachverhalten abzurufen, Sachverhalte und Beziehungen, z. B. zwischen Lebewesen oder Materialien, zu beschreiben und angemessene Beispiele zu nennen. Auch die korrekte Verwendung von Fachbegriffen und Kenntnisse über naturwissenschaftliche Instrumente gehören dazu. Die hier geforderten Wissenselemente legen den Grundstein für die beiden anderen Anforderungsbereiche. Beim *Anwenden* liegt der Schwerpunkt darauf, naturwissenschaftliche Fakten, Zusammenhänge, Konzepte und Methoden in verschiedenen Kontexten anzuwenden. In diesen Anforderungsbereich fällt auch das Nutzen von Repräsentationen wie Modellen oder Diagrammen, um einen Sachverhalt zu erläutern. Aufgaben im Bereich *Problemlösen* erfordern, dass Schüler:innen beim Analysieren von Daten oder Informationen oder beim Schlussfolgern argumentieren. Dieser kognitive Anforderungsbereich umfasst anspruchsvolle naturwissenschaftliche Aktivitäten wie das Planen von Experimenten oder das Entwickeln von Modellen.

Außer den Inhaltsbereichen (Tabelle 4.2) und den kognitiven Anforderungsbereichen (Tabelle 4.3) werden in der TIMSS-Rahmenkonzeption zusätzlich fünf zentrale naturwissenschaftliche Denk- und Arbeitsweisen (*science practices*) aufgeführt:

- Überprüfbare Fragen entwickeln, die auf Beobachtungen oder Theorien basieren
- Systematische Versuche planen, mit denen z. B. eine Vermutung belegt oder widerlegt werden kann, um Evidenz zu generieren
- Mit Daten arbeiten, was das Darstellen und Beschreiben von Daten sowie Erkennen und Interpretieren von Mustern und Zusammenhängen in den Daten umfasst

- Forschungsfragen beantworten, indem evidenzbasiert auf die Hypothesen Bezug genommen wird, wobei auch Limitationen von Untersuchungen, Evidenz oder Schlussfolgerungen reflektiert werden
- Evidenzbasiert argumentieren, um Erklärungen und Schlussfolgerungen zu entwickeln, zu begründen und auf neue Situationen anzuwenden

Diese Denk- und Arbeitsweisen stellen keine eigene Dimension in der Rahmenkonzeption dar, sondern sind in die Inhaltsbereiche und die kognitiven Anforderungen integriert. Sie werden also in den meisten Testaufgaben nicht isoliert erfasst, sondern durch Aufgaben im Kontext der Inhaltsgebiete, die verschiedene kognitive Aktivitäten erfordern. Der Schwerpunkt liegt dabei insbesondere auf dem Arbeiten mit Daten und dem evidenzbasierten Argumentieren. Hier lassen sich viele Überschneidungen zu den zwei Bereichen *Anwenden* und *Problemlösen* feststellen. Beispielsweise ist das Vergleichen und Feststellen von Unterschieden (Anforderungsbereich *Anwenden*) notwendig, um mit Daten zu arbeiten. Im Rahmen der computerbasierten Testung in TIMSS 2019 wurden spezifische Erkenntnisgewinnungsaufgaben erprobt, die in TIMSS 2023 weiter ausgebaut wurden. In diesen Aufgaben werden Alltags- oder laborspezifische Situationen simuliert, in denen die Schüler:innen neben dem Inhaltswissen auch Denk- und Arbeitsweisen anwenden müssen, um geeignete Versuche durchzuführen oder ein Problem zu lösen. Diese Aufgaben sind den beiden kognitiven Anforderungsbereichen *Anwenden* und *Problemlösen* zugeordnet und werden international als *Problem Solving and Inquiry Tasks* (PSI; Mullis et al., 2021) bezeichnet.

4.3.3 Beispielaufgaben

Um die Kernbereiche der TIMSS-Rahmenkonzeption zu visualisieren, sind in Abbildung 4.1 Beispielaufgaben für die drei Inhalts- und die drei Anforderungsbereiche dargestellt. Implizit sind darin auch die naturwissenschaftlichen Denk- und Arbeitsweisen enthalten. Anders als die Inhaltsbereiche sind die kognitiven Anforderungsbereiche nicht immer eindeutig voneinander abgrenzbar. So sind beispielsweise das *Anwenden* und das *Problemlösen* auch aus theoretischer Perspektive ähnliche kognitive Prozesse. Zudem hängen die kognitiven Prozesse, die zur Bearbeitung einer Aufgabe notwendig sind, stark vom Vorwissen der Schüler:innen ab. So kann eine Aufgabe für manche Schüler:innen aufgrund von mangelnder Erfahrung das *Anwenden* von Wissen in einem neuen Beispiel oder Kontext erfordern, während es für andere lediglich das *Reproduzieren* von Fakten bedeutet. Der Anforderungsbereich kann also nicht vollständig von der bearbeitenden Person getrennt betrachtet werden.

Abbildung 4.1: Beispielaufgaben für die Inhaltsbereiche und kognitive Anforderungsbereiche

	Reproduzieren	Anwenden	Problemlösen
Biologie	Welches Tier-Paar zeigt einen Räuber und seine Beute? A: Räuber: Schildkröte; Beute: Schmetterling B: Räuber: Gans; Beute: Schlange C: Räuber: Habicht; Beute: Maus D: Räuber: Katze; Beute: Hirsch SE71076 (int. .71/nat. .80) In dieser Aufgabe muss aus vier Beispielen eines wiedererkannt werden, das ein Räuber-Beute-Paar darstellt. Grundlegend ist hier ein erstes alltagsnahes Wissen über Nahrungsketten von Tieren.	Johann untersucht die Wirkung von unterschiedlichen Mengen Licht auf zwei gleiche Pflanzen. Er pflanzt die Pflanzen in gleiche Töpfe mit derselben Art Erde und derselben Menge Wasser. Johann stellt Pflanze A in die Nähe vom Fenster und Pflanze B in einen Schrank, dessen Tür geschlossen ist. Nach zwei Wochen sehen die Pflanzen so aus: Pflanze A Pflanze B Warum ist Pflanze B, die im Schrank stand, nicht so gesund wie Pflanze A, die beim Fenster stand? **Weil Pflanze B im Schrank kein Licht hatte** SE71005 (int. .71/nat. .77) Zum Lösen dieser Aufgabe muss aus einer Beobachtung der Unterschied zwischen beiden Bedingungen (mit und ohne Licht) identifiziert und dann der Zusammenhang zwischen Licht und Pflanzenwachstum hergestellt werden. Das Lösen der Aufgabe setzt ein Wissen über Pflanzen als Lebewesen und ihr Wachstum voraus.	Frösche legen Hunderte von Eiern auf einmal. Elefanten bekommen nur ein Kalb auf einmal. Warum hat ein einzelnes Elefantenkalb bessere Chancen zu überleben als ein einzelnes Frosch-Ei? Nenne einen Grund. **Weil sich der Elefant um sein Kind kümmert und der Frosch nicht** SE71043 (int. .23/nat. .29) Zum Lösen dieser Aufgabe müssen aus zwei sehr unterschiedlichen Beobachtungen – viele Froscheier versus ein Elefantenkalb – Schlussfolgerungen gezogen und diese erläutert werden. Das Lösen der Aufgabe erfordert ein Verständnis unterschiedlicher Lebensweisen von Arten.
Physik/Chemie	Max hat verschiedene Haufen mit Dingen, die er aufsammeln soll. Welche Dinge kann er mit einem Magnet aufsammeln? Klicke alle Dinge an, die Max mit einem Magnet aufsammeln kann. Glasmurmeln, Eisenspäne, Gummibänder, Stahlbüroklammern, Holzzahnstocher, Scheibenmagnete SE71114 (int. .65/nat. .76) In dieser Aufgabe müssen die Dinge wiedererkannt werden, die von einem Magneten angezogen werden. Die Aufgabe setzt Wissen darüber voraus, dass nur bestimmte Materialien mit Magnenten wechselwirken. Während das Identifizieren der Büroklammern und Magneten auf Alltagserfahrungen beruhen kann, erfordert das Erkennen der Eisenspäne die Kenntnis über den Fachbegriff Eisen.	Bei welchem Stromkreis wird die Glühbirne leuchten? A: Metalllöffel B: Holzlöffel C D: Glasbecher SE71129 (int. .79/nat. .86) In dieser Aufgabe müssen die vier„eingefügten" Materialien in den Stromkreisen verglichen werden (Metall, Holz, Luft, Glas) und ein Zusammenhang zwischen der elektrischen Leitfähigkeit dieser Materialien und dem geschlossenen Stromkreis hergestellt werden. Für die Lösung der Aufgabe ist u. a. Wissen über Materialien und ihre Eigenschaften notwendig.	Karl untersucht, wie man die gleiche Menge Zucker schnell in Wasser auflösen kann. Er bereitet Versuche vor. Klicke bei jedem Versuch das Glas an, in dem sich der Zucker schneller auflösen wird. **Versuch 1** unterschiedliche Temperatur: 25°C (A), 30°C (B) **Versuch 2** eins umgerührt: 25°C (A), 25°C (B) **Versuch 3** unterschiedlich große Würfel: 25°C (A), 25°C (B) Warum ist es wichtig, dass in jedem Glas die gleiche Menge Wasser ist? **Weil man nicht weiß, ob die Wassermenge nicht auch wichtig für das Auflösen ist** SO61142 (int. .24/nat. .22) Im zweiten Teil dieser Doppelaufgabe* muss erläutert werden, warum die Wassermenge konstant gehalten wird. Dies setzt ein erstes Wissen über die Denk- und Arbeitsweise ‚(kontrolliertes) Experimentieren' voraus. Als richtig werden kontextgebundene Antworten (siehe Beispiel) oder bereits generalisierte Antworten wie „um einen fairen Test zu machen" gewertet.
Geografie	Tobias wohnt in der Nähe der Wüste. Er bemerkt, dass die Sanddünen nicht immer an derselben Stelle sind. Die Bilder zeigen die Position von Sanddünen an zwei verschiedenen Tagen. Tag 1 Tag 2 Wie bewegen sich die Sanddünen von der Stelle? A Der Wind bewegt die Sanddünen. B Der Regen bewegt die Sanddünen. C Wüstentiere bewegen die Sanddünen. D Erdbeben bewegen die Sanddünen. SE71218 (int. .80/nat. .86) In dieser Aufgabe muss erkannt werden, welcher der aufgeführten Faktoren sich auf die Veränderung von Dünen auswirken kann. Dafür wird ein alltagsnahes Wissen über Landschaftsmerkmale und ihre Veränderungen vorausgesetzt.	Die Firma baut Möbel mit Holz aus einem Wald. Wie kann die Firma am besten dafür sorgen, dass sie auch weiterhin den Wald nutzen kann? A Sie sollte als Ersatz für die Bäume Gras und Blumen pflanzen. B Sie sollte die Tiere aus dem Wald in eine neue Umgebung bringen. C Sie sollte die Möbelfabrik weit weg vom Wald bauen. D Sie sollte als Ersatz für die Bäume junge Bäume anpflanzen. SE71214 (int. .60/nat. .61) In der Aufgabe muss aus den gegebenen Informationen interpretiert werden, dass die Firma Holz aus einem Wald entnimmt und dies zu (ungünstigen) Veränderungen in dem Wald führt. Es muss der Zusammenhang zu sinnvollen Maßnahmen hergestellt werden, um den Wald zu erhalten. Diese Aufgabe wird auch für die Skala der Nachhaltigkeitskompetenzen verwendet (siehe Abschnitt 4.3.4).	Ein anderer Bauer wohnt auf einer tropischen Insel im Ozean und will Bananen anbauen. Unten siehst du eine Abbildung der Insel. Jährliche Regenmenge auf der Insel Gebiet 1, Gebiet 2, Gebiet 4, Gebiet 3 Legende: 100 cm oder weniger; 101-140 cm; 141-180 cm; Mehr als 180 cm Sieh dir die Abbildung der Insel und die Information in der Abbildung aus dem Teil A dieser Frage an. In welchem Gebiet sollte der Bauer Bananen anbauen? A Gebiet 1 B Gebiet 2 C Gebiet 3 D Gebiet 4 SO61115_2 (int. .47/nat. .46) Dies ist der zweite Teil einer Aufgabe. Im ersten Teil erhalten die Schüler:innen Informationen über den Wasserbedarf von Bananenpflanzen. In dieser Aufgabe* müssen die Schüler:innen auf Basis dieser Informationen die (nicht typische) Abbildung interpretieren und zu einer Entscheidung kommen.

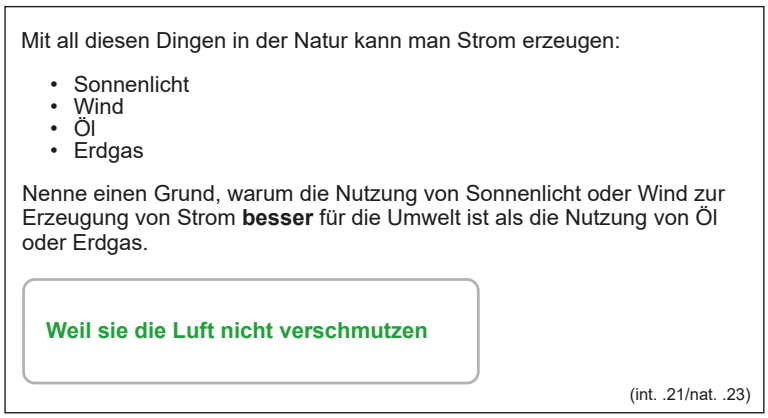
*= Aufgrund der begrenzten Aufgabenfreigabe sind Beispielaufgaben aus TIMSS 2019 (Steffensky et al., 2020) abgebildet.
Die Werte in Klammern geben die relativen internationalen und nationalen Lösungshäufigkeiten an.
Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: https://www.iea.nl/data-tools/permission-requests

 © TIMSS 2023

4.3.4 Nachhaltigkeitskompetenzen

Schon in den vergangenen Studienzyklen umfassten die TIMSS-Aufgaben auch solche, in denen es um den menschlichen Einfluss auf die Umwelt, natürliche Ressourcen sowie umweltverträgliche Handlungen ging. Dieser Bereich wurde stetig ausgebaut und wird erstmalig in TIMSS 2023 zusätzlich als eigene Skala berichtet. Hierfür wurden 44 Aufgaben aus dem bestehenden Itempool ausgewählt, die auch zu der Skala naturwissenschaftliche Kompetenz gehören. Diese Aufgaben verteilen sich folgendermaßen auf die drei Inhaltsbereiche: 25 aus dem Bereich *Biologie*, zwei aus *Physik/Chemie* und 17 aus *Geografie*. In Abbildung 4.1 ist die Beispielaufgabe aus dem Inhaltsbereich *Geografie* (kognitiver Anforderungsbereich *Anwenden*) eine der Aufgaben, die zu dieser Skala gehört. Hier müssen die Schüler:innen ein basales Verständnis davon haben, dass natürliche Ressourcen nicht unbegrenzt zur Verfügung stehen. Ein weiteres Beispiel einer Aufgabe der Nachhaltigkeitsskala aus dem Bereich *Biologie* findet sich in Abbildung 4.2. In dieser Aufgabe müssen die Schüler:innen einen Grund nennen, warum Sonnenlicht und Wind umweltverträglichere Energieträger sind als Öl und Erdgas. In beiden Fällen zeigen sich keine relevanten Unterschiede zwischen der internationalen und der nationalen Lösungshäufigkeit.

Abbildung 4.2: Aufgabe aus dem Inhaltsbereich *Biologie*, die zur Skala Nachhaltigkeitskompetenzen gehört



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

Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: https://www.iea.nl/data-tools/permission-requests

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

4.4 TIMSS-Test

In diesem Abschnitt werden der Aufbau des Tests, der zur Erfassung naturwissenschaftlicher Kompetenzen in TIMSS 2023 eingesetzt wurde (Abschnitt 4.4.1), die Ergebnisse aus den Analysen zur curricularen Validität der Testaufgaben (Abschnitt 4.4.2) sowie die Kompetenzstufen, die zur Illustration der Kompetenzausprägungen der Schüler:innen genutzt werden (Abschnitt 4.4.3), beschrieben.

4.4.1 Testentwicklung und -aufbau

Die TIMSS-Testaufgaben für die Naturwissenschaften wurden in einem kooperativen Prozess entwickelt, an dem neben dem *TIMSS & PIRLS International Study Center* am Boston College und den nationalen Studienleitungen auch Expert:innen aus der Naturwissenschaftsdidaktik beteiligt waren. Die in einem schrittweisen Prozess überarbeiteten Aufgaben wurden unter Berücksichtigung fach- und bildungssprachlicher Spezifika der jeweiligen Sprachen übersetzt und anschließend in nationalen Feldtests erprobt, bevor dann die finalen Testaufgaben für die Haupterhebung ausgewählt wurden. Ziel dieses aufwendigen Prozesses war es, Aufgaben zu erhalten, die sowohl in psychometrischer Hinsicht als auch für die Schüler:innen des jeweiligen Staates geeignet sind, naturwissenschaftliche Kompetenzen zu erfassen.

Um die in der Rahmenkonzeption verankerten Inhaltsbereiche *Biologie, Physik/Chemie* und *Geografie* und kognitiven Anforderungsbereiche *Reproduzieren*, *Anwenden* und *Problemlösen* breit abzudecken, ist eine große Zahl an Aufgaben erforderlich. Der finale Aufgabenpool im Bereich Naturwissenschaften bei TIMSS 2023 umfasst 173 Aufgaben für die 4. Klasse. Von diesen sind 78 *Selected Response*-Aufgaben, bei denen die Schüler:innen eine oder mehrere Antworten wählen müssen. Die übrigen 95 Testaufgaben sind so genannte *Constructed Response-Items*, die eine kurze schriftliche Antwort erfordern, oder bei denen die Schüler:innen Begriffe oder Bilder via drag-and-drop zu- oder anordnen müssen. Zwölf Aufgaben gehören zu dem neu entwickelten Format der Erkenntnisgewinnungsaufgaben. Für die Trendschätzungen wurden 95 Aufgaben (55 %) aus den vorherigen Studienzyklen übernommen. Die verbleibenden 78 Aufgaben (45 %) wurden für TIMSS 2023 neu entwickelt. Von den 173 Aufgaben erfassen 44 Items die Nachhaltigkeitskompetenzen.

Wie zuvor beschrieben lassen sich entsprechend der TIMSS-Rahmenkonzeption die einzelnen Aufgaben jeweils einem Inhaltsbereich und einem kognitiven Anforderungsbereich zuordnen. Dadurch können zusätzlich zur Erfassung der Kompetenz auf der Gesamtskala Naturwissenschaften auch Kompetenzen im Hinblick auf die einzelnen Inhalts- und Anforderungsbereiche untersucht werden. In Tabelle 4.4 ist die Verteilung der Aufgaben auf die Inhaltsbereiche *Biologie, Physik/Chemie* und *Geografie* dargestellt.

Tabelle 4.4: Verteilung der Testaufgaben auf die naturwissenschaftlichen Inhaltsbereiche

Naturwissenschaftliche Inhaltsbereiche	2007		2019		2023	
	n	%	*n*	%	*n*	%
Biologie	74	43	73	43	79	46
Physik/Chemie	64	37	61	36	61	35
Geografie	36	21	35	21	33	19
Gesamt	174	100	169	100	173	100

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

Die drei Bereiche sind nicht gleichmäßig berücksichtigt. Es lassen sich 46 Prozent der Aufgaben dem Inhaltsbereich *Biologie* zuordnen, 35 Prozent dem Bereich *Physik/Chemie* und lediglich 19 Prozent dem Inhaltsbereich *Geografie*. Diese Schwerpunktsetzung auf *Biologie* und – etwas weniger ausgeprägt – auf *Physik/Chemie* fand sich auch in den vorangegangenen Studienzyklen. Sie basiert auf der Rahmenkonzeption von TIMSS, die sich an den Curricula und Standards der teilnehmenden Staaten orientiert (Mullis et al., 2021). In diesem Themenkanon sind geografische Inhalte nicht so stark repräsentiert wie biologische oder physikalisch-chemische. Dies entspricht auch den Bildungsplänen in Deutschland. So deuten Analysen von Schulbüchern aus Deutschland zwar darauf hin, dass im Sachunterricht geografische Inhalte im Vergleich zu anderen Inhalten vergleichsweise stark repräsentiert sind (Blaseio, 2009). Allerdings liegt der Schwerpunkt vor allem auf regionalgeografischen und nicht auf naturgeografischen Themen.

In Tabelle 4.5 ist die Verteilung der Aufgaben auf die kognitiven Anforderungsbereiche *Reproduzieren, Anwenden* und *Problemlösen* dargestellt. Wie in den vergangenen Studienzyklen wird bei 40 Prozent der Aufgaben Faktenwissen abgefragt, während ca. 60 Prozent der Aufgaben kognitiv anspruchsvollere Anforderungen beinhalten. Trotz dieser angenommenen Hierarchie in den kognitiven Anforderungen bedeutet dies nicht, dass die Aufgaben im Anforderungsbereich *Reproduzieren* immer leichter sind als die in den anderen Anforderungsbereichen.

Tabelle 4.5: Verteilung der Testaufgaben auf die Anforderungsbereiche

Kognitive Anforderungsbereiche	2007		2019		2023	
	n	%	*n*	%	*n*	%
Reproduzieren	77	44	69	41	69	40
Anwenden	63	36	64	38	72	42
Problemlösen	34	20	36	21	32	19
Gesamt	174	100	169	100	173	100

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

4.4.2 Curriculare Validität

Der Themenkanon, der den TIMSS-Aufgaben zugrunde liegt, und auch die damit einhergehenden inhaltlichen Schwerpunktsetzungen orientieren sich an den Curricula und Standards der Staaten, die an TIMSS teilnehmen. Ähnlich wie in Mathematik gibt es in den Naturwissenschaften einen relativ großen nationalen wie internationalen Konsens über zentrale Konzepte und Themengebiete wie Kennzeichen von Lebewesen, Struktur und Funktion, Ökosysteme, Stoffe und ihre Eigenschaften, Energie, Kräfte oder Klima sowie Landschaft und ihre Merkmale, die sich z. B. an den *Basiskonzepten* (KMK, 2004, 2024), *Big ideas* (Harlen, 2015) oder *Core Concepts* und *Crosscutting Concepts* (National Research Council, 2012) ablesen lassen. Gleichwohl ist nicht jede einzelne

Facette eines Themenbereiches in allen Curricula der teilnehmenden Staaten abgebildet. Unterschiede im Umfang und in Gewichtungen ergeben sich z. B. aus der unterschiedlichen Organisation des naturwissenschaftlichen Unterrichts in der Grundschule oder aus der unterschiedlichen Verteilung von Inhalten auf den naturwissenschaftlichen Unterricht der Grundschule und der weiterführenden Schulen.

Um die curriculare Validität der Aufgaben, also die Passung der Aufgaben zu den curricularen Vorgaben in Deutschland, zu prüfen, wurden die Aufgaben wie in den bisherigen Studienzyklen durch verschiedene Expert:innen mit den Vorgaben in einer *Test-Curriculum Matching Analysis* (TCMA) abgeglichen. Da es anders als im Fach Mathematik für den Sachunterricht keine länderübergreifenden Bildungsstandards gibt, wurde für die Analyse in erster Linie der Lehrplan aus Nordrhein-Westfalen herangezogen (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2021). Nordrhein-Westfalen ist das bevölkerungsreichste Land der Bundesrepublik Deutschland, weshalb Schüler:innen aus diesem Land einen erheblichen Teil der TIMSS-Stichprobe ausmachen, sodass eine Orientierung an diesen Vorgaben sinnvoll erscheint. Zudem ist der Lehrplan vergleichbar mit denen vieler anderer Länder in der Bundesrepublik Deutschland. Zusätzlich wurden der (länderübergreifende) Perspektivrahmen Sachunterricht (GDSU, 2013) sowie die für den Erhebungsraum aktuellen Lehrpläne der Länder Bayern (Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst, 2014), Baden-Württemberg (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2016), Niedersachsen (Niedersächsischen Kultusministerium, 2017) und Sachsen (Sächsisches Staatsministerium für Kultus, 2019) berücksichtigt. Der Lehrplan aus Nordrhein-Westfalen wie auch die Lehrpläne der anderen Länder berücksichtigen neben naturwissenschaftlichen Inhalten auch naturwissenschaftliche Denk- und Arbeitsweisen als relevante Zielbereiche. In den konkreten Kompetenzerwartungen für das Ende von Klasse 2 und das Ende von Klasse 4, die in der Regel in den Lehrplänen formuliert sind, werden Inhalte und (fachspezifische) Denk- und Arbeitsweisen integriert formuliert. Zentrale Kriterien für die curriculare Passung waren die erläuterten Inhaltsgebiete und Konzepte sowie die verwendeten Fachbegriffe, die sich in den verbindlichen Inhalten und den Kompetenzerwartungen wiederfinden. Die Unterrichtsvalidität, also die Frage, inwieweit die Inhalte der TIMSS-Aufgaben auch tatsächlich im Unterricht behandelt wurden, lässt sich mit dieser Analyse nicht einschätzen. So ist generell unklar, welche Steuerungsfunktion die curricularen Vorgaben haben. Im mehrperspektivischen Sachunterricht kommt hinzu, dass Lehrkräfte möglicherweise auch individuelle Schwerpunkte auf einzelne Perspektiven in ihrem Unterricht setzen.

In Tabelle 4.6 sind die Anteile der curricular nicht validen Testaufgaben aufgeschlüsselt nach den drei Inhaltsbereichen *Biologie, Physik/Chemie* und *Geografie* dargestellt. Insgesamt wurden 12 Prozent der Aufgaben als curricular nicht valide eingeschätzt. Der größte prozentuale Anteil liegt dabei im Inhaltsbereich *Geografie*, während im Bereich *Biologie* lediglich drei Prozent der Aufgaben als nicht valide eingeschätzt wurden.

Die Aufgaben aus dem Inhaltsbereich *Biologie* adressieren die Fotosynthese im Kontext einer Nahrungskette, was so nicht Bestandteil der Curricula in Deutschland ist, sowie die Notwendigkeit des Abkochens von Wasser zum Erhalt von Trinkwasser, was für die meisten Schüler:innen in Deutschland wenig vertraut sein dürfte. Im Inhaltsbereich *Physik/Chemie* stammen die curricular nicht validen Aufgaben überwiegend aus dem Teilgebiet Kräfte und Bewegungen, ein

Tabelle 4.6: Curricular nicht valide Testaufgaben in den drei Inhaltsbereichen

Naturwissenschaftliche Inhaltsbereiche	2007			2019			2023		
	Gesamt	nicht valide		Gesamt	nicht valide		Gesamt	nicht valide	
	n	*n*	%	*n*	*n*	%	*n*	*n*	%
Biologie	74	12	16	73	6	8	79	2	3
Physik/Chemie	64	3	5	61	4	7	61	10	16
Geografie	36	15	42	35	2	6	33	8	24
Gesamt	174	30	17	169	12	7	173	20	12

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

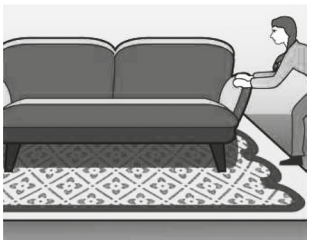
Bereich, der im Sachunterricht in der Regel unter der technischen Perspektive auf der Phänomenebene, beispielsweise im Hinblick auf einfache Kraftwandler oder Stabilität beim Bauen, thematisiert wird. Fachbegriffe wie Erdanziehungskraft oder auch die unterschiedlichen Richtungen von Kräften und deren detailliertes Zusammenspiel werden nur in Ausnahmen erarbeitet. Im Inhaltsgebiet *Geografie* wurde ein Teil der Aufgaben als curricular nicht valide eingeschätzt, die dem Teilgebiet Erde und Sonnensystem zugeordnet sind. Dieses Thema ist so nicht vorgesehen, auch wenn durchaus didaktische Materialien dazu vorliegen und es in manchen Bildungsplänen als Wahlpflichtthema aufgeführt wird. Weitere als nicht curricular valide eingeschätzte Aufgaben beziehen sich auf Aspekte der Erdgeschichte (Inhaltsgebiet: charakteristische Landschaftsmerkmale, Ressourcen und Erdgeschichte) wie die Bildung von Fossilien oder Landschaften, die in Deutschland in der Regel erst in den weiterführenden Schulen bearbeitet werden.

Die meisten Aufgaben, die als nicht curricular valide eingeschätzt wurden, stammen aus dem Anforderungsbereich *Reproduzieren* (zwölf Aufgaben). In vielen Fällen ist dies darauf zurückzuführen, dass die Schüler:innen zur Lösung der Aufgaben über Faktenwissen verfügen müssen, das im Lehrplan Nordrhein-Westfalens bzw. den weiteren Dokumenten nicht vorgesehen ist, z. B. Wissen über spezifische Tiere und Pflanzen oder spezifische Fachbegriffe. Aus den anderen Anforderungsbereichen *Anwenden* und *Problemlösen* wurden sieben bzw. zwei Aufgaben als nicht curricular valide eingeschätzt.

In Abbildung 4.3 ist als Beispiel eine Aufgabe dargestellt, die als nicht curricular valide eingeschätzt wurde, weil das Konzept Kraft im Regelfall nicht explizit eingeführt wird. Die Aufgabe kann vermutlich von Schüler:innen auch ohne schulische Lerngelegenheiten gelöst werden, wenn sie über ein erfahrungsbasiertes Wissen verfügen oder auch den Begriff Reibung kennen – ohne, dass sie ihn mit dem Kraft-Konzept in Beziehung setzen. Daher ist die Lösungshäufigkeit von 63 Prozent nicht überraschend.

Zusammenfassend kann festgehalten werden, dass die große Mehrzahl der Aufgaben in TIMSS 2023 als curricular valide eingeschätzt wurde. Der Anspruch der TIMSS-Rahmenkonzeption, Kompetenzen zu erfassen, die in der Schule aufgrund von curricularen Vorgaben vermittelt werden sollen, wurde also in Deutschland weitestgehend erfüllt. Im Vergleich zu den hier berücksichtigten vorherigen Studien zeigen sich eher kleinere Unterschiede: So schwanken die als nicht curricular valide eingeschätzten Aufgaben zwischen ca. 10 und 20 Prozent.

Abbildung 4.3: Beispiel für eine curricular nicht valide Aufgabe



Dana verschiebt ein Sofa. Sie findet es schwieriger, das Sofa zu schieben, wenn es auf dem Teppich steht, als wenn es auf dem Holzboden steht.

Wie heißt die Kraft, die zwischen dem Sofa und dem Teppich wirkt und die es für Dana schwierig macht, das Sofa zu schieben?

A Luftwiderstand

B Reibung

C Magnetkraft

D Erdanziehungskraft

(int. .58/nat. .63)

SE71152

Aufgabenbeispiel zu *Physik/Chemie*
Teilgebiet: Kräfte und Bewegungen

Die Werte in Klammern geben die relativen internationalen und nationalen Lösungshäufigkeiten an.
Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: https://www.iea.nl/data-tools/permission-requests

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Der Unterschied zu TIMSS 2007, als der Anteil nicht curricular valider Aufgaben höher eingeschätzt wurde, beruht vermutlich auf dem damaligen höheren Anteilen von Items im Anforderungsbereich Reproduzieren. Auch Weiterentwicklungen der curricularen Vorgaben in Deutschland in den letzten 10 Jahren sowie die Anpassungen der TIMSS-Rahmenkonzeptionen an die Curricula der verschiedenen Teilnehmerstaaten haben hier einen Einfluss auf die Unterschiede zwischen den Studienzyklen.

In TIMSS 2007, 2011 und 2015 wurden zusätzliche Analysen durchgeführt, bei denen die als nicht curricular valide eingeschätzten Aufgaben von den internationalen Vergleichen ausgeschlossen wurden. In TIMSS 2015, wo 10 Prozent der Aufgaben als nicht curricular valide in Deutschland eingeschätzt wurden, zeigten sich kaum Unterschiede in den Ergebnissen zwischen Analysen mit allen und Analysen mit lediglich den curricular validen Aufgaben. So erhöhte sich in den Analysen mit den ausgeschlossenen Aufgaben der Mittelwert der Schüler:innen aus Deutschland um 3 Punkte und auch der Rangplatz von Deutschland verschob sich um drei Plätze nach oben (Steffensky et al., 2016). Diese leichten Verschiebungen werden als wenig bedeutsam eingeschätzt. Aus diesem Grund wurde wie schon in TIMSS 2019 auf die entsprechenden aufwendigen Analysen, die keinen substanziellen Erkenntnisgewinn versprechen, verzichtet. In der Projektdokumentation der internationalen Studienleitung sind für die einzelnen teilnehmenden Staaten die Berechnungen mit den curricular validen Testaufgaben für TIMSS 2023 aber einzusehen (von Davier, Kennedy et al., 2024).

4.4.3 Kompetenzstufen in den Naturwissenschaften

Wie bereits in den vorherigen Studienzyklen wurden in TIMSS 2023 die Testwerte der Schüler:innen fünf Kompetenzstufen zugeordnet. Diese Stufen spiegeln die inhaltlichen und kognitiven Leistungsniveaus der Schüler:innen wider (für Details zur Skalierungsmethode: von Davier, Fishbein & Kennedy, 2024). Die Analyse der inhaltlichen und kognitiven Anforderungen, die zur erfolgreichen Bearbeitung der Aufgaben einer Kompetenzstufe erforderlich sind, lassen eine stärker inhaltlich-kriteriale Beschreibung und Interpretation der erreichten Testwerte zu.

Wie in Mathematik gibt es vier Bezugspunkte, sogenannte Benchmarks, die die Kompetenzskala in fünf Abschnitte einteilen, die den fünf Kompetenzstufen entsprechen. Diese werden von der internationalen TIMSS-Studienleitung festgelegt und sind identisch zu den bisherigen Studienzyklen, sodass Vergleiche der Verteilung auf die Kompetenzstufen über die Zeit möglich sind. In Tabelle 4.7 ist beschrieben, welches Wissen und Verständnis für die fünf Kompetenzstufen charakteristisch ist. Die Kompetenzstufen sind hierarchisch geordnet, d. h., Schüler:innen, die sich beispielsweise auf Kompetenzstufe IV befinden, lösen Testaufgaben, die typisch für diese Stufe sind, mit einer höheren Wahrscheinlichkeit als Testaufgaben der Stufe V. Leistungsschwache Schüler:innen befinden sich auf den unteren Kompetenzstufen I (< 400 Punkte) und II (400–474 Punkte). Sie verfügen lediglich über ein rudimentäres Wissen einzelner isolierter Fakten in den Naturwissenschaften. Schüler:innen in der mittleren Kompetenzstufe III (475–549 Punkte) können ihr Basiswissen nicht nur reproduzieren, sondern in einfachen Kontexten anwenden. Auf den obersten Kompetenzstufen IV (550–624 Punkte) und V (≥ 625 Punkte) befinden sich leistungsstarke Schüler:innen. Sie können Aufgaben mit mehreren Zwischenschritten lösen, Wissen in komplexeren Kontexten sowie naturwissenschaftliche Denk- und Arbeitsweisen anwenden.

Tabelle 4.7: Beschreibung der fünf Kompetenzstufen für die Gesamtskala Naturwissenschaften

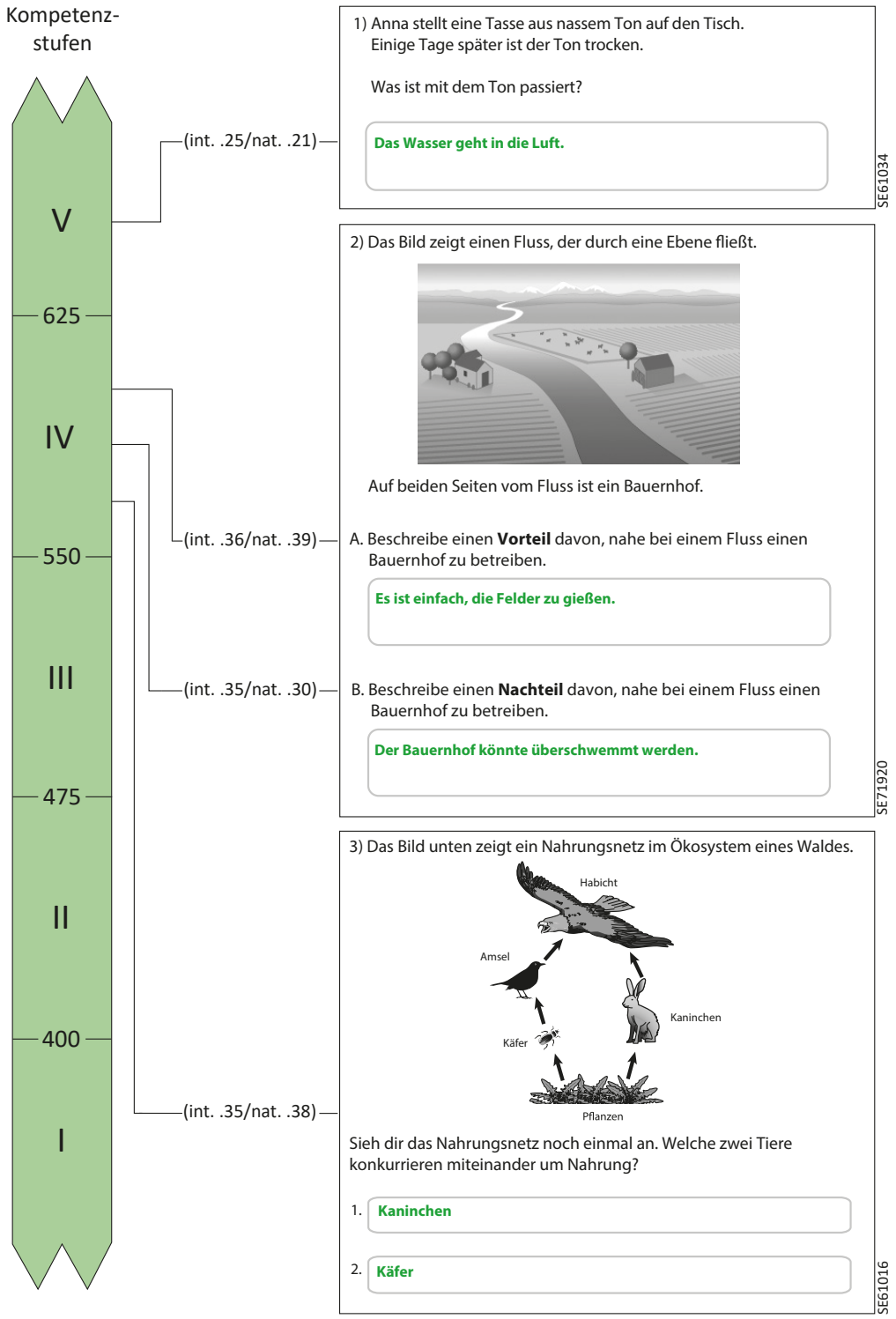
Kompetenzstufe V (ab 625): fortgeschritten
Die Schüler:innen verfügen über ein grundlegendes Verständnis in den drei Inhaltsbereichen (Biologie, Physik/Chemie und Geografie) sowie den naturwissenschaftlichen Denk- und Arbeitsweisen und können vielfältige Denk- und Arbeitsweisen nutzen. Sie zeigen z.B. Wissen über die Kennzeichen von Lebewesen und argumentieren über die Beziehungen von Organismen in Ökosystemen, haben Wissen über Vererbung, Krankheitserreger und Umweltverschmutzung. Sie kennen Eigenschaften von Stoffen, Veränderungen von Stoffen und argumentieren über die Löslichkeit in einem Laborkontext. Weiterhin verfügen sie über ein Verständnis der Eigenschaften der Erde und dort ablaufende Prozesse, das Sonnensystem und können erklären, wie menschliches Verhalten die Umwelt beeinflusst. Im Bereich der Denk- und Arbeitsweisen können sie einfache kontrollierte Versuche planen, Vorhersagen treffen und Schlussfolgerungen bewerten.
Kompetenzstufe IV (550–624): hoch
Die Schüler:innen verfügen über Wissen in den drei Inhaltsbereichen Biologie, Physik/Chemie und Geografie sowie zu den naturwissenschaftlichen Denk- und Arbeitsweisen und können dieses anwenden. Sie verfügen z.B. über Wissen über Kennzeichen des Lebendigen, über die Fortpflanzung und das Überleben von Tieren und wenden Wissen über Merkmale und Lebenszyklen von Pflanzen und Tieren sowie über die Ausbreitung von Krankheiten an. Sie können weiterhin Wissen über Aggregatzustände und Eigenschaften von Materialien, Magneten, Schall und Wärme anwenden und können mit dem Wissen über die Löslichkeit von Stoffen in einem alltäglichen Kontext argumentieren. Im Bereich Geografie haben sie Wissen über verschiedene Fakten über die Eigenschaften der Erde und wenden Wissen über die verschiedenen Klimazonen der Erde und die Veränderungen im Laufe der Zeit an. Die Schüler:innen beschreiben Beobachtungen und interpretieren Modelle und grafische Darstellungen.
Kompetenzstufe III (475–549): durchschnittlich
Die Schüler:innen haben ein Basiswissen in den drei Inhaltsbereichen Biologie, Physik/Chemie und Geografie und wenden dieses zum Teil an. Sie haben z.B. ein Basiswissen über Pflanzen und Tiere, die menschliche Gesundheit, Eigenschaften von Materie, Energie und Licht, die Erdoberfläche und wenden Grundkenntnisse über Kräfte und Bewegung an. Die Schüler:innen können Beobachtungen teilweise beschreiben und Beobachtungen und Daten mit naturwissenschaftlichen Fakten in Verbindung bringen.
Kompetenzstufe II (400–474): niedrig
Die Schüler:innen verfügen über ein naturwissenschaftliches Basiswissen. Sie haben z.B. Grundkenntnisse, was Pflanzen zum Wachsen brauchen, über einzelne Eigenschaften von Stoffen in alltäglichen Kontexten. Sie können menschliche Verhaltensweise in umweltschädlich oder nicht einordnen. Sie beschreiben das Vorkommen von Salzwasser auf der Erde und ausgewählte Folgen der globalen Erwärmung.
Kompetenzstufe I (unter 400): rudimentär
Die Schüler:innen verfügen über rudimentäres schulisches Anfangswissen. Selbst einfache Aufgaben werden nur gelegentlich oder ansatzweise gelöst.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Die folgenden Abbildungen 4.4 und 4.5 zeigen Beispiele für Aufgaben auf den unterschiedlichen Kompetenzstufen. Die in Klammern gezeigten Werte entsprechen den internationalen und nationalen relativen Lösungshäufigkeiten. Beispielsweise wurde die erste Aufgabe von 25 Prozent der Schüler:innen im internationalen Durchschnitt und von 21 Prozent der Schüler:innen in Deutschland gelöst. Die Zuordnung einer Aufgabe zu den Kompetenzstufen erfolgt anhand der internationalen Lösungshäufigkeiten (von Davier, Fishbein & Kennedy, 2024). Im Folgenden werden für die sieben Aufgaben die Kompetenzen beschrieben, die Schüler:innen zu ihrer Lösung benötigen.[3]

3 Zum Zeitpunkt der Erstellung des nationalen Berichtsbandes standen die für die Zuordnung der TIMSS-2023-Aufgaben zu den Kompetenzstufen erforderlichen Informationen noch nicht zur Verfügung, sodass in den Abbildungen 4.4 und 4.5 Beispielaufgaben und Daten aus dem vorangegangen Studienzyklus von 2019 (Steffensky et al., 2020) verwendet wurden.

Abbildung 4.4: Kompetenzstufen und Beispielaufgaben (Gesamtskala Naturwissenschaften) I

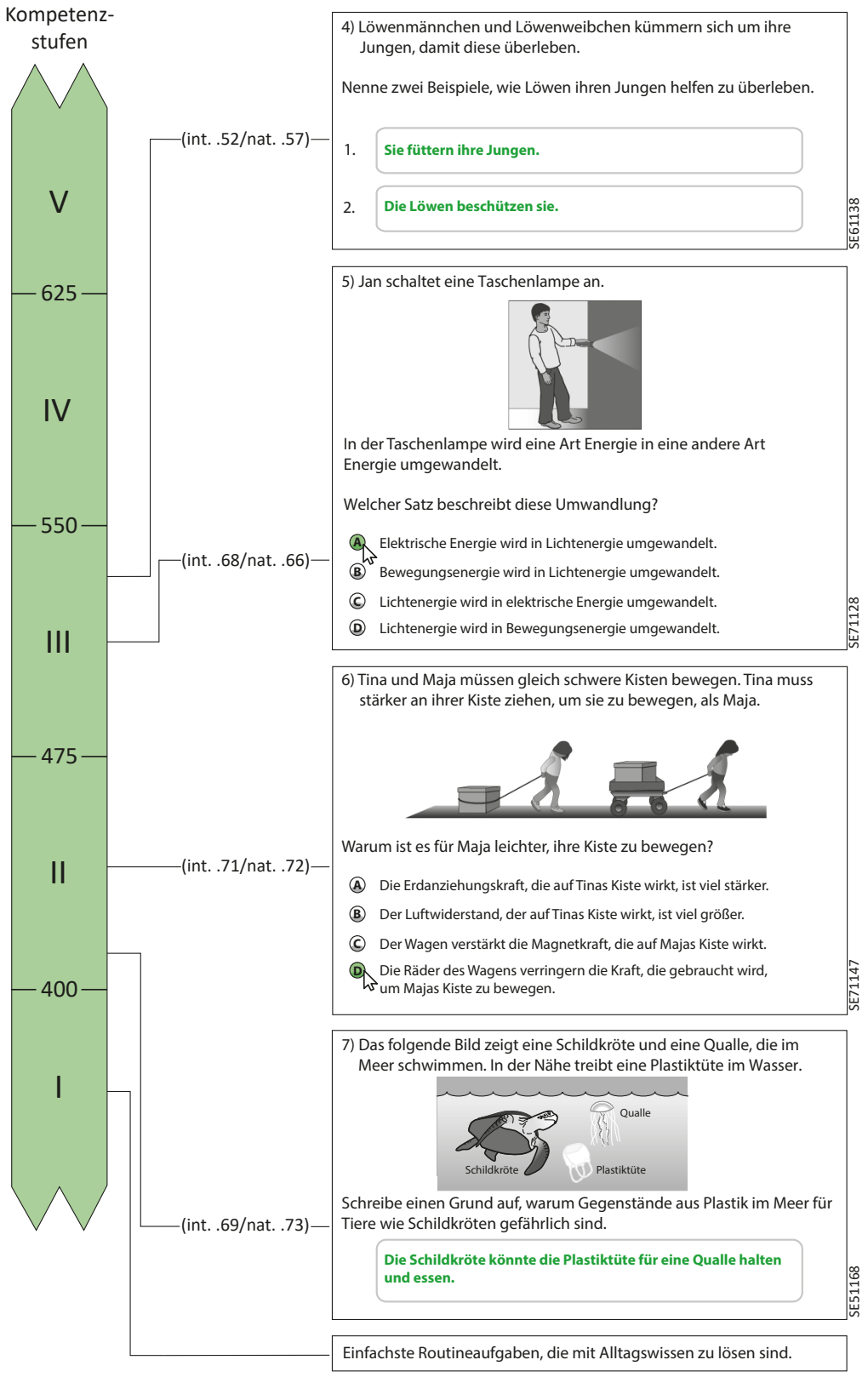


Da die erforderlichen Informationen zur Erstellung der Abbildung für TIMSS 2023 durch die internationale Studienleitung erst mit dem Datum der Veröffentlichung des Berichts bereitgestellt werden, wird zur Illustration der Aufgaben auf den verschiedenen Kompetenzstufen die Abbildung aus TIMSS 2019 (Steffensky et al., 2020) genutzt.

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

Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: https://www.iea.nl/publications/form/iea-permission-request-form

Abbildung 4.5: Kompetenzstufen und Beispielaufgaben (Gesamtskala Naturwissenschaften) II



Da die erforderlichen Informationen zur Erstellung der Abbildung für TIMSS 2023 durch die internationale Studienleitung erst mit dem Datum der Veröffentlichung des Berichts bereitgestellt werden, wird zur Illustration der Aufgaben auf den verschiedenen Kompetenzstufen die Abbildung aus TIMSS 2019 (Steffensky et al., 2020) genutzt.

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

Abdruck und Nutzung der Aufgaben nur mit ausdrücklicher Genehmigung der IEA: https://www.iea.nl/publications/form/iea-permission-request-form

Beispiele für Kompetenzstufe V (fortgeschritten)

In der Aufgabe 1 sollen die Schüler:innen erläutern, was beim Trocknen von Ton passiert. Das Beispiel ist zwar (relativ) alltagsnah, aber es ist in der Regel kein typisches Unterrichtsbeispiel, an dem der Prozess der Verdunstung thematisiert wird. Dieser wird häufig am Beispiel des Wasserkreislaufs erarbeitet. Die Schüler:innen müssen also ihr Wissen über die Verdunstung von Wasser auf ein (in der Regel) neues Beispiel anwenden. Sie können dabei den Fachbegriff Verdunsten nutzen oder den Vorgang umschreiben („Das Wasser geht in die Luft").

Beispiel für Kompetenzstufe IV (hoch)

Aufgabe 2 stammt aus dem Inhaltsbereich *Geografie*. Die Schüler:innen können im ersten Teil verschiedene Vorteile anführen, die auf die Bedeutung von Wasser für Tiere und Pflanzen verweisen oder aber auch auf die Möglichkeit, das bewegte Wasser mithilfe eines Wasserrads zur Stromerzeugung zu nutzen. Dabei wenden sie naturwissenschaftliches Wissen an. Als richtig werden auch Antworten gewertet, die auf die Qualität des Bodens nahe dem Wasser eingehen. Im zweiten Teil der Aufgabe soll ein Nachteil genannt werden. Mögliche Nachteile sind die Überflutungsgefahr, aber auch die Gefahr, dass das Wasser aus dem Fluss mit Schadstoffen belastet sein könnte und dann auf die Felder oder zu den Tieren gelangt oder umgekehrt der Dung der Tiere in das Wasser gelangt und somit den Fluss belastet. Dieser Teil der Aufgabe fällt den Schüler:innen aus Deutschland etwas schwerer, möglicherweise, weil diese Aspekte im geografiebezogenen Sachunterricht seltener im Mittelpunkt stehen.

In der Aufgabe 3 ist ein einfaches Nahrungsnetz dargestellt. Aus der Abbildung müssen die Schüler:innen schlussfolgern, welche beiden Tiere um Nahrung konkurrieren. Die Aufgabe erfordert ein erstes Verständnis über Nahrungsbeziehungen in einem Ökosystem (hier Wald).

Beispiele für Kompetenzstufe III (durchschnittlich)

Aufgabe 4 ist eine eher schwierige Aufgabe aus dem Anforderungsbereich *Reproduzieren*. Die Schüler:innen müssen zwei unterschiedliche Gründe nennen, wie Löwen Brutpflege betreiben. Als richtig werden Antworten gewertet, in denen auf das Füttern oder den Schutz vor Feinden verwiesen wird oder wie man jagt oder Wasserstellen findet. Auch hier sind konkrete Antworten („Beschützen vor Hyänen") und generalisierte Antworten möglich („Beschützen vor Feinden").

Auch Aufgabe 5 ist eine eher schwierige Aufgabe aus dem Anforderungsbereich *Reproduzieren*. Hier müssen die Schüler:innen Wissen über die Umwandlung verschiedener Energieformen in einem alltagsnahen Kontext (Taschenlampe) haben.

Beispiele für Kompetenzstufe II (niedrig)

In Aufgabe 6 müssen die Schüler:innen aus vorgegebenen Antworten diejenige wählen, die erklärt, warum sich ein schwerer Gegenstand leichter mit einem Karren ziehen lässt. Selbst wenn die Schüler:innen noch keine differenziertere Vorstellung von Kräften haben, können sie hier Alltagserfahrungen anwenden, um die Aufgabe zu lösen.

In der Aufgabe 7 müssen die Schüler:innen einen Grund angeben, warum Gegenstände aus Plastik für Meerestiere, wie eine Schildkröte, gefährlich sind. Als richtig werden Antworten gewertet, die auf die Gefahr des Erstickens, des

Fressens oder der eingeschränkten Bewegungsfreiheit verweisen. Bei dieser Aufgabe fällt auf, dass die Schüler:innen in Deutschland im internationalen Vergleich einen etwas höheren Mittelwert haben. Dies liegt vermutlich daran, dass das Thema Plastik im Meer in Deutschland in den letzten Jahren viel Aufmerksamkeit erfahren hat und im Zusammenhang mit Bildung für nachhaltige Entwicklung auch im Sachunterricht häufig thematisiert wurde.

4.5 Kompetenzen im internationalen Vergleich

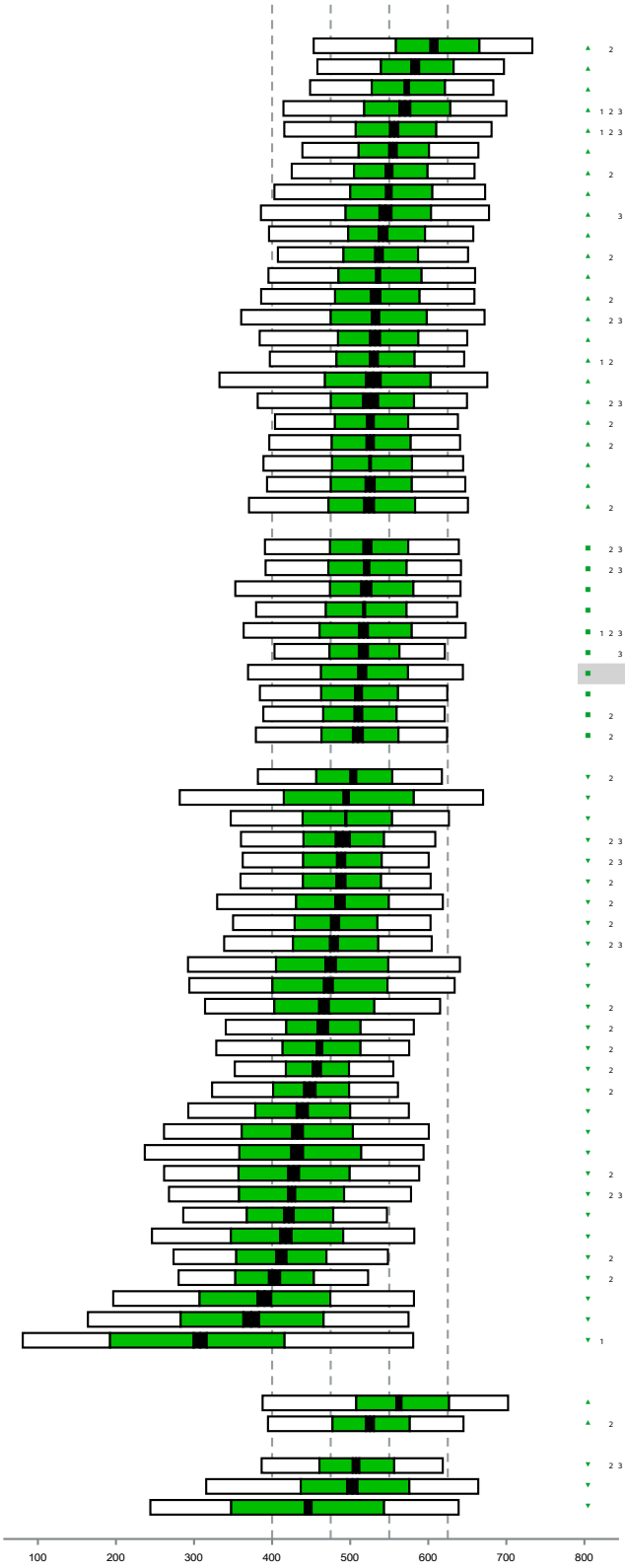
Im Folgenden wird auf die Ergebnisse der Gesamtskala Naturwissenschaften im internationalen Vergleich eingegangen. Dabei werden die Leistungsmittelwerte, die Streuung der Werte sowie der Vergleich mit TIMSS 2007 und 2019 dargestellt.

4.5.1 Unterschiede im Niveau der naturwissenschaftlichen Kompetenz im internationalen Vergleich

In Abbildung 4.6 sind die Mittelwerte (*M*) und Standardabweichungen (*SD*) mit ihren jeweiligen Standardfehlern (*SE*) für die naturwissenschaftlichen Kompetenzwerte der teilnehmenden Staaten und Benchmark-Teilnehmer angegeben. Die Standardabweichung ist ein Maß der Streuung der Werte innerhalb eines Staates, also der Heterogenität der Kompetenzwerte. Je größer die Standardabweichung, desto breiter ist die Kompetenzverteilung. Die Standardfehler geben Auskunft über die Präzision der Mittelwerts- und Standardabweichungsschätzung. Je kleiner die Standardfehler, desto präziser die Schätzungen. Zusätzlich zu den Mittelwerten einzelner Staaten sind die Mittelwerte der zuvor beschriebenen Vergleichsgruppen (alle Teilnehmerstaaten, OECD- und EU-Staaten) angegeben. Die Staaten sind entsprechend ihres Abschneidens in TIMSS 2023 angeordnet.

Der internationale Mittelwert der erreichten Kompetenzwerte liegt bei 494 Punkten. 33 Staaten liegen über dem Mittelwert und 25 darunter. Die höchsten mittleren Kompetenzwerte zeigen Schüler:innen aus Singapur (607 Punkte). Mit einem Abstand von 24 Punkten folgt die Republik Korea (Südkorea; 583 Punkte) und wiederum mit einem Abstand von ca. 10 Punkten folgen Taiwan und die Türkei (573 bzw. 570 Punkte). Am schwächsten schneiden Marokko, Kuwait und mit Abstand Südafrika ab, deren mittlere Kompetenzwerte zum Teil deutlich mehr als 100 Punkte unter dem internationalen Mittelwert liegen (390, 373 bzw. 308 Punkte). In Deutschland liegt der mittlere Kompetenzwert der Schüler:innen bei 515 Punkten, also 21 Punkte über dem internationalen Mittelwert. Staaten, die vergleichbar wie Deutschland abschneiden, zwischen denen sich die Mittelwerte also statistisch nicht signifikant unterscheiden, sind Dänemark, Kanada, Italien, die Niederlande, Neuseeland, Portugal, Serbien und die Slowakei. Der Mittelwert in Deutschland liegt sehr nah an dem Mittelwert der EU-Staaten (518), aber mit 11 Punkten signifikant unter dem der OECD-Staaten (526). Im Vergleich zu Singapur als Spitzenreiter zeigt sich eine Differenz von 92 Punkten, ein Unterschied, der mehr als einer Kompetenzstufe entspricht.

Abbildung 4.6: Testleistung der Schüler:innen im internationalen Vergleich – Gesamtskala Naturwissenschaften



	Teilnehmer	*M*	*(SE)*	*SD*	*(SE)*
▲ 2	*Singapur*	607	(2.3)	84	(1.3)
▲	Republik Korea (Südkorea)	583	(2.5)	74	(2.2)
▲	Taiwan	573	(1.4)	71	(1.4)
▲ 1 2 3	*Türkei (5. Jgst.)*	570	(3.4)	86	(1.7)
▲ 1 2 3	*England*	556	(2.6)	80	(2.0)
▲	Japan	555	(2.4)	68	(1.4)
▲ 2	*Polen*	550	(2.2)	71	(1.1)
▲	Australien	550	(1.8)	81	(0.9)
▲ 3	*Hongkong*	545	(3.8)	88	(2.7)
▲	Finnland	542	(2.8)	79	(1.8)
▲ 2	*Litauen*	537	(2.4)	74	(1.6)
▲	Macau	536	(1.3)	81	(0.8)
▲ 2	*Schweden*	533	(3.1)	82	(2.0)
▲ 2 3	*USA*	532	(2.4)	93	(0.9)
▲	Irland	532	(3.2)	80	(1.6)
▲ 1 2	*Norwegen (5. Jgst.)*	530	(2.5)	75	(1.1)
▲	Bulgarien	530	(4.8)	104	(4.1)
▲ 2 3	*Rumänien*	526	(4.8)	81	(2.4)
▲ 2	*Tschechien*	526	(2.2)	71	(1.2)
▲ 2	*Slowenien*	526	(2.3)	74	(1.4)
▲	**VG OECD**	**526**	**(0.5)**	**77**	**(0.3)**
▲	Lettland	526	(2.9)	77	(1.9)
▲ 2	*Ungarn*	524	(3.1)	85	(2.5)
■ 2 3	*Dänemark*	522	(2.5)	75	(1.2)
■ 2 3	*Kanada*	521	(1.7)	76	(1.1)
■	Slowakei	521	(3.3)	87	(2.5)
■	**VG EU**	**518**	**(0.6)**	**78**	**(0.4)**
■ 1 2 3	*Neuseeland*	517	(2.8)	87	(1.7)
■ 3	*Niederlande*	517	(2.9)	67	(1.5)
■	Deutschland	515	(2.6)	84	(1.7)
■	Portugal	511	(2.3)	73	(1.1)
■ 2	*Italien*	511	(2.5)	70	(1.1)
■ 2	*Serbien*	510	(3.2)	74	(1.6)
▼ 2	*Spanien*	504	(1.9)	72	(0.9)
▼	Vereinigte Arabische Emirate (VAE)	495	(1.7)	118	(0.7)
▼	**Internationaler Mittelwert**	**494**	**(0.4)**	**85**	**(0.2)**
▼ 2 3	*Albanien*	491	(4.5)	77	(2.8)
▼ 2 3	*Belgien (Fläm. Gem.)*	488	(2.6)	73	(1.8)
▼ 2	*Frankreich*	488	(2.9)	74	(1.6)
▼ 2	*Zypern*	487	(3.0)	88	(1.6)
▼ 2	*Belgien (Frz. Gem.)*	481	(2.6)	77	(1.6)
▼ 2 3	*Chile*	479	(2.5)	80	(1.4)
▼	Bahrain	475	(3.4)	106	(2.5)
▼	Katar	472	(2.8)	104	(1.5)
▼ 2	*Kasachstan*	467	(3.1)	91	(1.8)
▼ 2	*Georgien*	465	(3.4)	72	(1.9)
▼ 2	*Montenegro*	461	(1.8)	75	(1.1)
▼ 2	*Armenien*	457	(2.6)	61	(1.3)
▼ 2	*Bosnien und Herzegowina*	448	(3.7)	72	(2.0)
▼	Nordmazedonien	439	(3.6)	86	(1.8)
▼	Oman	433	(3.4)	103	(1.8)
▼	Iran	432	(3.8)	109	(1.8)
▼ 2	*Saudi-Arabien*	428	(3.5)	100	(1.9)
▼ 2 3	*Brasilien*	425	(2.3)	94	(1.0)
▼	Aserbaidschan	422	(3.0)	80	(1.3)
▼	Jordanien	418	(3.6)	102	(2.3)
▼ 2	*Usbekistan*	412	(3.2)	83	(1.1)
▼ 2	*Kosovo*	403	(3.6)	74	(1.8)
▼	Marokko	390	(4.3)	118	(1.9)
▼ 4	*Kuwait*	373	(5.0)	126	(2.9)
▼ 1 4	*Südafrika (5. Jgst.)*	308	(4.2)	153	(3.0)
	Benchmark-Teilnehmer				
▲	Dubai, VAE	562	(1.7)	95	(1.0)
▲ 2	*Ontario, Kanada*	525	(2.6)	76	(1.2)
▼ 2 3	*Québec, Kanada*	508	(2.1)	71	(1.4)
▼	Schardscha, VAE	503	(3.3)	105	(1.7)
▼	Abu Dhabi, VAE	446	(2.2)	125	(1.2)

▲ = Mittelwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Mittelwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu Deutschland nicht statistisch signifikant ($p > .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten

© TIMSS 2023

Streuung der naturwissenschaftlichen Kompetenz im internationalen Vergleich:
Neben dem Mittelwert sind die Streuung und die Verteilung der Kompetenzwerte wichtige Indikatoren für die Leistungsfähigkeit von Bildungssystemen. Wünschenswert ist eine möglichst geringe Streuung, also eine geringe Heterogenität der Leistungen der Schüler:innen, bei gleichzeitig hohen Kompetenzwerten. Streuung und Verteilung lassen sich an den Standardabweichungen und der Breite der Perzentilbänder ablesen (Abbildung 4.6). Die Perzentilwerte teilen die Bandbreite der Kompetenzwerte in definierte Prozentbereiche ein. Der Kompetenzwert, der dem 75. Perzentil entspricht, beschreibt, dass 75 Prozent der Schüler:innen einen niedrigeren oder identischen Kompetenzwert und entsprechend 25 Prozent einen höheren Kompetenzwert erzielt haben. Die Breite des Perzentilbandes vom 5. bis zum 95. Perzentil repräsentiert die Streuung der Kompetenzwerte von 90 Prozent der Schüler:innen, ausgeschlossen sind hierbei die jeweils 5 Prozent Schüler:innen mit den niedrigsten und den höchsten Kompetenzwerten. Entsprechend gibt die Breite des Perzentilbandes vom 25. bis zum 75. Perzentil die Bandbreite der 50 Prozent der Schüler:innen an, die nicht zu den jeweils 25 Prozent mit den niedrigsten und höchsten Kompetenzwerten gehören.

Abbildung 4.6 zeigt, dass in allen Staaten die Kompetenzwerte stark streuen, es aber trotzdem zwischen den Staaten große Unterschiede gibt. Sehr große Streuungen finden sich vor allem in den Staaten mit eher niedrigen Kompetenzwerten. Beispielsweise zeigen sich Kompetenzunterschiede zwischen dem 5. und dem 95. Perzentil in Südafrika von 500 Punkten (*SD* = 153) und in Kuwait von 411 Punkten (*SD* = 126). Bei den europäischen Staaten zeigt sich in Bulgarien eine relativ große Streuung (343 Punkte, *SD* = 104) bei einem Kompetenzwert über dem internationalen Durchschnitt (530 Punkte). Zwei Staaten mit einem sehr hohen Niveau und gleichzeitig geringen Streuungen zwischen dem 5. und dem 95. Perzentil sind Taiwan und die Republik Korea (Südkorea; 234 bzw. 239 Punkte; *SD* = 71 bzw. 74). Diesen Staaten gelingt es wie schon in den vorherigen Studienzyklen, die naturwissenschaftliche Kompetenz ihrer Grundschüler:innen in der Breite sehr gut zu fördern, aber gleichzeitig die leistungsstärkste Gruppe nicht zu vernachlässigen. Von den Staaten im oberen Drittel der Rangreihe ist Japan hervorzuheben, das eine besonders geringe Streuung von 225 Punkten aufweist (*SD* = 68). Sehr geringe Streuungen zeigen sich auch in den Niederlanden, Italien und Tschechien mit 218, 232 bzw. 234 Punkten (*SD* = 67, 70 bzw. 71). Allerdings sind hier die Kompetenzwerte, die die 5 Prozent der leistungsstärksten Schüler:innen erreichen, niedriger im Vergleich zu denen der vorher genannten Staaten.

In Deutschland erreichen die 5 Prozent der leistungsschwächsten Schüler:innen maximal einen Kompetenzwert von 369 Punkten und die 5 Prozent der leistungsstärksten minimal einen Kompetenzwert von 644 Punkten. Der Unterschied liegt entsprechend bei 275 Punkten. Die Standardabweichung in Deutschland liegt bei 84 Punkten. Dies bedeutet, dass unter der Annahme normalverteilter Testwerte ca. zwei Drittel der Schüler:innen Kompetenzwerte zwischen 431 (515 - 84) und 599 (515 + 84) Punkten erreichen. Damit ist die Standardabweichung etwas höher als die der EU- und OECD-Vergleichsgruppe (*SD* = 78 bzw. 77).

Leistungsstärkste und leistungsschwächste Schüler:innen:

Für die Einschätzung, wie gut die Förderung der leistungsstärksten und -schwächsten Schüler:innen in einem Staat gelingt, eignet sich die Betrachtung der Kompetenzwerte des 95-Prozent- und des 5-Prozent-Perzentils. Eine besonders gelungene Spitzenförderung zeigt sich in Singapur (hier beginnt die Spitzengruppe bei 733 Punkten), der Türkei (700 Punkte) und der Republik Korea (Südkorea; 699 Punkte). Weitere relativ kompetenzstarke Staaten wie Japan, Polen, Australien und England zeigen Werte zwischen 658 (Polen) und 672 (Australien) Punkten. Die 5 Prozent leistungsstärksten Schüler:innen in Deutschland erreichen mindestens 644 Punkte und liegen damit (sehr) deutlich unter den genannten Werten. Auch bei der Betrachtung der 5 Prozent leistungsschwächsten Schüler:innen zeigen sich große Unterschiede zwischen den Teilnehmerstaaten. So erreichen diese in den (sehr) kompetenzstarken Staaten maximale Werte von 458 (Singapur), 453 (Republik Korea [Südkorea]) oder 425 (Polen) Punkten, während in Südafrika beispielsweise der Wert bei 81 Punkten liegt. In Deutschland liegt der maximale Wert der 5 Prozent leistungsschwächsten Schüler:innen bei 369 Punkten. Wie auch in den vergangenen Studienzyklen deuten die Befunde aus Deutschland darauf hin, dass es weniger gut gelingt, hochkompetente Schüler:innen sowie solche mit ungünstigen Voraussetzungen hinreichend zu fördern.

Unterschiede in den mittleren Kompetenzen zwischen TIMSS 2007, 2019 und 2023:

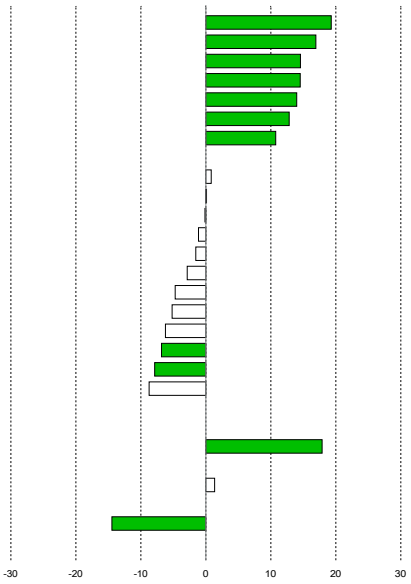
Ein wichtiges Element der TIMS-Studie sind die Trendanalysen über Studienzyklen hinweg. An dieser Stelle werden die aktuellen Ergebnisse aus einer lang- und kurzfristigen Perspektive (TIMSS 2007 und TIMSS 2019) verglichen. In Abbildung 4.7 sind die nationalen Mittelwerte und die zugehörigen Standardfehler der naturwissenschaftlichen Kompetenz für die Staaten gegenübergestellt, die seit dem Jahr 2007 kontinuierlich an TIMSS teilgenommen haben. Im rechten Teil der Abbildung sind die Differenzbalken für den Vergleich TIMSS 2019 und TIMSS 2023 dargestellt.

Während die mittlere Kompetenz in den Naturwissenschaften in Deutschland zwischen 2007 bis 2015 stabil war (528 Punkte), zeigt sich in TIMSS 2019 eine signifikante Abnahme um 10 Punkte. Auch wenn in TIMSS 2019 erstmalig die computerbasierte Testung eingeführt wurde, zeigen Analysen, dass diese Umstellung nicht ursächlich für diese Veränderung ist (Robitzsch et al., 2020). Im aktuellen Studienzyklus hat sich der mittlere Kompetenzwert nochmals um 3 Punkte verringert. Die Veränderung gegenüber 2019 ist allerdings statistisch nicht signifikant.

Die Veränderungen zwischen 2019 und 2023 in den anderen teilnehmenden Staaten sind unterschiedlich (siehe Abbildung 4.7). Betrachtet man alle an TIMSS 2023 teilnehmenden Staaten, so zeigt sich in 26 Staaten eine Zunahme der Testwerte und in 27 Staaten eine Abnahme im Vergleich zu TIMSS 2019, wobei die Unterschiede oft sehr klein und damit nicht statistisch bedeutsam sind. In drei Staaten zeigen sich keine Veränderungen. Eindeutige Muster lassen sich nicht erkennen. In vielen EU-Staaten sind kaum Unterschiede zwischen den letzten beiden Studienzyklen sichtbar, z. B. in Dänemark, Frankreich, Ungarn, Irland, Italien, den Niederlanden und Schweden. Das bedeutet im Umkehrschluss, dass sich potenzielle negative Effekte der COVID-19-Pandemie nicht in den mittleren Kompetenzwerten widerspiegeln. Etwas auffälligere Entwicklungen zeigen

Abbildung 4.7: Vergleich der Testleistungen zwischen TIMSS 2007, 2019 und 2023 – Gesamtskala Naturwissenschaften

Fußnoten 2007	2019	2023	T	Teilnehmer	2007 M_{07}	(SE)	2019 M_{19}	(SE)	2023 M_{23}	(SE)	Veränderungen[A] M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		
1 3	1 2	1 2 3	8	England	542	(2.9)	537	(2.7)	556	(2.6)	15	(3.9)	+	19	(3.8)	+	▲
				Australien	527	(3.3)	533	(2.4)	550	(1.8)	22	(3.8)	+	17	(3.0)	+	▲
				Taiwan	557	(2.0)	558	(1.8)	573	(1.4)	16	(2.4)	+	15	(2.2)	+	▲
1 2	1 2	1 2 3	8	Neuseeland	504	(2.6)	503	(2.3)	517	(2.8)	13	(3.9)	+	15	(3.6)	+	▲
2 3	3	3	8	Hongkong	554	(3.5)	531	(3.3)	545	(3.8)	-9	(5.2)		14	(5.1)	+	▲
2	2	2	8	Singapur	587	(4.1)	595	(3.4)	607	(2.3)	21	(4.7)	+	13	(4.1)	+	▲
2	2	2		Georgien	418	(4.6)	454	(3.9)	465	(3.4)	47	(5.7)	+	11	(5.2)	+	▲
2		2	8	Italien	535	(3.2)	510	(3.0)	511	(2.5)	-25	(4.0)	–	1	(3.9)		■
3	3	2 3	8	Dänemark	517	(2.9)	522	(2.4)	522	(2.5)	5	(3.8)		0	(3.5)		■
	2			Slowakei	526	(4.8)	521	(3.7)	521	(3.3)	-5	(5.8)		0	(4.9)		■
2	2	2	8	Litauen	514	(2.4)	538	(2.5)	537	(2.4)	23	(3.4)	+	-1	(3.5)		■
3	3	3	8	Niederlande	523	(2.6)	518	(2.9)	517	(2.9)	-6	(3.9)		-2	(4.1)		■
				Deutschland	528	(2.4)	518	(2.2)	515	(2.6)	-12	(3.5)	–	-3	(3.4)		■
	2	2		Schweden	525	(2.9)	537	(3.3)	533	(3.1)	8	(4.2)		-5	(4.5)		■
		2		Ungarn	536	(3.3)	529	(2.7)	524	(3.1)	-12	(4.6)	–	-5	(4.1)		■
2 3	2 3	2 3		USA	539	(2.7)	539	(2.7)	532	(2.4)	-6	(3.6)		-6	(3.6)		■
	3		8	Japan	548	(2.1)	562	(1.8)	555	(2.4)	7	(3.2)	+	-7	(3.0)	–	■
		2		Tschechien	515	(3.1)	534	(2.6)	526	(2.2)	11	(3.8)	+	-8	(3.4)	–	■
			8	Iran	436	(4.3)	441	(4.1)	432	(3.8)	-4	(5.7)		-9	(5.6)		■
				Benchmark-Teilnehmer													
2 3 6	2		8	Dubai, VAE	460	(2.8)	545	(1.7)	562	(1.7)	103	(3.2)	+	18	(2.4)	+	▲
2		2	8	Ontario, Kanada	536	(3.7)	524	(3.2)	525	(2.6)	-11	(4.5)	–	1	(4.1)		■
2		3	2 3 8	Québec, Kanada	517	(2.7)	522	(2.5)	508	(2.1)	-10	(3.4)	–	-14	(3.3)	–	▼



Keine statistisch signifikanten Veränderungen zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Veränderungen zwischen 2019 und 2023 ($p \leq .05$)
▲ = Differenzwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
\+ = Differenzwert in 2023 statistisch signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Differenzwert in 2023 statistisch signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

sich in England und Polen (+19 Punkte) sowie in Staaten, die bislang eher unter dem internationalen Mittelwert lagen wie Qatar oder die Vereinigten Arabischen Emirate (+23 Punkte). Auch der langfristige Trend zeigt über die teilnehmenden Staaten hinweg ein eindeutiges Bild.

Unterschiede in der Streuung zwischen TIMSS 2007, 2019 und 2023:

Neben Unterschieden in den Mittelwerten ist auch der Vergleich der Heterogenität der Leistungen in den verschiedenen Studienzyklen von Interesse. In Tabelle 4.8 sind neben den Angaben zu den Perzentilen für die Kompetenzen in TIMSS 2023, die zuvor bereits beschrieben wurden, auch die entsprechenden Werte für die vorherigen Studienzyklen dargestellt.

Die Staaten sind in der Rangfolge des mittleren Kompetenzwertes in 2023 geordnet. Für Deutschland zeigen die Ergebnisse, dass die Streuung der Leistungen (zwischen 5. und 95. Perzentil) im Vergleich zu TIMSS 2007 (254 Punkte) und TIMSS 2019 (251 Punkte) mit 275 Punkten zugenommen hat. An dieser Stelle ist zu berücksichtigen, dass in TIMSS 2011 und 2015 die Streuung deutlich geringer war (230 bzw. 229 Punkte) als in TIMSS 2007 und 2019. Die Unterschiede ergeben sich vor allem durch die breitere Streuung im unteren Leistungsbereich. Betrachtet man das 5. Perzentil, also den Kompetenzwert, den die 5 Prozent der leistungsschwächsten Schüler:innen maximal erreichen, so liegt der Wert 2023 bei 369 Punkten, 2019 lag er bei 383 Punkten und 2007 bei 393 Punkten. Im oberen Leistungsspektrum sind die Werte im Vergleich zu den früheren

Tabelle 4.8: Testleistungen nach Perzentilen im Vergleich: TIMSS 2007, 2019 und 2023 – Gesamtskala Naturwissenschaften

Fußnoten 2007	2019	2023	T	Teilnehmer	TIMSS 2007 Perzentile											TIMSS 2019 Perzentile											TIMSS 2023 Perzentile										
					5	(SE)		25	(SE)		75	(SE)		95	(SE)		5	(SE)		25	(SE)		75	(SE)		95	(SE)		5	(SE)		25	(SE)		75	(SE)		95	(SE)	
	2	2	8	Singapur	418	(7.0)	+	531	(6.0)	+	652	(4.4)	+	727	(3.9)		454	(6.3)		548	(4.5)	+	649	(3.3)	+	708	(3.7)	+	453	(6.7)	▲	559	(2.6)	▲	666	(2.1)	▲	733	(2.6)	▲
				Taiwan	423	(6.5)	+	508	(2.0)	+	609	(2.4)	+	679	(3.0)		444	(4.7)		516	(2.7)	+	604	(1.9)	+	659	(4.1)	+	449	(4.7)	▲	528	(1.8)	▲	621	(1.6)	▲	683	(2.7)	▲
1 3	1 2	1 2 3	8	England	403	(6.4)		491	(6.0)	+	596	(3.6)	+	666	(6.3)		413	(6.8)		491	(3.3)	+	587	(3.3)	+	648	(4.8)	+	416	(7.2)	▲	507	(3.0)	▲	610	(3.3)	▲	681	(6.4)	▲
	3		8	Japan	428	(7.0)		505	(3.1)		595	(1.4)		655	(3.2)		442	(4.5)		519	(2.7)	–	609	(2.1)	–	668	(3.4)		439	(4.1)	▲	511	(2.5)	▲	601	(3.2)	▲	664	(5.0)	▲
				Australien	384	(9.0)		478	(4.4)	+	583	(2.6)	+	651	(4.8)	+	389	(6.0)		484	(2.8)	+	587	(2.4)	+	653	(4.3)	+	403	(5.4)	▲	500	(2.9)	▲	605	(2.7)	▲	673	(2.6)	▲
2 3		3	8	Hongkong	437	(4.6)	–	511	(3.7)	–	601	(3.6)		659	(4.2)		408	(6.5)		486	(4.0)		580	(3.2)	+	643	(4.1)	+	386	(9.6)	■	494	(5.0)	▲	604	(4.1)	▲	678	(8.8)	▲
2	2	2	8	Litauen	401	(5.2)		473	(2.8)	+	559	(3.2)	+	615	(4.2)	+	414	(6.0)		492	(3.4)		588	(2.2)		651	(4.7)		407	(6.5)	▲	491	(3.9)	▲	587	(2.6)	▲	651	(4.3)	■
	2	2		Schweden	400	(3.3)		478	(4.0)		575	(3.0)	+	642	(4.0)	+	409	(7.2)	–	490	(4.6)		588	(3.8)		652	(3.8)		386	(8.7)	■	481	(4.5)	▲	589	(2.8)	▲	659	(3.1)	▲
2 3	2 3	2 3		USA	392	(4.9)	–	484	(3.2)	–	597	(3.4)		668	(3.2)		387	(5.6)	–	486	(3.6)	–	598	(2.2)		664	(2.6)	+	360	(4.0)	■	475	(3.0)	▲	598	(2.7)	▲	672	(3.1)	▲
		2		Tschechien	386	(8.9)		467	(4.4)	+	567	(3.8)		635	(4.3)		414	(5.1)		490	(3.4)	–	582	(2.2)	–	640	(3.2)		404	(4.8)	▲	480	(2.6)	▲	574	(2.7)	■	638	(4.3)	■
		2		Ungarn	383	(9.4)		485	(3.8)	–	595	(3.4)	–	661	(4.7)		395	(5.6)	–	479	(3.9)		584	(3.2)		647	(3.2)		370	(7.7)	■	472	(5.0)	■	583	(3.3)	▲	651	(4.8)	■
	3	3	8	Dänemark	383	(6.5)		468	(2.6)		570	(5.1)		636	(4.7)		405	(6.6)		478	(3.1)		569	(2.7)		629	(3.4)		391	(3.3)	▲	474	(3.2)	▲	574	(3.3)	■	639	(5.1)	■
	2			Slowakei	376	(13.9)		476	(4.7)		584	(2.8)		652	(6.9)		367	(13.8)		479	(4.4)		575	(2.6)		634	(4.1)		353	(9.8)	■	474	(4.8)	■	581	(2.9)	■	641	(3.6)	■
1 2	1 2	1 2 3	8	Neuseeland	344	(4.9)	+	447	(3.7)	+	568	(2.9)	+	643	(2.6)		360	(4.0)		448	(3.2)	+	561	(2.5)	+	629	(3.7)	+	364	(7.0)	■	461	(4.1)	■	579	(3.4)	■	648	(3.0)	■
	3	3	8	Niederlande	421	(6.3)	–	484	(3.6)	–	565	(3.9)		617	(5.0)		407	(5.4)		476	(3.9)		564	(2.8)		619	(5.7)		403	(5.9)	▲	473	(3.9)	▲	563	(2.2)	▼	621	(3.7)	▼
				Deutschland	393	(7.6)	–	479	(4.2)	–	582	(1.6)	–	647	(6.0)		383	(4.1)		468	(3.4)		573	(2.7)		634	(4.8)		369	(6.9)	■	463	(3.5)	■	574	(2.5)	■	644	(3.7)	■
2		2	8	Italien	395	(5.0)		484	(5.2)	–	590	(2.9)	–	664	(4.4)	–	400	(5.7)		467	(3.1)		555	(3.7)		615	(7.0)		389	(4.9)	▲	465	(3.6)	■	559	(2.5)	▼	621	(3.1)	▼
2	2	2		Georgien	273	(8.6)	+	361	(5.2)	+	477	(4.4)	+	552	(7.9)	+	307	(10.0)	+	400	(5.0)	+	512	(3.6)		582	(5.1)		341	(5.3)	▼	418	(4.5)	▼	513	(2.9)	▼	582	(6.4)	▼
			8	Iran	267	(7.1)	–	371	(6.5)		506	(5.3)		587	(3.3)		265	(10.4)	–	377	(6.0)	–	511	(3.9)		587	(4.4)		237	(7.5)	▼	358	(5.9)	▼	514	(3.4)	▼	594	(3.5)	▼
				Benchmark-Teilnehmer																																			
2 3 6	2		8	Dubai, VAE	267	(7.6)	+	390	(5.2)	+	537	(3.3)	+	620	(3.3)	+	389	(5.6)		492	(2.8)	+	605	(1.9)	+	671	(2.9)	+	388	(5.1)	▲	508	(2.5)	▲	627	(1.8)	▲	702	(4.6)	▲
2		2	8	Ontario, Kanada	396	(8.1)		487	(3.0)		590	(4.7)	–	657	(5.2)		397	(6.3)		474	(4.7)		577	(3.8)		641	(5.3)		395	(3.1)	▲	477	(4.3)	▲	576	(2.7)	■	645	(4.6)	■
2		3	8	Québec, Kanada	405	(6.8)	–	474	(3.6)	–	563	(3.4)		623	(9.2)		413	(5.3)	–	479	(4.4)	–	566	(3.2)	–	627	(4.7)		387	(5.7)	■	461	(3.8)	■	556	(2.8)	▼	619	(2.6)	▼

▲ = Punktwerte in den Perzentilen im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Punktwerte in den Perzentilen im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu Punktwerten in den Perzentilen im Vergleich zu Deutschland nicht statistisch signifikant ($p > .05$)
\+ = Mittelwert in 2023 statistisch signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 statistisch signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007

© TIMSS 2023

Studienzyklen sehr ähnlich. Hier liegt der Kompetenzwert, den die fünf Prozent der leistungsstärksten Schüler:innen mindestens erreichen, bei 644 Punkten, in TIMSS 2019 waren es 634 Punkte und in TIMSS 2007 647 Punkte. Zusammengefasst zeigt sich also, dass die Leistungen in Deutschland heterogener geworden sind und sich die Streuung im unteren Leistungsbereich vergrößert hat. In den meisten anderen Teilnehmerstaaten ist die Streuung der Leistungen relativ stabil. Ähnliche Veränderungen wie in Deutschland zeigen sich z. B. in Spanien, Frankreich und Schweden.

4.5.2 Kompetenzstufen

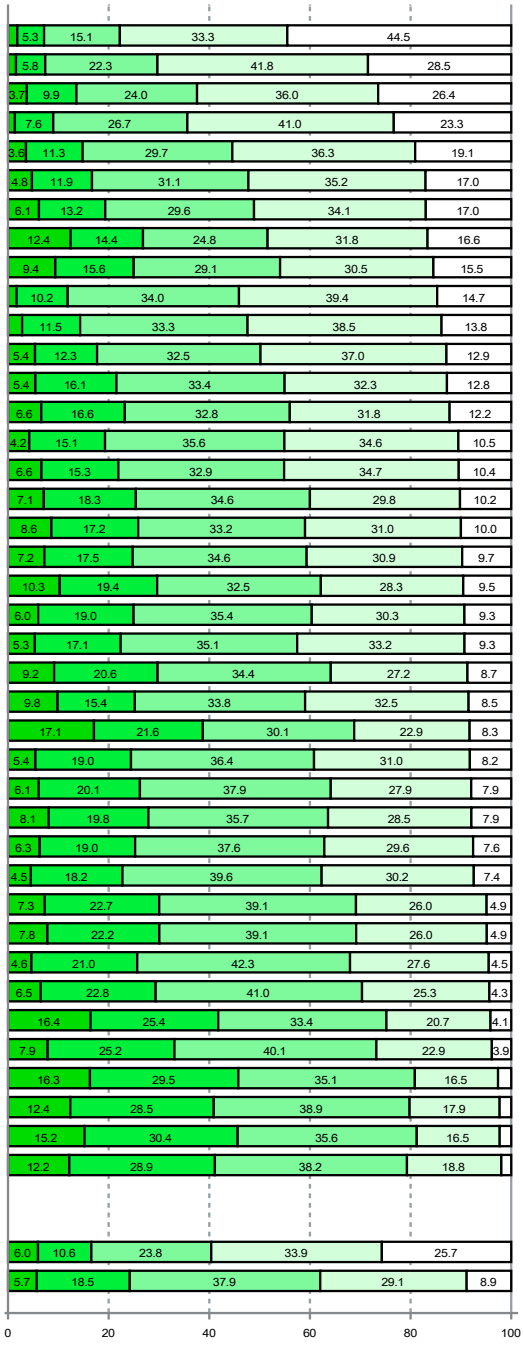
Die Kompetenzstufen leiten sich wie die Perzentile aus der Metrik der Gesamtskala Naturwissenschaften ab. Sie erweitern den Blick um inhaltliche Aspekte und lassen Aussagen darüber zu, welcher Anteil von Schüler:innen bestimmte Anforderungen einer Kompetenzstufe erfüllt. Die Erläuterungen zur inhaltlichen Bedeutung der fünf Kompetenzstufen, die von einem sporadischen Alltagswissen bis zu einem grundlegenden Verständnis von Inhalten sowie Denk- und Arbeitsweisen reichen, finden sich im Abschnitt 4.4.3. Die Abbildung 4.8 zeigt die Verteilung der Schüler:innen auf die fünf Kompetenzstufen für alle teilnehmenden Staaten. Die Kompetenzausprägungen, die auf den einzelnen Stufen zur erfolgreichen Lösung der entsprechenden Aufgaben erforderlich sind, bauen aufeinander auf. Schüler:innen der Kompetenzstufe IV lösen dementsprechend mit hinreichender Sicherheit Aufgaben der Stufen I bis IV. Entsprechend sind auf der rechten Seite der Abbildung 4.8 die kumulierten prozentualen Anteile der Kompetenzstufen dargestellt. Die Staaten sind entsprechend der Höhe der naturwissenschaftlichen Kompetenz in einer Rangreihe angeordnet.

Über alle teilnehmenden Staaten hinweg erreichen 15 Prozent der Schüler:innen lediglich die Kompetenzstufe I und befinden sich damit am untersten Ende des Leistungsspektrums. Diese Gruppe von Schüler:innen verfügt allenfalls über sporadisches naturwissenschaftliches Alltagswissen. Am anderen Ende des Leistungsspektrums auf der Kompetenzstufe V befinden sich 7 Prozent der Schüler:innen. Diese verfügen bereits über ein vertieftes Verständnis von Konzepten in allen Inhaltsbereichen und auch von Denk- und Arbeitsweisen. Die Unterschiede in den prozentualen Anteilen auf der höchsten und niedrigsten Kompetenzstufe zwischen den teilnehmenden Staaten sind enorm. So erreichen 45 Prozent der Schüler:innen in Singapur die Kompetenzstufe V und immerhin 29, 26, 23 bzw. 19 Prozent der Schüler:innen aus der Republik Korea (Südkorea), der Türkei, Taiwan bzw. England. In einigen anderen Staaten, z. B. Kosovo, Armenien oder Bosnien-Herzegowina, wird diese Kompetenzstufe von keinem Kind erreicht. In Staaten mit sehr niedrigen mittleren Kompetenzwerten erreichen allenfalls kleine Anteile die höchste Kompetenzstufe. Ein entsprechendes Bild zeigt sich auch in Bezug auf die Kompetenzstufe I, in die 2 Prozent (Singapur) bis 48 Prozent (Kosovo) der Schüler:innen eines Staates eingeordnet werden.

Der Anteil leistungsstarker Schüler:innen in Deutschland beträgt 9 Prozent und weicht damit nicht signifikant vom Durchschnitt der EU- und OECD-Staaten ab (8 bzw. 10 %). Einen vergleichbaren Anteil leistungsstarker Schüler:innen haben z. B. Dänemark, Norwegen, Kanada, Slowenien oder Lettland. Nimmt man die Kompetenzstufe IV hinzu, liegt der Anteil der Schüler:innen in Deutschland bei 36 Prozent, was ähnlich hoch wie in Kanada, Slowenien oder Neuseeland ist. Damit ist der Anteil höher als der internationale Durchschnitt (31 %) und ver-

Abbildung 4.8: Prozentuale Verteilung der Schüler:innen auf die fünf Kompetenzstufen (Naturwissenschaften) im internationalen Vergleich

	Anteil (%) derer, die mindestens folgende Kompetenzstufe erreichen							
	II		III		IV		V	
Teilnehmer	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
2 *Singapur*	98.2	(0.3)	92.8	(0.6)	77.8	(0.9)	44.5	(1.1)
Republik Korea (Südkorea)	98.4	(0.4)	92.6	(0.9)	70.3	(1.3)	28.5	(1.2)
1 2 3 *Türkei (5. Jgst.)*	96.3	(0.5)	86.4	(1.1)	62.4	(1.8)	26.4	(1.4)
Taiwan	98.6	(0.2)	91.0	(0.7)	64.4	(1.0)	23.3	(0.7)
1 2 3 *England*	96.4	(0.5)	85.2	(0.9)	55.4	(1.5)	19.1	(1.1)
Australien	95.2	(0.3)	83.3	(0.7)	52.2	(1.1)	17.0	(0.7)
3 *Hongkong*	93.9	(0.7)	80.7	(1.5)	51.1	(1.8)	17.0	(1.6)
Bulgarien	87.6	(1.7)	73.2	(1.9)	48.4	(1.6)	16.6	(1.2)
2 3 *USA*	90.6	(0.6)	75.0	(0.9)	45.9	(1.2)	15.5	(0.7)
Japan	98.3	(0.4)	88.1	(1.1)	54.1	(1.4)	14.7	(1.1)
2 *Polen*	97.2	(0.4)	85.7	(0.8)	52.4	(1.3)	13.8	(0.9)
Finnland	94.6	(0.6)	82.3	(1.3)	49.8	(1.4)	12.9	(1.0)
Macau	94.6	(0.3)	78.4	(0.5)	45.0	(0.8)	12.8	(0.4)
2 *Schweden*	93.4	(0.9)	76.8	(1.5)	44.0	(1.4)	12.2	(0.8)
2 *Litauen*	95.8	(0.5)	80.7	(1.5)	45.1	(1.3)	10.5	(0.7)
Irland	93.4	(0.8)	78.1	(1.3)	45.1	(1.9)	10.4	(0.8)
VG OECD	**92.9**	**(0.1)**	**74.6**	**(0.2)**	**40.0**	**(0.2)**	**10.2**	**(0.1)**
2 *Ungarn*	91.4	(1.0)	74.1	(1.5)	41.0	(1.4)	10.0	(0.8)
2 3 *Rumänien*	92.8	(1.3)	75.2	(2.4)	40.7	(2.3)	9.7	(1.3)
1 2 3 *Neuseeland*	89.7	(1.0)	70.4	(1.3)	37.8	(1.4)	9.5	(0.7)
Lettland	94.0	(0.8)	75.1	(1.4)	39.6	(1.7)	9.3	(0.8)
1 2 *Norwegen (5. Jgst.)*	94.7	(0.5)	77.6	(1.1)	42.5	(1.2)	9.3	(0.7)
Deutschland	90.8	(0.7)	70.3	(1.3)	35.9	(1.0)	8.7	(0.7)
Slowakei	90.2	(1.1)	74.8	(1.4)	41.0	(1.5)	8.5	(0.7)
Internationaler Mittelwert	**82.9**	**(0.1)**	**61.3**	**(0.2)**	**31.2**	**(0.2)**	**8.3**	**(0.1)**
2 *Slowenien*	94.6	(0.5)	75.6	(1.4)	39.2	(1.3)	8.2	(0.7)
2 3 *Kanada*	93.9	(0.4)	73.8	(0.9)	35.9	(1.0)	7.9	(0.4)
VG EU	**91.9**	**(0.2)**	**72.1**	**(0.3)**	**36.4**	**(0.3)**	**7.9**	**(0.2)**
2 3 *Dänemark*	93.7	(0.6)	74.7	(1.2)	37.1	(1.5)	7.6	(0.7)
2 *Tschechien*	95.5	(0.4)	77.3	(1.1)	37.7	(1.2)	7.4	(0.7)
Portugal	92.7	(0.6)	70.0	(1.2)	30.9	(1.2)	4.9	(0.5)
2 *Serbien*	92.2	(0.9)	69.9	(1.7)	30.8	(1.6)	4.9	(0.6)
3 *Niederlande*	95.4	(0.7)	74.3	(1.7)	32.0	(1.4)	4.5	(0.5)
2 *Italien*	93.5	(0.6)	70.7	(1.6)	29.7	(1.2)	4.3	(0.5)
2 *Zypern*	83.6	(1.1)	58.2	(1.4)	24.8	(1.1)	4.1	(0.4)
2 *Spanien*	92.1	(0.6)	66.9	(1.1)	26.8	(0.8)	3.9	(0.3)
2 3 *Chile*	83.7	(0.9)	54.2	(1.4)	19.2	(1.1)	2.6	(0.3)
2 *Frankreich*	87.6	(1.3)	59.1	(1.6)	20.2	(1.2)	2.3	(0.5)
2 *Belgien (Frz. Gem.)*	84.8	(1.0)	54.4	(1.7)	18.8	(0.9)	2.3	(0.4)
2 3 *Belgien (Fläm. Gem.)*	87.8	(1.2)	58.9	(1.5)	20.7	(1.0)	2.0	(0.3)
Benchmark-Teilnehmer								
Dubai, VAE	94.0	(0.4)	83.4	(0.6)	59.6	(0.9)	25.7	(0.8)
2 *Ontario, Kanada*	94.3	(0.5)	75.8	(1.5)	38.0	(1.6)	8.9	(0.7)



% der Schüler:innen, die Kompetenzstufe I erreichen
% der Schüler:innen, die Kompetenzstufe II erreichen
% der Schüler:innen, die Kompetenzstufe III erreichen
% der Schüler:innen, die Kompetenzstufe IV erreichen
% der Schüler:innen, die Kompetenzstufe V erreichen

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

© TIMSS 2023

gleichbar mit dem der anderen EU-Staaten (36 %). Die OECD-Staaten erreichen im Mittel höhere Prozentsätze (40 %). Die Gruppe der Schüler:innen auf den beiden höchsten Kompetenzstufen verfügt über (sehr) gute Voraussetzungen für die erfolgreiche Teilnahme am naturwissenschaftlichen Unterricht in den weiterführenden Schulen. Es sollte allerdings beachtet werden, dass es anderen Staaten gelingt, dass sogar ca. zwei Drittel der Schüler:innen die beiden höchsten Kompetenzstufen erreichen.

Am unteren Ende des Leistungsspektrums, auf der Kompetenzstufe I, befinden sich 9 Prozent der Schüler:innen in Deutschland. Ähnliche Anteile zeigen sich z. B. in Portugal, der Slowakei oder Neuseeland. Fasst man die beiden unteren Kompetenzstufen I und II zusammen, ergeben sich für Deutschland 30 Prozent (EU-Staaten: 28 %; OECD-Staaten: 25 %). Dieser relativ großen Gruppe (sehr) leistungsschwacher Schüler:innen fehlen zentrale Grundlagen für den anschließenden Biologie-, Chemie- und Physikunterricht sowie die naturwissenschaftlichen Anteile des Geografieunterrichts der weiterführenden Schulen. Anderen Staaten wie beispielsweise Polen, Finnland und der Türkei gelingt es besser, am unteren Ende des Leistungsspektrums zu fördern und die Anzahl der Schüler:innen auf den beiden unteren Kompetenzstufen kleinzuhalten. So finden sich hier lediglich ca. 15 Prozent auf den beiden untersten Kompetenzstufen.

Unterschiede in den Anteilen der Schüler:innen auf den Kompetenzstufen zwischen 2007, 2019 und 2023:

In Tabelle 4.9 sind die prozentualen Anteile der Schüler:innen auf den verschiedenen Kompetenzstufen bei den bisherigen TIMSS-Zyklen dargestellt, um so Unterschiede und möglicherweise Trends abschätzen zu können. In der Tabelle sind nur Staaten aufgeführt, die an allen drei Studienzyklen teilgenommen haben. Die Staaten sind entsprechend ihres Mittelwertes der naturwissenschaftlichen Kompetenz in einer Rangreihe angeordnet. Für Deutschland zeigt sich dabei folgendes Bild: Der Anteil der Schüler:innen auf den beiden oberen Kompetenzstufen IV und V hat sich im Vergleich zu TIMSS 2019 und 2007 um ca. 4 Prozentpunkte verringert. Gleichzeitig nimmt der Anteil der Schüler:innen auf den beiden unteren Kompetenzstufen zu. Diese Veränderungen sind eher klein (1–3 %) und nur teilweise statistisch bedeutsam. Gleichwohl deuten sie in der Tendenz auf eine ungünstige Entwicklung hin. So scheint es etwas schlechter zu gelingen, die naturwissenschaftliche Kompetenz leistungsschwacher Schüler:innen zu fördern. Auch in den meisten anderen Staaten zeigen sich im Hinblick auf die Verteilung auf die Kompetenzstufen nur geringfügige Verschiebungen von wenigen Prozenten. Staaten, in denen sich substanzielle Änderungen in der Verteilung zeigen, sind vor allem solche, in denen sich auch das Niveau deutlich verändert hat. Hierzu gehören beispielsweise England und Hongkong, wo ca. 9 Prozent mehr Schüler:innen als in den früheren Studienzyklen die Kompetenzstufe IV (hoch) erreichen.

Tabelle 4.9: Prozentuale Verteilung der Schüler:innen auf die Kompetenzstufen in TIMSS 2007, 2019 und 2023

Fußnoten				Teilnehmer	TIMSS 2007 Kompetenzstufen										TIMSS 2019 Kompetenzstufen										TIMSS 2023 Kompetenzstufen									
2007	2019	2023	T		I		II		III		IV		V		I		II		III		IV		V		I		II		III		IV		V	
					%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
	2	2	8	Singapur	3.8	(0.5)	– 7.9	(0.8)	– 20.1	(1.2)	– 32.2	(1.4)	35.9	(1.9)	+ 1.8	(0.4)	5.6	(0.6)	18.3	(1.2)	– 36.7	(1.3)	– 37.6	(1.9)	+ 1.8	(0.3)	5.3	(0.4)	15.1	(0.7)	33.3	(1.0)	44.5	(1.1)
				Taiwan	3.0	(0.4)	– 11.0	(0.6)	– 30.9	(1.1)	– 36.4	(1.3)	+ 18.7	(1.0)	+ 1.4	(0.3)	9.3	(0.9)	31.8	(1.1)	– 42.4	(1.1)	15.0	(0.9)	+ 1.4	(0.2)	7.6	(0.6)	26.7	(0.8)	41.0	(0.9)	23.3	(0.7)
1	3	1 2	1 2 3	8 England	4.7	(0.6)	14.7	(0.8)	– 32.7	(1.2)	33.5	(1.1)	14.2	(1.2)	+ 3.7	(0.6)	14.9	(0.9)	– 37.2	(1.3)	– 33.9	(1.6)	10.3	(1.1)	+ 3.6	(0.5)	11.3	(0.7)	29.7	(1.2)	36.3	(1.3)	19.1	(1.1)
		3		8 Japan	2.7	(0.4)	11.3	(0.9)	35.2	(0.9)	38.4	(0.9)	12.4	(1.0)	1.8	(0.4)	8.5	(0.7)	30.3	(1.0)	+ 42.0	(1.1)	17.4	(0.8)	1.7	(0.4)	10.2	(1.0)	34.0	(1.2)	39.4	(1.0)	14.7	(1.1)
				Australien	6.7	(0.8)	– 16.9	(1.1)	– 35.5	(1.3)	– 30.6	(2.1)	+ 10.2	(0.7)	+ 6.2	(0.7)	15.5	(1.0)	– 34.4	(1.3)	– 32.8	(1.3)	11.1	(0.9)	+ 4.8	(0.3)	11.9	(0.6)	31.1	(0.8)	35.2	(0.9)	17.0	(0.7)
2 3		3	3 8	Hongkong	2.0	(0.4)	+ 10.1	(1.0)	+ 33.1	(1.4)	– 40.7	(1.5)	– 14.1	(1.4)	3.9	(0.6)	+ 16.8	(1.4)	– 38.3	(1.3)	– 32.5	(1.6)	8.5	(0.9)	+ 6.1	(0.7)	13.2	(1.0)	29.6	(1.1)	34.1	(1.5)	17.0	(1.6)
2		2	2 8	Litauen	4.9	(0.6)	21.2	(1.2)	– 43.6	(1.3)	– 27.0	(1.3)	+ 3.3	(0.4)	+ 3.5	(0.4)	15.1	(1.2)	36.0	(1.2)	34.6	(1.3)	10.9	(0.9)	4.2	(0.5)	15.1	(1.2)	35.6	(1.2)	34.6	(1.0)	10.5	(0.7)
		2	2	Schweden	5.0	(0.6)	19.0	(1.1)	38.5	(1.0)	– 29.4	(1.3)	8.0	(0.6)	+ 4.0	(0.6)	+ 15.7	(1.2)	35.2	(1.1)	33.9	(1.4)	11.2	(1.0)	6.6	(0.9)	16.6	(1.0)	32.8	(1.4)	31.8	(1.1)	12.2	(0.8)
2 3		2 3	2 3	USA	6.0	(0.6)	+ 16.1	(0.7)	31.2	(0.9)	31.8	(0.9)	15.0	(0.9)	6.5	(0.6)	+ 15.0	(0.7)	30.6	(0.8)	33.3	(1.0)	– 14.7	(0.8)	9.4	(0.6)	15.6	(0.7)	29.1	(0.7)	30.5	(0.7)	15.5	(0.7)
			2	Tschechien	7.0	(0.8)	– 21.5	(1.1)	– 38.5	(1.5)	26.1	(1.7)	+ 7.0	(0.7)	3.4	(0.5)	15.9	(0.9)	37.9	(1.9)	34.5	(1.9)	– 8.3	(0.9)	4.5	(0.4)	18.2	(0.9)	39.6	(0.9)	30.2	(0.9)	7.4	(0.7)
			2	Ungarn	6.6	(0.8)	15.1	(1.3)	31.1	(1.5)	33.7	(1.5)	13.4	(1.0)	– 5.7	(0.7)	+ 17.9	(1.0)	34.5	(1.1)	32.2	(1.0)	9.8	(0.6)	8.6	(1.0)	17.2	(1.2)	33.2	(1.0)	31.0	(1.1)	10.0	(0.8)
	3		3 2 3	8 Dänemark	7.1	(0.8)	20.8	(1.0)	37.6	(1.6)	27.7	(1.1)	6.9	(0.8)	4.5	(0.5)	+ 19.2	(1.1)	40.8	(1.1)	– 29.9	(1.0)	5.7	(0.7)	6.3	(0.6)	19.0	(0.9)	37.6	(1.1)	29.6	(1.1)	7.6	(0.7)
		2		Slowakei	7.5	(1.3)	17.0	(1.2)	33.7	(1.1)	31.1	(1.6)	10.6	(0.8)	– 8.0	(1.2)	15.5	(1.1)	37.9	(1.1)	– 31.7	(1.2)	6.9	(0.8)	9.8	(1.1)	15.4	(1.0)	33.8	(1.1)	32.5	(1.1)	8.5	(0.7)
1 2		1 2	1 2 3	8 Neuseeland	13.4	(1.0)	– 21.7	(0.8)	32.7	(0.9)	24.5	(0.9)	+ 7.8	(0.5)	+ 11.6	(0.8)	24.0	(0.9)	– 34.5	(1.4)	24.3	(1.1)	+ 5.6	(0.5)	+ 10.3	(1.0)	19.4	(1.1)	32.5	(1.0)	28.3	(1.0)	9.5	(0.7)
	3		3	3 8 Niederlande	2.5	(0.5)	+ 18.2	(1.3)	45.4	(1.4)	30.2	(1.5)	3.7	(0.8)	4.2	(0.6)	20.1	(1.4)	43.0	(1.3)	28.8	(1.8)	3.9	(0.9)	4.6	(0.7)	21.0	(1.4)	42.3	(1.4)	27.6	(1.3)	4.5	(0.5)
				Deutschland	5.8	(0.6)	+ 17.9	(0.9)	+ 35.4	(0.7)	31.3	(1.1)	– 9.6	(0.7)	7.2	(0.7)	+ 20.5	(1.0)	35.9	(1.0)	29.6	(1.3)	6.9	(0.9)	9.2	(0.7)	20.6	(1.0)	34.4	(1.0)	27.2	(0.9)	8.7	(0.7)
2			2 8	Italien	5.5	(0.7)	16.1	(0.9)	+ 34.5	(1.0)	+ 30.9	(1.2)	– 13.0	(1.0)	– 5.0	(0.8)	23.9	(1.1)	43.6	(1.1)	24.2	(1.7)	3.3	(0.7)	6.5	(0.6)	22.8	(1.3)	41.0	(1.2)	25.3	(1.1)	4.3	(0.5)
2		2	2	Georgien	41.1	(2.1)	– 32.9	(1.5)	20.6	(1.7)	+ 4.9	(0.7)	+ 0.5	(0.2)	25.1	(1.8)	– 32.2	(1.1)	+ 30.9	(1.7)	10.3	(1.1)	1.4	(0.4)	18.0	(1.3)	36.6	(1.2)	34.1	(1.2)	9.9	(1.0)	1.4	(0.5)
			8	Iran	34.7	(1.9)	29.0	(1.1)	– 24.6	(1.2)	10.1	(0.9)	1.6	(0.3)	31.8	(1.8)	+ 28.5	(1.1)	– 26.8	(1.1)	11.5	(0.9)	1.3	(0.3)	36.6	(1.5)	25.0	(0.7)	24.2	(1.0)	12.4	(0.8)	1.8	(0.3)
				Benchmark-Teilnehmer																														
2 3 6		2	8	Dubai, VAE	27.8	(1.4)	– 24.2	(1.3)	– 27.4	(1.2)	– 16.3	(0.8)	+ 4.2	(0.5)	+ 6.2	(0.5)	13.5	(0.6)	– 28.4	(0.7)	– 35.3	(1.0)	16.5	(0.8)	+ 6.0	(0.4)	10.6	(0.6)	23.8	(0.8)	33.9	(1.0)	25.7	(0.8)
2			2 8	Ontario, Kanada	5.4	(1.0)	15.6	(1.1)	33.6	(1.4)	+ 33.5	(2.0)	11.9	(1.2)	– 5.3	(0.8)	19.9	(1.6)	36.7	(1.5)	30.1	(1.3)	7.9	(1.0)	5.7	(0.5)	18.5	(1.3)	37.9	(1.2)	29.1	(1.4)	8.9	(0.7)
2		3	2 3 8	Québec, Kanada	4.3	(0.6)	+ 21.4	(1.5)	42.3	(1.9)	27.2	(1.8)	4.8	(0.6)	3.4	(0.5)	+ 19.8	(1.5)	+ 42.6	(1.6)	29.0	(1.3)	– 5.3	(0.7)	7.0	(0.6)	24.3	(1.0)	40.4	(1.2)	24.1	(1.2)	4.2	(0.3)

\+ = Prozentwert in 2023 statistisch signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Prozentwert in 2023 statistisch signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007

© TIMSS 2023

4.5.3 Inhaltsbereiche

Die naturwissenschaftlichen Kompetenzen werden in der TIMSS-Rahmenkonzeption in die drei Inhaltsbereiche *Biologie, Physik/Chemie* und *Geografie* aufgeteilt. Dies erlaubt einen genaueren Blick auf die spezifischen Stärken und Schwächen der Schüler:innen in den drei Bereichen, die auch in der Sekundarstufe I durch die Fächer Biologie, Geografie und Physik/Chemie abgebildet werden. Aus nationaler Perspektive ist diese Differenzierung wichtig, um abzuschätzen, inwiefern die drei Inhaltsbereiche gleichmäßig im Sachunterricht behandelt werden und so vergleichbare Leistungen erreicht werden.

Kompetenzen in den drei Inhaltsgebieten im internationalen Vergleich:

In Abbildung 4.9 sind die Testleistungen (Mittelwerte und Standardabweichungen sowie deren Standardfehler) getrennt nach den Inhaltsbereichen *Biologie, Physik/Chemie* und *Geografie* aufgeführt. Berücksichtigt sind hier nur OECD- und EU-Staaten sowie Staaten, die signifikant höhere Kompetenzen als oder vergleichbare Kompetenzen wie Deutschland auf der Gesamtskala aufweisen (siehe Kapitel 2 in diesem Band für die Vergleichsgruppen). Die Anordnung entspricht der Rangfolge auf der Gesamtskala Naturwissenschaften.

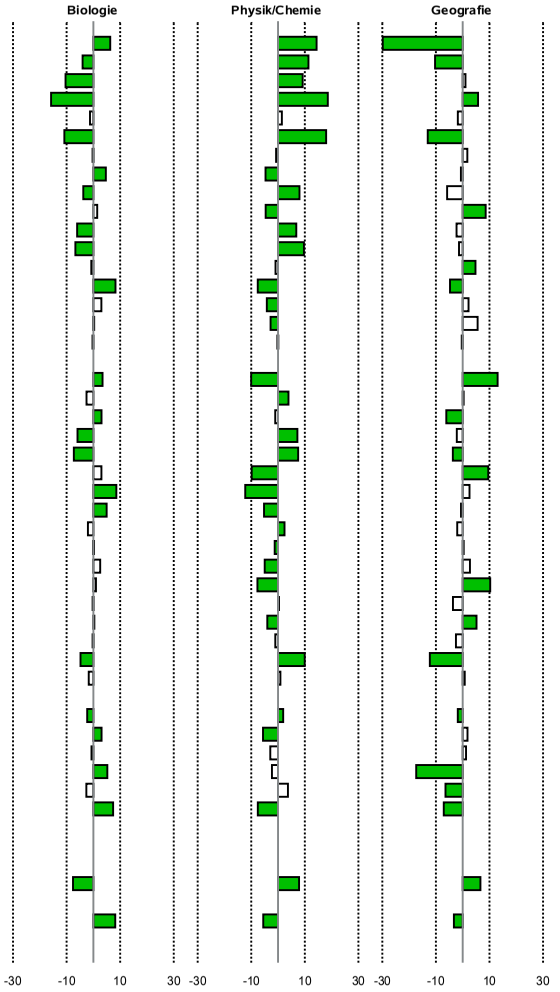
Werden jeweils die Ergebnisse in den drei Inhaltsbereichen mit der Gesamtskala verglichen, so zeigt sich über alle berücksichtigten Staaten kein einheitliches Muster der Abweichungen vom Gesamtwert. Vielfach lassen sich einzelne positive und negative Abweichungen feststellen, die maximal 30 Punkte betragen. Generelle Tendenzen, dass z. B. im Inhaltsbereich *Biologie* höhere Werte erzielt werden als im Bereich *Physik/Chemie*, sind nicht erkennbar. Für Deutschland zeigt sich, dass die Werte in *Biologie* sowie *Physik/Chemie* exakt dem Gesamtwert für die Naturwissenschaften entsprechen (M = 515), lediglich in *Geografie* fällt der Wert mit M = 512 etwas niedriger aus. Diese Abweichung ist aber nicht statistisch signifikant, sodass insgesamt keine Unterschiede zwischen den drei Inhaltsbereichen in Deutschland nachweisbar sind. Die etwas größere Standardabweichung im Bereich *Geografie* (SD = 86) könnte damit verbunden sein, dass, wie eingangs erläutert, naturgeografische Themen im Sachunterricht weniger in den Lehrplänen und damit auch im Unterricht verankert sind. Betrachtet man schließlich die mittleren Testleistungen der EU- und OECD-Staaten in den drei Inhaltsbereichen, so lassen sich auch hier keine substanziellen Abweichungen zum jeweiligen Gesamtwert in den Naturwissenschaften nachweisen.

Unterschiede in den Kompetenzen in den naturwissenschaftlichen Inhaltsbereichen zwischen TIMSS 2007, 2019 und 2023:

Ähnlich wie bei den Niveauunterschieden in den einzelnen Inhaltsbereichen (siehe Abbildung 4.10) zeigt sich auch bei der Betrachtung der Veränderungen seit 2007 kein einheitliches Muster. In den meisten Staaten sind die relativen Schwächen und Stärken eher klein. Existieren stärkere Abweichungen vom Gesamtmittelwert, so zeigt sich dies typischerweise in allen drei betrachteten Studienzyklen. Weiterhin zeigt sich, dass Veränderungen im Gesamtwert seit 2007 mit entsprechenden Veränderungen in allen drei Inhaltsbereichen einhergehen, d. h., steigt der Gesamtwert seit 2007 in einem Staat an, so zeigt sich dieser Trend ebenso in *Biologie*, *Physik/Chemie* und *Geografie*. Große positive oder negative Veränderungen, die nur einen Inhaltsbereich betreffen, sind dagegen nicht erkennbar.

Abbildung 4.9: Leistungen der Schüler:innen in den Inhaltsbereichen *Biologie*, *Physik/Chemie* und *Geografie* im internationalen Vergleich

	Gesamtskala				Biologie^						Physik/Chemie^						Geografie^					
Teilnehmer	M_{GS}	*(SE)*	*SD*	*(SE)*	M_B	*(SE)*		*SD*	*(SE)*	$M_B - M_{GS}$	*(SE)*	M_P	*(SE)*		*SD*	*(SE)*	$M_P - M_{GS}$	*(SE)*	M_G	*(SE)*		*SD*	*(SE)*	$M_G - M_{GS}$	*(SE)*
2 *Singapur*	607	(2.3)	84	(1.3)	614	(2.5)	+	91	(1.7)	6	(0.6)	622	(2.3)	+	94	(1.9)	14	(1.1)	578	(2.3)	+	87	(1.9)	-30	(1.3)
Republik Korea (Südkorea)	583	(2.5)	74	(2.2)	579	(2.8)	+	78	(2.6)	-4	(1.5)	595	(2.6)	+	77	(2.6)	11	(1.4)	573	(3.5)	+	85	(2.9)	-10	(2.1)
Taiwan	573	(1.4)	71	(1.4)	562	(1.9)	+	78	(1.4)	-10	(1.1)	582	(1.4)	+	70	(1.3)	9	(0.8)	574	(2.2)	+	80	(1.4)	1	(1.5)
123 *Türkei (5. Jgst.)*	570	(3.4)	86	(1.7)	554	(3.6)	+	92	(2.0)	-16	(1.0)	589	(3.9)	+	92	(2.0)	19	(1.9)	576	(4.5)	+	93	(2.5)	6	(2.8)
123 *England*	556	(2.6)	80	(2.0)	555	(3.0)	+	82	(1.8)	-1	(1.7)	558	(3.1)	+	85	(1.8)	1	(1.5)	554	(3.5)	+	90	(1.9)	-2	(1.8)
Japan	555	(2.4)	68	(1.4)	544	(3.1)	+	75	(1.2)	-11	(2.5)	573	(2.9)	+	70	(1.2)	18	(1.7)	542	(2.4)	+	75	(1.4)	-13	(1.5)
2 *Polen*	550	(2.2)	71	(1.1)	550	(2.7)	+	73	(1.2)	0	(1.6)	549	(2.3)	+	77	(1.5)	-1	(1.0)	552	(2.7)	+	80	(1.4)	2	(2.8)
Australien	550	(1.8)	81	(0.9)	554	(2.1)	+	82	(1.5)	5	(2.1)	545	(1.9)	+	83	(1.5)	-5	(1.5)	549	(2.3)	+	92	(2.7)	-1	(2.6)
3 *Hongkong*	545	(3.8)	88	(2.7)	541	(4.0)	+	98	(2.7)	-4	(1.9)	553	(3.9)	+	91	(2.7)	8	(1.7)	539	(4.9)	+	96	(3.0)	-6	(3.5)
Finnland	542	(2.8)	79	(1.8)	543	(2.8)	+	83	(1.9)	1	(1.5)	537	(2.7)	+	79	(1.9)	-5	(0.7)	551	(2.9)	+	92	(2.1)	9	(1.8)
2 *Litauen*	537	(2.4)	74	(1.6)	531	(2.2)	+	77	(1.3)	-6	(1.7)	544	(2.5)	+	75	(1.2)	7	(1.6)	535	(2.9)	+	88	(2.8)	-2	(1.2)
Macau	536	(1.3)	81	(0.8)	529	(1.3)	+	84	(1.5)	-7	(1.0)	545	(1.1)	+	88	(1.5)	10	(1.4)	534	(2.2)	+	92	(1.0)	-1	(1.9)
2 *Schweden*	533	(3.1)	82	(2.0)	532	(3.3)	+	85	(2.0)	-1	(1.1)	532	(3.2)	+	84	(2.1)	-1	(1.6)	537	(4.1)	+	95	(2.4)	5	(1.9)
23 *USA*	532	(2.4)	93	(0.9)	541	(2.4)	+	94	(0.9)	8	(1.1)	525	(2.9)	+	96	(0.9)	-8	(1.3)	528	(3.1)	+	105	(0.9)	-5	(1.7)
Irland	532	(3.2)	80	(1.6)	535	(3.6)	+	88	(1.9)	3	(1.7)	528	(3.4)	+	80	(2.0)	-4	(1.5)	534	(4.1)	+	88	(2.5)	2	(2.7)
Bulgarien	530	(4.8)	104	(4.1)	530	(5.0)	+	108	(4.5)	0	(1.5)	527	(4.9)	+	109	(4.6)	-3	(1.3)	535	(6.0)	+	108	(4.5)	6	(3.6)
VG OECD	**526**	**(0.5)**	**77**	**(0.3)**	**525**	**(0.5)**	+	**81**	**(0.3)**	**0**	**(0.3)**	**525**	**(0.5)**	+	**79**	**(0.3)**	**0**	**(0.3)**	**525**	**(0.6)**	+	**88**	**(0.4)**	**0**	**(0.4)**
12 *Norwegen (5. Jgst.)*	530	(2.5)	75	(1.1)	534	(3.0)	+	76	(1.4)	3	(1.5)	520	(2.8)		81	(1.1)	-10	(1.1)	543	(2.9)	+	84	(1.5)	13	(1.4)
23 *Rumänien*	526	(4.8)	81	(2.4)	524	(5.1)		81	(2.2)	-3	(1.6)	530	(5.0)	+	85	(2.2)	4	(1.9)	526	(5.2)	+	92	(3.6)	0	(2.6)
2 *Tschechien*	526	(2.2)	71	(1.2)	529	(1.9)	+	74	(1.1)	3	(0.8)	525	(2.3)	+	74	(1.3)	-1	(1.1)	520	(2.9)		83	(1.2)	-6	(2.0)
2 *Slowenien*	526	(2.3)	74	(1.4)	520	(2.3)		77	(1.1)	-6	(0.9)	533	(2.3)	+	76	(1.2)	7	(1.7)	523	(2.8)	+	84	(1.9)	-2	(1.6)
Lettland	526	(2.9)	77	(1.9)	518	(3.3)		81	(2.2)	-7	(1.4)	533	(3.6)	+	78	(1.6)	8	(2.0)	522	(3.7)	+	95	(2.3)	-4	(1.6)
2 *Ungarn*	524	(3.1)	85	(2.5)	527	(3.3)	+	88	(2.7)	3	(1.7)	514	(3.1)		86	(2.3)	-10	(1.1)	534	(3.8)	+	98	(2.6)	10	(2.3)
23 *Dänemark*	522	(2.5)	75	(1.2)	531	(2.6)	+	79	(1.4)	9	(1.5)	510	(2.9)		79	(1.3)	-12	(1.2)	525	(2.9)	+	83	(2.0)	3	(1.3)
23 *Kanada*	521	(1.7)	76	(1.1)	526	(2.0)	+	77	(0.9)	5	(1.3)	516	(1.4)		78	(0.7)	-5	(1.2)	520	(2.3)	+	84	(1.4)	-1	(1.4)
Slowakei	521	(3.3)	87	(2.5)	519	(3.7)		89	(2.7)	-2	(1.4)	523	(3.6)		88	(2.7)	2	(1.1)	519	(3.7)		101	(2.6)	-2	(2.0)
VG EU	**518**	**(0.6)**	**78**	**(0.4)**	**518**	**(0.7)**		**82**	**(0.4)**	**0**	**(0.3)**	**517**	**(0.7)**		**80**	**(0.4)**	**-1**	**(0.3)**	**518**	**(0.8)**	+	**89**	**(0.5)**	**0**	**(0.5)**
123 *Neuseeland*	517	(2.8)	87	(1.7)	520	(2.9)		91	(1.6)	3	(1.9)	512	(3.0)		89	(1.6)	-5	(1.5)	520	(2.4)	+	93	(2.0)	3	(1.9)
3 *Niederlande*	517	(2.9)	67	(1.5)	518	(3.2)		69	(2.3)	1	(1.2)	509	(2.5)		68	(2.0)	-8	(2.4)	527	(2.9)	+	79	(1.9)	10	(3.3)
Deutschland	515	(2.6)	84	(1.7)	515	(2.5)		90	(1.8)	0	(0.9)	515	(2.6)		83	(1.3)	0	(1.4)	512	(3.2)		91	(1.7)	-4	(2.8)
Portugal	511	(2.3)	73	(1.1)	511	(2.8)		80	(1.5)	0	(1.4)	507	(2.7)	–	73	(1.4)	-4	(1.5)	516	(2.9)		81	(2.0)	5	(1.5)
2 *Italien*	511	(2.5)	70	(1.1)	510	(2.9)		74	(1.3)	0	(1.1)	510	(2.6)		69	(1.6)	-1	(1.1)	508	(3.1)		83	(1.4)	-3	(2.0)
2 *Serbien*	510	(3.2)	74	(1.6)	505	(3.9)	–	80	(2.1)	-5	(2.2)	520	(3.4)		76	(2.2)	10	(2.5)	498	(3.1)	–	83	(3.4)	-12	(2.3)
2 *Spanien*	504	(1.9)	72	(0.9)	502	(2.2)	–	73	(1.0)	-2	(0.9)	505	(2.1)	–	75	(1.3)	1	(1.4)	505	(2.4)		86	(1.2)	1	(1.9)
Internationaler Mittelwert	**494**	**(0.4)**	**85**	**(0.2)**	**495**	**(0.4)**	–	**87**	**(0.3)**	**-2**	**(0.2)**	**500**	**(0.4)**	–	**88**	**(0.3)**	**2**	**(0.2)**	**496**	**(0.5)**	–	**93**	**(0.3)**	**-2**	**(0.3)**
23 *Belgien (Fläm. Gem.)*	488	(2.6)	73	(1.8)	491	(2.6)	–	78	(1.8)	3	(1.0)	483	(2.8)	–	78	(1.7)	-6	(1.3)	490	(3.4)	–	76	(2.9)	2	(2.8)
2 *Frankreich*	488	(2.9)	74	(1.6)	487	(3.4)	–	79	(1.8)	-1	(1.4)	485	(3.0)	–	71	(1.8)	-3	(2.0)	489	(4.2)	–	95	(1.8)	1	(2.2)
2 *Zypern*	487	(3.0)	88	(1.6)	492	(3.1)	–	88	(1.6)	5	(1.1)	485	(3.4)	–	98	(1.9)	-2	(1.4)	470	(3.7)	–	99	(2.4)	-17	(1.9)
2 *Belgien (Frz. Gem.)*	481	(2.6)	77	(1.6)	478	(2.7)	–	85	(1.6)	-3	(1.5)	484	(2.0)	–	76	(1.3)	4	(2.1)	474	(3.2)	–	84	(2.1)	-6	(2.2)
23 *Chile*	479	(2.5)	80	(1.4)	487	(2.3)	–	82	(1.5)	7	(0.9)	472	(2.6)	–	85	(1.5)	-8	(2.0)	472	(3.1)	–	93	(1.4)	-7	(1.3)
Benchmark-Teilnehmer																									
Dubai, VAE	562	(1.7)	95	(1.0)	555	(1.3)	+	93	(1.3)	-8	(1.4)	570	(2.0)	+	104	(1.1)	8	(1.2)	569	(2.4)	+	98	(1.2)	7	(1.6)
2 *Ontario, Kanada*	525	(2.6)	76	(1.2)	533	(2.7)	+	77	(1.1)	8	(1.4)	520	(2.3)		78	(1.0)	-6	(1.0)	522	(3.3)	+	85	(1.8)	-3	(1.6)

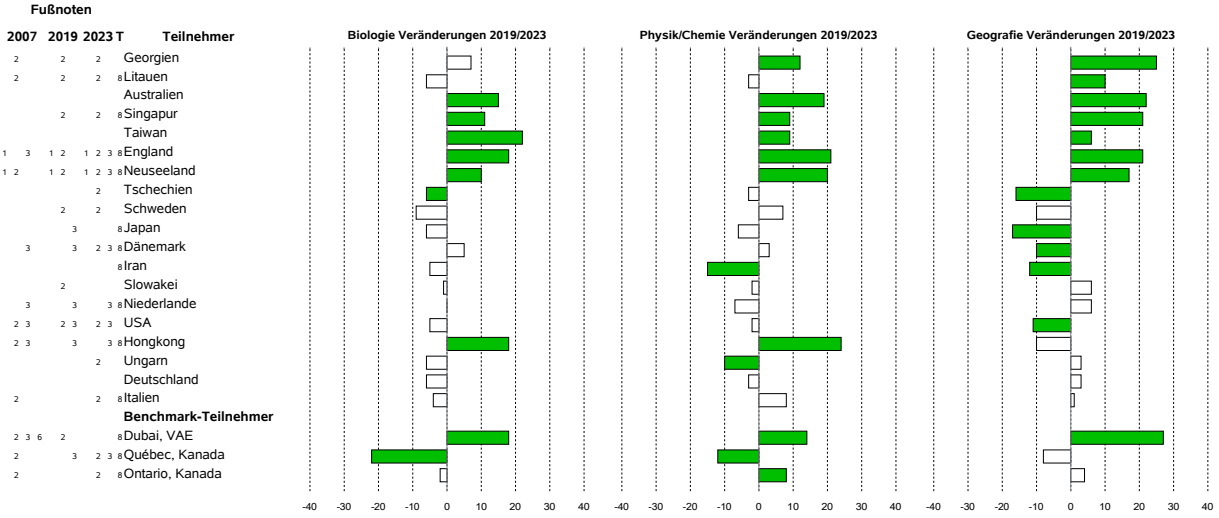


Keine statistisch signifikanten Unterschiede zur Gesamtskala ($p > .05$)
Statistisch signifikante Unterschiede zur Gesamtskala ($p \leq .05$)
\+ = Mittelwert statistisch signifikant höher als der Mittelwert von Deutschland ($p \leq .05$)
– = Mittelwert statistisch signifikant niedriger als der Mittelwert von Deutschland ($p \leq .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Abbildung 4.10: Leistungen in den Inhaltsbereichen *Biologie, Physik/Chemie* und *Geografie* in TIMSS 2007, 2019 und 2023

Fußnoten 2007	2019	2023	T	Teilnehmer	Biologie Veränderungen^ M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		Physik/Chemie Veränderungen^ M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		Geografie Veränderungen^ M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)	
2	2	2		Georgien	464	(3.5)	43	(5.3)	+	7	(5.3)		464	(3.3)	61	(5.9)	+	12	(5.7)	+	460	(4.5)	44	(7.0)	+	25	(6.2)	+
2	2	2	8	Litauen	531	(2.2)	13	(3.1)	+	-6	(3.6)		544	(2.5)	33	(3.2)	+	-3	(3.9)		535	(2.9)	27	(4.1)	+	10	(4.2)	+
				Australien	554	(2.1)	25	(4.2)	+	15	(3.5)	+	545	(1.9)	24	(4.2)	+	19	(3.3)	+	549	(2.3)	13	(5.1)	+	22	(3.6)	+
	2	2	8	Singapur	614	(2.5)	19	(5.3)	+	11	(4.4)	+	622	(2.3)	25	(4.9)	+	9	(4.4)	+	578	(2.3)	13	(4.6)	+	21	(4.5)	+
				Taiwan	562	(1.9)	15	(3.4)	+	22	(2.8)	+	582	(1.4)	18	(2.9)	+	9	(2.4)	+	574	(2.2)	11	(3.6)	+	6	(2.8)	+
1 3	1 2	1 2 3	8	England	555	(3.0)	19	(4.3)	+	18	(4.0)	+	558	(3.1)	12	(4.5)	+	21	(4.5)	+	554	(3.5)	12	(4.9)	+	21	(4.5)	+
1 2	1 2	1 2 3	8	Neuseeland	520	(2.9)	14	(4.0)	+	10	(3.7)	+	512	(3.0)	18	(4.5)	+	20	(3.7)	+	520	(2.4)	7	(4.2)		17	(3.9)	+
		2		Tschechien	529	(1.9)	7	(4.0)		-6	(2.9)	–	525	(2.3)	16	(4.1)	+	-3	(3.4)		520	(2.9)	6	(4.5)		-16	(4.2)	–
	2	2		Schweden	532	(3.3)	0	(4.3)		-9	(4.7)		532	(3.2)	23	(4.5)	+	7	(4.6)		537	(4.1)	-2	(5.5)		-10	(5.6)	
	3		8	Japan	544	(3.1)	8	(3.9)	+	-6	(3.7)		573	(2.9)	2	(4.1)		-6	(3.5)		542	(2.4)	10	(4.2)	+	-17	(3.1)	–
3	3	2 3	8	Dänemark	531	(2.6)	4	(4.1)		5	(3.4)		510	(2.9)	8	(4.0)	+	3	(3.7)		525	(2.9)	6	(4.4)		-10	(4.0)	–
			8	Iran	425	(3.7)	-12	(6.4)		-5	(5.8)		438	(4.7)	-2	(6.8)		-15	(6.6)	–	426	(3.6)	10	(6.2)		-12	(5.5)	–
	2			Slowakei	519	(3.7)	-16	(5.9)	–	-1	(5.4)		523	(3.6)	11	(6.3)		-2	(5.3)		519	(3.7)	-13	(7.5)		6	(5.7)	
3	3	3	8	Niederlande	518	(3.2)	-21	(4.2)	–	0	(4.6)		509	(2.5)	6	(4.0)		-7	(3.8)		527	(2.9)	3	(4.4)		6	(4.5)	
2 3	2 3	2 3		USA	541	(2.4)	-3	(3.8)		-5	(3.5)		525	(2.9)	-10	(4.2)	–	-2	(4.0)		528	(3.1)	-9	(4.5)	–	-11	(4.5)	–
2 3	3	3	8	Hongkong	541	(4.0)	1	(5.7)		18	(5.4)	+	553	(3.9)	-9	(5.6)		24	(5.2)	+	539	(4.9)	-29	(6.5)	–	-10	(6.7)	
		2		Ungarn	527	(3.3)	-26	(4.7)	–	-6	(4.7)		514	(3.1)	-15	(4.8)	–	-10	(4.2)	–	534	(3.8)	17	(5.7)	+	3	(5.0)	
				Deutschland	515	(2.5)	-16	(3.3)	–	-6	(3.4)		515	(2.6)	-12	(4.1)	–	-3	(4.0)		512	(3.2)	-12	(4.3)	–	3	(5.1)	
2		2	8	Italien	510	(2.9)	-45	(4.6)	–	-4	(4.4)		510	(2.6)	-10	(4.5)	–	8	(4.3)		508	(3.1)	-19	(5.1)	–	1	(4.8)	
				Benchmark-Teilnehmer																								
2 3 6	2		8	Dubai, VAE	555	(1.3)	99	(3.1)	+	18	(2.3)	+	570	(2.0)	114	(4.0)	+	14	(2.9)	+	569	(2.4)	108	(4.5)	+	27	(3.3)	+
2	3	2 3	8	Québec, Kanada	508	(2.5)	-16	(3.9)	–	-22	(3.5)	–	502	(1.9)	-7	(3.6)		-12	(3.4)	–	511	(2.6)	-11	(4.0)	–	-8	(4.1)	
2		2	8	Ontario, Kanada	533	(2.7)	-6	(4.8)		-2	(4.0)		520	(2.3)	-15	(4.0)	–	8	(3.7)	+	522	(3.3)	-11	(5.3)	–	4	(4.7)	



Keine statistisch signifikanten Veränderungen zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Veränderungen zwischen 2019 und 2023 ($p \leq .05$)
\+ = Mittelwert in 2023 statistisch signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 statistisch signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

In Deutschland zeigt sich seit 2007 für die Inhaltsbereiche *Biologie* und *Physik/Chemie* eine vergleichbare Entwicklung mit einem Rückgang von 16 bzw. 12 Punkten. Dagegen lässt sich für *Geografie* erkennen, dass der deutliche Rückgang zwischen 2007 und 2019 in TIMSS 2023 aufgehalten werden konnte. Die leichten, wenn auch nicht signifikanten Zuwächse in TIMSS 2023 sorgen dafür, dass sich *Geografie* – anders als noch in TIMSS 2019 – in den mittleren Leistungen nicht mehr signifikant von den anderen beiden Inhaltsbereichen unterscheidet.

4.5.4 Kognitive Anforderungsbereiche

Neben den Inhaltsbereichen wird in der TIMSS-Rahmenkonzeption auch zwischen den drei kognitiven Anforderungsbereichen *Reproduzieren* von Wissen (*Knowing*), *Anwenden* von Wissen (*Applying*) und *Problemlösen* (*Reasoning*) differenziert. Diese Trennung erlaubt es, unterschiedliche kognitive Facetten naturwissenschaftlicher Kompetenzen zu analysieren. Dieser Abschnitt ist analog zu dem vorherigen Abschnitt zu den drei Inhaltsbereichen gegliedert. Es werden zuerst die Ergebnisse im internationalen Vergleich und dann im Vergleich zu den vorherigen Studienzyklen 2007 und 2019 berichtet.

Kompetenzen in den kognitiven Anforderungsbereichen im internationalen Vergleich:

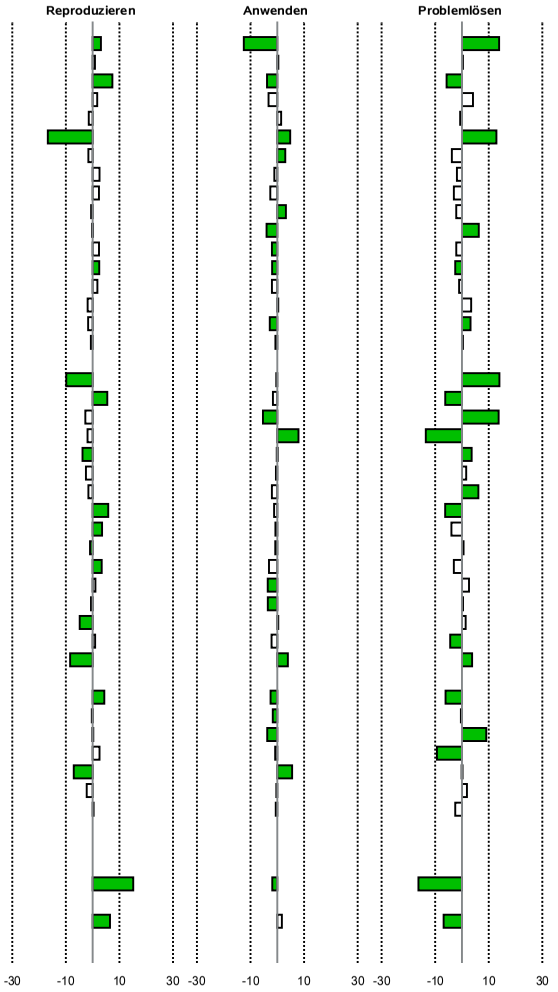
In Abbildung 4.11 sind die Testleistungen (Mittelwerte und Standardabweichungen mit ihren Standardfehlern) aufgeschlüsselt nach den kognitiven Anforderungsbereichen dargestellt. Aufgeführt sind auch hier nur OECD- und EU-Staaten sowie Staaten, die eine signifikant höhere oder gleich hohe Kompetenz auf der Gesamtskala wie Deutschland aufweisen. Die Anordnung entspricht der Rangfolge auf der Gesamtskala Naturwissenschaften. Für die Schüler:innen in Deutschland zeigt sich beim *Anwenden* von Wissen eine kleine relative Schwäche (512 Punkte), die auch statistisch signifikant vom deutschen Gesamtmittelwert in den Naturwissenschaften (515 Punkte) abweicht. Die Differenz ist allerdings so klein, dass sie im Folgenden nicht weiter interpretiert wird. Vielmehr weisen die Befunde für Deutschland insgesamt keine spezifischen Stärken und Schwächen in den drei kognitiven Anforderungsbereichen auf.

Ein ähnliches Bild wie für Deutschland zeigt sich auch für die Vergleichsgruppe der EU- und OECD-Staaten, in denen keine relativen Stärken und Schwächen in den drei Anforderungsbereichen beobachtbar sind. Die Staaten, die ähnliche Gesamtleistungen wie Deutschland erreichen, zeigen auch eher geringe bis moderate Unterschiede zwischen den drei Anforderungsbereichen, Abweichungen von über 10 Punkten treten hier kaum auf. Auch in den meisten anderen Teilnehmerstaaten finden sich in den drei Anforderungsbereichen allenfalls kleinere positive oder negative relative Verschiebungen von 1 bis 17 Punkten im Vergleich zu der jeweiligen Gesamtskala.

Wie schon in TIMSS 2019 stellt Japan eine Ausnahme dar, hier zeigt sich ein Unterschied von 30 Punkten zwischen dem stärksten Anforderungsbereich (*Problemlösen*) und dem schwächsten Anforderungsbereich (*Reproduzieren*). Im Jahr 2019 betrug diese Differenz sogar 45 Punkte. Auch bei Schweden fällt die große Differenz zwischen *Reproduktion* (523 Punkte) und *Problemlösen* (546 Punkte) auf. Diese relativen Stärken und Schwächen hängen vermutlich mit unterschiedlichen Unterrichtstraditionen in den verschiedenen Teilnehmerstaaten zusammen. Zusammenfassend zeigt sich aber, dass in den meisten Staaten eine gleichmäßige Förderung der Kompetenzen im Unterricht gelingt, die notwendig sind, um die Aufgaben in den verschiedenen Anforderungsbereichen zu bearbeiten.

Abbildung 4.11: Leistungen der Schüler:innen in den kognitiven Anforderungsbereichen *Reproduzieren*, *Anwenden* und *Problemlösen* im internationalen Vergleich

	Gesamtskala				Reproduzieren^						Anwenden^						Problemlösen^					
Teilnehmer	M_{GS}	*(SE)*	*SD*	*(SE)*	M_R	*(SE)*		*SD*	*(SE)*	M_R - M_{GS}	*(SE)*	M_A	*(SE)*		*SD*	*(SE)*	M_A - M_{GS}	*(SE)*	M_P	*(SE)*		*SD*	*(SE)*	M_P - M_{GS}	*(SE)*
2 *Singapur*	607	(2.3)	84	(1.3)	610	(2.4)	+	89	(1.7)	3	(1.4)	595	(2.2)	+	84	(1.5)	-13	(1.4)	621	(2.2)	+	90	(1.8)	14	(1.5)
Republik Korea (Südkorea)	583	(2.5)	74	(2.2)	584	(2.9)	+	78	(2.2)	1	(1.6)	584	(2.5)	+	79	(2.2)	0	(1.1)	583	(2.4)	+	75	(2.1)	0	(1.2)
Taiwan	573	(1.4)	71	(1.4)	580	(1.7)	+	76	(1.1)	7	(1.0)	569	(2.2)	+	73	(1.2)	-4	(1.4)	567	(1.6)	+	80	(1.4)	-6	(1.2)
123 *Türkei (5. Jgst.)*	570	(3.4)	86	(1.7)	572	(4.1)	+	98	(2.2)	2	(2.5)	567	(4.0)	+	83	(2.0)	-3	(2.5)	574	(3.9)	+	92	(2.6)	4	(2.4)
123 *England*	556	(2.6)	80	(2.0)	555	(3.1)	+	83	(2.0)	-1	(2.1)	558	(3.0)	+	81	(2.1)	1	(1.9)	556	(2.8)	+	87	(2.3)	-1	(2.0)
Japan	555	(2.4)	68	(1.4)	538	(3.3)	+	70	(1.4)	-17	(2.2)	560	(2.9)	+	73	(1.1)	5	(1.2)	568	(2.5)	+	73	(1.5)	13	(1.2)
2 *Polen*	550	(2.2)	71	(1.1)	548	(1.9)	+	72	(1.6)	-2	(1.2)	553	(2.3)	+	76	(1.8)	3	(1.2)	546	(2.4)	+	78	(1.3)	-4	(2.1)
Australien	550	(1.8)	81	(0.9)	552	(2.3)	+	86	(1.5)	3	(2.4)	548	(2.0)	+	79	(1.4)	-1	(1.4)	548	(3.2)	+	88	(1.3)	-2	(2.3)
3 *Hongkong*	545	(3.8)	88	(2.7)	548	(4.9)	+	99	(3.1)	2	(2.5)	543	(3.8)	+	84	(2.6)	-3	(1.6)	542	(4.6)	+	96	(3.4)	-3	(2.4)
Finnland	542	(2.8)	79	(1.8)	541	(2.7)	+	81	(1.6)	-1	(1.2)	545	(2.8)	+	81	(1.8)	3	(0.7)	540	(2.9)	+	86	(2.3)	-2	(1.5)
2 *Litauen*	537	(2.4)	74	(1.6)	537	(2.2)	+	78	(2.0)	0	(1.2)	533	(2.2)	+	78	(1.8)	-4	(1.4)	543	(2.0)	+	70	(2.0)	6	(2.0)
Macau	536	(1.3)	81	(0.8)	538	(1.6)	+	89	(1.1)	2	(1.4)	534	(1.5)	+	79	(1.2)	-2	(0.6)	534	(1.3)	+	87	(1.1)	-2	(1.2)
23 *USA*	532	(2.4)	93	(0.9)	535	(2.9)	+	95	(1.1)	2	(1.2)	530	(2.6)	+	95	(1.1)	-2	(0.8)	530	(2.5)	+	98	(1.2)	-3	(0.6)
Irland	532	(3.2)	80	(1.6)	534	(3.1)	+	85	(2.0)	2	(1.6)	530	(3.7)	+	82	(1.7)	-2	(2.1)	531	(3.3)	+	86	(1.9)	-1	(1.1)
12 *Norwegen (5. Jgst.)*	530	(2.5)	75	(1.1)	529	(2.5)	+	75	(1.6)	-2	(1.0)	530	(2.3)	+	80	(1.1)	0	(1.2)	534	(3.1)	+	82	(1.4)	3	(2.2)
2 *Tschechien*	526	(2.2)	71	(1.2)	524	(2.5)	+	73	(1.0)	-2	(1.2)	523	(2.2)	+	73	(0.9)	-3	(1.2)	529	(2.8)	+	78	(1.5)	3	(1.5)
VG OECD	**526**	**(0.5)**	**77**	**(0.3)**	**525**	**(0.5)**	+	**81**	**(0.3)**	**-1**	**(0.3)**	**525**	**(0.5)**	+	**80**	**(0.3)**	**-1**	**(0.3)**	**526**	**(0.6)**	+	**84**	**(0.4)**	**0**	**(0.4)**
2 *Schweden*	533	(3.1)	82	(2.0)	523	(3.6)		86	(2.0)	-10	(1.6)	532	(3.1)	+	84	(2.3)	0	(1.2)	546	(3.4)	+	85	(2.4)	14	(1.3)
Bulgarien	530	(4.8)	104	(4.1)	535	(4.9)	+	108	(4.5)	5	(1.9)	528	(5.0)	+	111	(4.4)	-2	(1.6)	523	(5.2)		103	(4.8)	-6	(2.1)
23 *Rumänien*	526	(4.8)	81	(2.4)	523	(4.7)		81	(2.5)	-3	(1.9)	521	(4.9)		85	(2.2)	-5	(1.7)	540	(4.9)	+	86	(2.4)	14	(1.3)
2 *Slowenien*	526	(2.3)	74	(1.4)	524	(2.2)	+	80	(2.1)	-2	(2.0)	534	(2.6)	+	77	(1.4)	8	(1.6)	512	(2.6)		83	(1.9)	-13	(2.2)
Lettland	526	(2.9)	77	(1.9)	522	(2.9)		79	(1.6)	-4	(1.9)	525	(3.3)	+	80	(1.9)	0	(1.2)	529	(3.5)	+	90	(2.1)	4	(1.8)
2 *Ungarn*	524	(3.1)	85	(2.5)	522	(3.7)		87	(2.3)	-3	(1.5)	524	(3.5)	+	89	(2.7)	-1	(1.3)	526	(3.5)	+	89	(3.1)	2	(1.6)
23 *Dänemark*	522	(2.5)	75	(1.2)	521	(2.7)		77	(1.1)	-2	(2.1)	520	(3.3)	+	79	(1.3)	-2	(1.6)	528	(2.7)	+	79	(1.1)	6	(1.7)
23 *Kanada*	521	(1.7)	76	(1.1)	527	(2.0)	+	78	(0.8)	6	(1.4)	520	(2.2)	+	78	(0.6)	-1	(1.1)	515	(2.6)		82	(1.1)	-6	(1.6)
Slowakei	521	(3.3)	87	(2.5)	524	(3.4)	+	90	(2.8)	4	(1.7)	520	(3.5)		92	(2.9)	-1	(1.2)	517	(4.5)		91	(2.8)	-4	(2.9)
VG EU	**518**	**(0.6)**	**78**	**(0.4)**	**517**	**(0.7)**		**82**	**(0.4)**	**-1**	**(0.3)**	**517**	**(0.7)**	+	**81**	**(0.4)**	**-1**	**(0.3)**	**519**	**(0.7)**		**83**	**(0.5)**	**1**	**(0.4)**
123 *Neuseeland*	517	(2.8)	87	(1.7)	520	(2.2)		87	(1.6)	3	(1.5)	514	(2.7)		90	(2.1)	-3	(1.7)	514	(3.0)		98	(1.5)	-3	(2.2)
3 *Niederlande*	517	(2.9)	67	(1.5)	518	(3.1)		72	(1.9)	1	(2.2)	513	(2.5)		69	(1.4)	-4	(1.8)	520	(3.0)		72	(1.8)	3	(3.1)
Deutschland	515	(2.6)	84	(1.7)	515	(2.2)		87	(1.6)	-1	(1.2)	512	(2.4)		84	(1.8)	-4	(1.3)	516	(2.6)		89	(1.9)	0	(2.5)
Portugal	511	(2.3)	73	(1.1)	506	(2.6)	–	73	(1.1)	-5	(1.4)	511	(3.5)		75	(1.4)	0	(2.3)	512	(2.8)		87	(2.2)	1	(1.6)
2 *Italien*	511	(2.5)	70	(1.1)	511	(2.9)		73	(2.0)	1	(1.2)	508	(2.8)		73	(1.5)	-2	(1.9)	506	(2.7)	–	76	(2.0)	-4	(1.4)
2 *Serbien*	510	(3.2)	74	(1.6)	501	(3.0)	–	74	(1.5)	-8	(1.7)	514	(3.5)		79	(2.1)	4	(1.8)	514	(3.8)		81	(2.4)	4	(1.5)
2 *Spanien*	504	(1.9)	72	(0.9)	508	(2.0)	–	77	(1.0)	4	(0.5)	502	(1.8)	–	73	(1.0)	-3	(1.0)	498	(2.2)	–	77	(1.4)	-6	(1.5)
Internationaler Mittelwert	**494**	**(0.4)**	**85**	**(0.2)**	**497**	**(0.4)**	–	**88**	**(0.3)**	**0**	**(0.2)**	**496**	**(0.4)**	–	**86**	**(0.3)**	**-2**	**(0.2)**	**497**	**(0.4)**	–	**89**	**(0.3)**	**0**	**(0.2)**
23 *Belgien (Fläm. Gem.)*	488	(2.6)	73	(1.8)	488	(3.3)	–	78	(1.7)	0	(1.8)	485	(2.5)	–	75	(1.8)	-4	(1.3)	497	(3.4)	–	79	(2.5)	9	(1.8)
2 *Frankreich*	488	(2.9)	74	(1.6)	491	(2.8)	–	74	(1.6)	3	(2.0)	487	(3.5)	–	77	(1.7)	-1	(1.7)	479	(3.4)	–	82	(1.9)	-9	(1.7)
2 *Zypern*	487	(3.0)	88	(1.6)	480	(3.5)	–	102	(1.8)	-7	(2.0)	493	(3.0)	–	89	(1.6)	6	(0.8)	487	(2.9)	–	82	(2.2)	0	(1.9)
2 *Belgien (Frz. Gem.)*	481	(2.6)	77	(1.6)	478	(2.6)	–	80	(1.2)	-2	(1.6)	480	(2.9)	–	77	(2.1)	0	(1.7)	483	(2.3)	–	82	(1.7)	2	(2.1)
23 *Chile*	479	(2.5)	80	(1.4)	479	(3.1)	–	81	(1.6)	0	(1.7)	479	(2.3)	–	84	(1.4)	-1	(1.8)	477	(2.8)	–	88	(2.1)	-3	(2.4)
Benchmark-Teilnehmer																									
Dubai, VAE	562	(1.7)	95	(1.0)	578	(1.5)	+	105	(1.1)	15	(1.0)	561	(1.5)	+	96	(1.0)	-2	(0.9)	546	(1.7)	+	90	(1.0)	-16	(1.3)
2 *Ontario, Kanada*	525	(2.6)	76	(1.2)	532	(2.7)	+	77	(1.3)	6	(1.5)	527	(2.9)	+	77	(1.1)	2	(1.3)	518	(3.4)		82	(1.5)	-7	(2.1)



□ Keine statistisch signifikanten Unterschiede zur Gesamtskala ($p > .05$)
■ Statistisch signifikante Unterschiede zur Gesamtskala ($p \leq .05$)
\+ = Mittelwert statistisch signifikant höher als der Mittelwert von Deutschland ($p \leq .05$)
– = Mittelwert statistisch signifikant niedriger als der Mittelwert von Deutschland ($p \leq .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Unterschiede in den Kompetenzen in den kognitiven Anforderungsbereichen zwischen TIMSS 2007, 2019 und 2023:

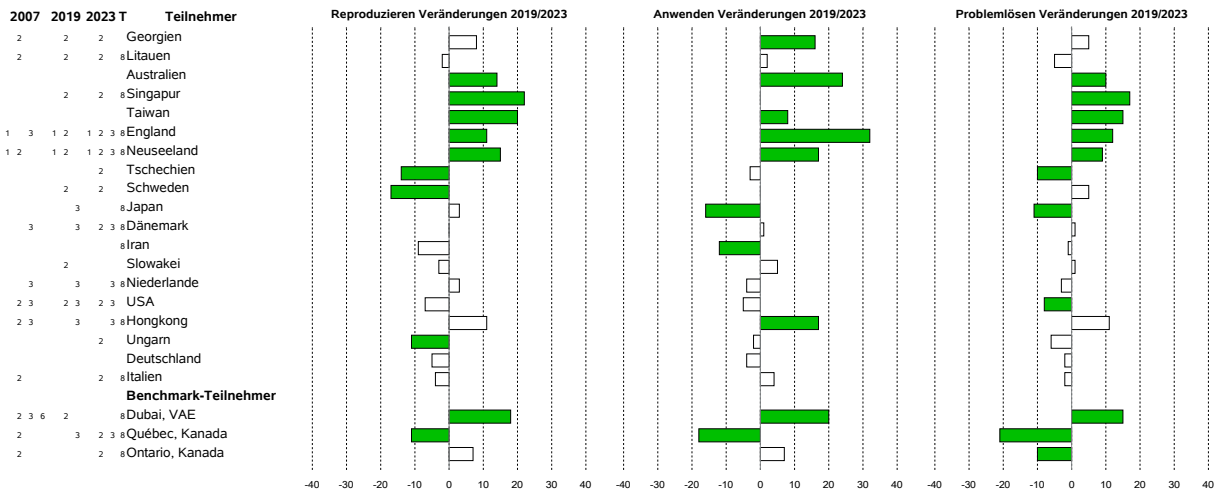
In Abbildung 4.12 sind die mittleren Testleistungen in den drei kognitiven Anforderungsbereichen aus TIMSS 2007, 2019 sowie 2023 dargestellt. Positive Differenzwerte in den Spalten *Veränderungen* im oberen Teil der Abbildung 4.12 bedeuten, dass höhere Kompetenzwerte im jeweils erstgenannten Studienzyklus erzielt wurden. Im unteren Teil von Abbildung 4.12 folgt eine grafische Darstellung mittels Differenzbalken. Während weiße Balken statistisch unbedeutende Veränderungen repräsentieren, markieren die grünen Balken statistisch signifikante Differenzen zwischen den Erhebungszyklen von TIMSS 2019 und 2023. Im Vergleich zu TIMSS 2007 wird in Abbildung 4.12 sichtbar, dass Schüler:innen in Deutschland sich in allen drei Anforderungsbereichen signifikant verschlechtert haben. Die Veränderungen zwischen 2019 und 2023 sind sehr klein, was sich auch schon beim Gesamtwert angedeutet hat. Vergleicht man die drei Anforderungsbereiche, so sind die Leistungen in Deutschland im *Problemlösen* seit 2007 weniger zurückgegangen (9 Punkte) als in den beiden anderen Anforderungsbereichen (jeweils 14 Punkte). Für die übrigen Staaten ergibt sich *cum grano salis*, dass Veränderungen über die Zeit – wenn sie auftreten – relativ homogen in allen drei kognitiven Anforderungsbereichen vorkommen.

4.5.5 Nachhaltigkeitskompetenzen im internationalen Vergleich

Die 44 Items aus der Gesamtskala Naturwissenschaften mit einem starken Bezug zu Umwelt und nachhaltigem Handeln sind (zusätzlich) zu einer eigenen Skala zusammengefasst. Im internationalen Vergleich (Abbildung 4.13; aufgrund von Einschränkungen in der Stichprobengenerierung wurde Südafrika nicht bei den Analysen berücksichtigt) zeigt sich, dass die Spannbreite der Mittelwerte etwas niedriger ist als bei der Gesamtskala Naturwissenschaften und nur 188 Punkte im Vergleich zu 234 Punkten beträgt. Es zeigen sich kleinere Unterschiede in der Rangfolge. 16 Staaten schneiden signifikant schwächer ab, zwölf besser und bei 25 Staaten lässt sich kein statistisch bedeutsamer Unterschied zur Gesamtskala Naturwissenschaften feststellen. Die meisten Abweichungen liegen bei 10 oder weniger Punkten. Größere Abweichungen zeigen sich für Singapur (30 Punkte), die Türkei (18 Punkte) und Lettland (22 Punkte), in diesen drei Staaten liegt der Mittelwert für die Gesamtskala Naturwissenschaften höher. In Deutschland ist der Mittelwert der Gesamtskala Naturwissenschaften um 2 Punkte höher als der der Skala Nachhaltigkeitskompetenzen, der Unterschied ist statistisch nicht bedeutsam. Die, von wenigen Ausnahmen abgesehen, eher geringen Unterschiede könnten ein Hinweis darauf sein, dass Schüler:innen im naturwissenschaftlichen Unterricht auch Lerngelegenheiten für umweltbezogene Themen erhalten. Nicht auszuschließen ist allerdings, dass es auch viele Gelegenheiten gibt, sich mit diesen Themen außerhalb des Unterrichts auseinanderzusetzen, z. B. im häuslichen Umfeld.

Abbildung 4.12: Leistungen in den kognitiven Anforderungsbereichen *Reproduzieren, Anwenden* und *Problemlösen* in TIMSS 2007, 2019 und 2023

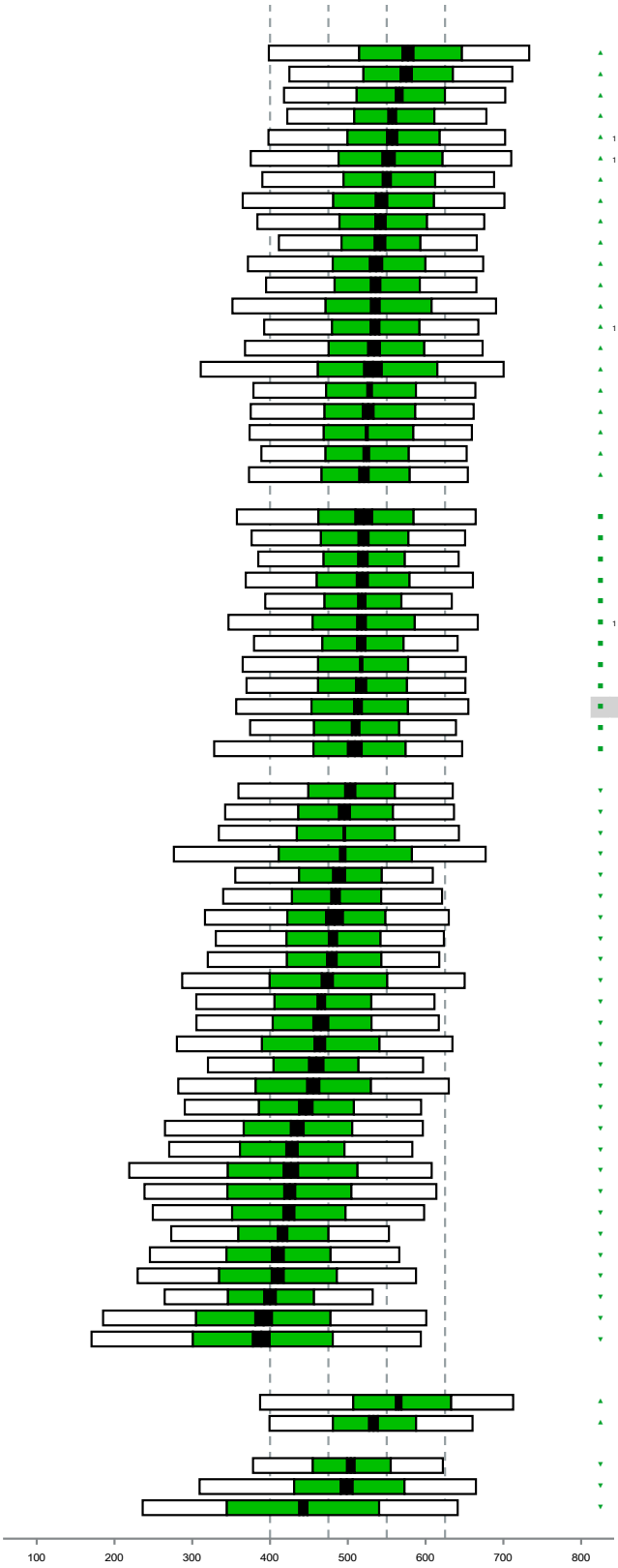
Fußnoten 2007	2019	2023	T	Teilnehmer	Reproduzieren M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		Anwenden M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		Problemlösen M_{23}	SE_{23}	M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)	
2	2	2		Georgien	460	(4.4)	31	(6.2)	+	8	(5.9)		461	(3.5)	46	(5.7)	+	16	(5.1)	+	470	(3.0)	91	(6.7)	+	5	(5.3)	
2	2	2	8	Litauen	537	(2.2)	26	(3.2)	+	-2	(3.8)		533	(2.2)	20	(4.0)	+	2	(3.2)		543	(2.0)	22	(3.5)	+	-5	(3.5)	
				Australien	552	(2.3)	20	(4.3)	+	14	(3.8)	+	548	(2.0)	26	(4.3)	+	24	(3.8)	+	548	(3.2)	20	(5.2)	+	10	(4.4)	+
	2	2	8	Singapur	610	(2.4)	11	(5.0)	+	22	(4.4)	+	595	(2.2)	8	(4.7)		0	(4.3)		621	(2.2)	45	(4.6)	+	17	(4.1)	+
				Taiwan	580	(1.7)	36	(3.3)	+	20	(2.5)	+	569	(2.2)	9	(3.1)	+	8	(3.0)	+	567	(1.6)	-7	(3.6)	–	15	(3.1)	+
1 3	1 2	1 2 3	8	England	555	(3.1)	8	(4.6)		11	(4.5)	+	558	(3.0)	21	(4.4)	+	32	(4.2)	+	556	(2.8)	16	(4.0)	+	12	(4.6)	+
1 2	1 2	1 2 3	8	Neuseeland	520	(2.2)	9	(4.0)	+	15	(3.5)	+	514	(2.7)	18	(3.8)	+	17	(3.7)	+	514	(3.0)	11	(5.0)	+	9	(4.0)	+
		2		Tschechien	524	(2.5)	3	(3.8)		-14	(3.8)	–	523	(2.2)	8	(4.0)	+	-3	(3.3)		529	(2.8)	22	(4.6)	+	-10	(4.3)	–
	2	2		Schweden	523	(3.6)	-5	(4.7)		-17	(5.0)	–	532	(3.1)	12	(4.3)	+	0	(4.4)		546	(3.4)	18	(5.5)	+	5	(4.7)	
	3		8	Japan	538	(3.3)	4	(4.3)		3	(4.2)		560	(2.9)	14	(4.3)	+	-16	(3.6)	–	568	(2.5)	-5	(3.4)		-11	(3.5)	–
3	3	2 3	8	Dänemark	521	(2.7)	4	(4.3)		0	(3.4)		520	(3.3)	7	(4.6)		1	(4.1)		528	(2.7)	4	(5.2)		1	(3.8)	
			8	Iran	435	(3.9)	4	(6.3)		-9	(6.0)		428	(3.8)	-15	(6.2)	–	-12	(5.7)	–	431	(3.6)	4	(5.8)		-1	(6.1)	
	2			Slowakei	524	(3.4)	-7	(5.9)		-3	(5.2)		520	(3.5)	-7	(6.0)		5	(5.5)		517	(4.5)	5	(7.0)		1	(6.2)	
3	3	3	8	Niederlande	518	(3.1)	-3	(4.0)		3	(4.2)		513	(2.5)	-12	(3.5)	–	-4	(4.0)		520	(3.0)	-6	(4.0)		-3	(4.4)	
2 3	2 3	2 3		USA	535	(2.9)	-11	(4.0)	–	-7	(4.0)		530	(2.6)	-4	(4.0)		-5	(4.0)		530	(2.5)	-5	(3.9)		-8	(3.7)	–
2 3	3	3	8	Hongkong	548	(4.9)	-5	(6.3)		11	(5.9)		543	(3.8)	-9	(5.1)		17	(4.9)	+	542	(4.6)	-21	(6.7)	–	11	(5.8)	
		2		Ungarn	522	(3.7)	-22	(5.1)	–	-11	(4.6)	–	524	(3.5)	-8	(5.2)		-2	(4.7)		526	(3.5)	-2	(5.5)		-6	(4.4)	
				Deutschland	515	(2.2)	-14	(3.3)	–	-5	(3.2)		512	(2.4)	-14	(3.5)	–	-4	(3.5)		516	(2.6)	-9	(3.7)	–	-2	(3.9)	
2		2	8	Italien	511	(2.9)	-24	(5.1)	–	-4	(4.2)		508	(2.8)	-33	(4.4)	–	4	(3.9)		506	(2.7)	-17	(4.5)	–	-2	(3.8)	
				Benchmark-Teilnehmer																								
2 3 6	2		8	Dubai, VAE	578	(1.5)	117	(3.0)	+	18	(2.6)	+	561	(1.5)	103	(4.0)	+	20	(2.7)	+	546	(1.7)	90	(3.4)	+	15	(2.7)	+
2	3	2 3	8	Québec, Kanada	512	(2.2)	-5	(3.7)		-11	(3.6)	–	502	(2.5)	-13	(3.9)	–	-18	(4.4)	–	504	(3.0)	-22	(4.8)	–	-21	(4.2)	–
2		2	8	Ontario, Kanada	532	(2.7)	-10	(4.6)	–	7	(4.1)		527	(2.9)	-2	(4.9)		7	(4.2)		518	(3.4)	-22	(4.8)	–	-10	(4.5)	–



Keine statistisch signifikanten Veränderungen zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Veränderungen zwischen 2019 und 2023 ($p \leq .05$)
\+ = Mittelwert in 2023 statistisch signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 statistisch signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Abbildung 4.13: Testleistung der Schüler:innen im internationalen Vergleich – Skala Nachhaltigkeitskompetenzen



	Teilnehmer[A]	*M*	*(SE)*	*SD*	*(SE)*
▲ 2	*Singapur*	577	(3.5)	102	(1.5)
▲	Republik Korea (Südkorea)	575	(3.6)	89	(2.5)
▲	Taiwan	566	(2.1)	86	(1.6)
▲ 2	*Polen*	557	(2.5)	78	(1.6)
▲ 1 2 3	*England*	557	(3.2)	91	(2.4)
▲ 1 2 3	*Türkei (5. Jgst.)*	553	(3.9)	102	(2.6)
▲	Australien	550	(2.6)	90	(1.5)
▲ 3	*Hongkong*	543	(3.7)	102	(3.2)
▲	Finnland	542	(3.3)	89	(2.0)
▲	Japan	541	(3.4)	77	(1.5)
▲	Irland	536	(3.9)	92	(2.8)
▲ 2 3	*Dänemark*	536	(3.0)	82	(1.5)
▲ 2 3	*USA*	535	(2.8)	102	(1.2)
▲ 1 2	*Norwegen (5. Jgst.)*	535	(2.8)	84	(1.4)
▲ 2	*Ungarn*	534	(3.8)	93	(2.9)
▲	Bulgarien	532	(5.6)	118	(5.1)
▲	Macau	528	(1.4)	87	(1.6)
▲ 2	*Schweden*	526	(3.5)	87	(2.5)
▲	**VG OECD**	**524**	**(0.6)**	**87**	**(0.4)**
▲ 2 3	*Kanada*	524	(1.8)	80	(1.2)
▲ 2	*Litauen*	521	(3.0)	85	(1.8)
■ 2 3	*Rumänien*	521	(5.3)	93	(2.8)
■ 2	*Slowenien*	520	(3.2)	83	(1.8)
■ 2	*Tschechien*	519	(3.0)	79	(1.1)
■	Portugal	519	(3.5)	89	(1.9)
■ 3	*Niederlande*	518	(2.3)	73	(1.7)
■ 1 2 3	*Neuseeland*	518	(2.7)	97	(2.2)
■ 2	*Italien*	517	(2.7)	79	(1.3)
■	**VG EU**	**517**	**(0.7)**	**87**	**(0.5)**
■ 2	*Serbien*	517	(3.2)	86	(2.0)
■	Deutschland	513	(2.4)	91	(1.6)
■ 2	*Spanien*	510	(2.4)	81	(1.4)
■	Slowakei	509	(4.4)	96	(3.5)
▼	Lettland	503	(3.3)	84	(2.3)
▼ 2	*Frankreich*	496	(3.5)	90	(1.7)
▼	**Internationaler Mittelwert**	**496**	**(0.5)**	**94**	**(0.3)**
▼	Vereinigte Arabische Emirate (VAE)	494	(1.7)	122	(0.8)
▼ 2 3	*Belgien (Fläm. Gem.)*	489	(3.8)	77	(2.1)
▼ 2 3	*Chile*	484	(2.7)	86	(1.6)
▼ 2 3	*Albanien*	483	(5.4)	96	(3.1)
▼ 2	*Belgien (Frz. Gem.)*	481	(2.6)	89	(1.7)
▼ 2	*Zypern*	480	(3.0)	90	(1.9)
▼	Bahrain	474	(3.6)	110	(2.9)
▼ 2	*Montenegro*	466	(2.4)	93	(2.5)
▼ 2	*Armenien*	465	(4.8)	95	(1.7)
▼	Katar	464	(3.4)	108	(1.6)
▼ 2	*Georgien*	459	(4.6)	84	(3.2)
▼ 2	*Kasachstan*	456	(3.9)	106	(1.8)
▼ 2	*Bosnien und Herzegowina*	446	(4.3)	92	(3.2)
▼	Nordmazedonien	435	(4.1)	101	(2.6)
▼ 2	*Saudi-Arabien*	428	(3.7)	96	(2.1)
▼	Iran	427	(4.6)	118	(2.1)
▼ 2 3	*Brasilien*	425	(3.5)	114	(1.7)
▼	Oman	424	(3.5)	106	(1.5)
▼ 2	*Usbekistan*	416	(3.0)	85	(2.1)
▼	Aserbaidschan	410	(3.8)	98	(2.7)
▼	Jordanien	410	(3.8)	109	(2.0)
▼ 2	*Kosovo*	400	(3.7)	82	(2.3)
▼	Marokko	392	(5.3)	126	(3.1)
▼ 4	*Kuwait*	389	(5.4)	129	(3.0)
	Benchmark-Teilnehmer				
▲	Dubai, VAE	565	(1.7)	98	(1.0)
▲ 2	*Ontario, Kanada*	533	(2.8)	79	(1.6)
▼ 2 3	*Québec, Kanada*	504	(2.4)	75	(1.4)
▼	Schardscha, VAE	499	(3.8)	107	(1.9)
▼	Abu Dhabi, VAE	443	(2.6)	127	(1.3)

▲ = Mittelwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Mittelwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu Deutschland nicht statistisch signifikant ($p > .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten
A = Aufgrund zu hoher Anteile nicht skalierbarer Leistungswerte werden die Ergebnisse von Südafrika nicht berichtet.

© TIMSS 2023

4.5.6 Einstellungen und Selbstkonzept

Der Sachunterricht (wie andere Fächer auch) zielt nicht nur auf kognitive Kompetenzen ab, sondern auch auf motivationale Merkmale von Schüler:innen wie Interesse, Einstellungen oder fachbezogene Selbstkonzepte. Entsprechende multikriteriale Ziele sind auch in den meisten Lehrplänen und dem Perspektivrahmen verankert. Diese motivationalen Merkmale gelten als zentrale Bedingungsfaktoren für erfolgreiche Bildungsverläufe (Kriegbaum et al., 2018). Schüler:innen, die sich als kompetent in einem Fach und die Inhalte als interessant und bedeutsam einschätzen, sind motivierter, sich mit einem Inhalt auseinanderzusetzen, und entwickeln wiederum höhere Kompetenzen (Wigfield & Eccles, 2000). Gleichzeitig beeinflussen die Kompetenzstände der Schüler:innen, welche fachbezogenen Fähigkeitsselbstbilder sie entwickeln und wie stark sie interessiert sowie motiviert sind. Es besteht also eine wechselseitige Beziehung zwischen motivationalen Merkmalen und kognitiven Kompetenzen. Im Rahmen der TIMSS-Konzeption werden die Einstellungen gegenüber den Naturwissenschaften und das naturwissenschaftsbezogene Selbstkonzept als zwei zentrale motivationale Merkmale aufgegriffen. Einstellungen sind affektive Bewertungen der Naturwissenschaften (Osborne et al., 2003), während unter dem naturwissenschaftsbezogenen Selbstkonzept die Selbsteinschätzung der eigenen Leistungsstärke in diesem Bereich verstanden wird (Möller & Trautwein, 2009). Diese Selbsteinschätzungen beruhen auf Kompetenzerfahrungen, die jemand in einem Schulfach oder einem Kompetenzbereich macht, dabei spielen soziale Vergleichsprozesse mit Mitschüler:innen eine wichtige Rolle.

In diesem Abschnitt werden die Ergebnisse zu beiden Merkmalen der Schüler:innen aus Deutschland dargestellt und mit den Ergebnissen aus den vorherigen Studienzyklen verglichen. Da bei Einstellungs- und Selbstkonzeptskalen kulturspezifische Referenzgruppeneffekte zum Tragen kommen (Seaton et al., 2009; Van de gaer et al., 2012), sind die verwendeten Messmodelle in der Regel nicht universell gültig. Mittelwertunterschiede zwischen Staaten lassen sich oftmals kaum valide interpretieren, weswegen hier auf den internationalen Vergleich verzichtet wird.

Die Herausforderung bei der Erhebung von naturwissenschaftsbezogenen Einstellungen und Selbstkonzepten von Grundschüler:innen in Deutschland ist, dass die Kinder zwar den Sachunterricht kennen, aber möglicherweise nicht zwischen den verschiedenen Perspektiven des Sachunterrichts differenzieren. Sachunterrichtslehrkräfte benennen in der Regel die einzelnen Perspektiven nicht gezielt oder grenzen diese voneinander ab, da das Selbstverständnis des Sachunterrichts ein perspektivenintegrierender Ansatz ist. Fragt man lediglich nach den Naturwissenschaften, ist wiederum unklar, ob alle Schüler:innen eine klare Vorstellung von den dazugehörigen Inhalten und Vorgehensweisen haben. Aus diesem Grund wurden die Schüler:innen in Deutschland vor den Fragen zu ihren Einstellungen und ihrem Selbstkonzept auf die *naturwissenschaftlichen* Anteile des Sachunterrichts verwiesen, indem ihnen konkrete naturwissenschaftliche Themen, z. B. Tiere, Pflanzen, Wetter, elektrischer Strom, Licht und Schatten, genannt wurden. Dieses Vorgehen hat sich in anderen Studien bewährt (Kleickmann et al., 2020). Zudem ist anzunehmen, dass die Schüler:innen durch den naturwissenschaftlichen Kompetenztest, den sie vor dieser Befragung bearbeiten, ein gewisses Priming auf die Naturwissenschaften erfahren haben. Gleichwohl sind die Ergebnisse mit einer gewissen Unsicherheit behaftet, welche Vorstellung die

Schüler:innen für ihre Einschätzungen zugrunde gelegt haben. Aus diesem Grund wird von Einstellungen gegenüber dem Sachunterricht gesprochen.

Die Skala zur Erhebung der Einstellungen zum Sachunterricht umfasst drei Aussagen, bei denen die Schüler:innen angeben mussten, ob sie ihnen *völlig zustimmen*, *eher zustimmen*, *eher nicht zustimmen* oder *überhaupt nicht zustimmen.* Folgende Aussagen mussten dabei eingeschätzt werden:

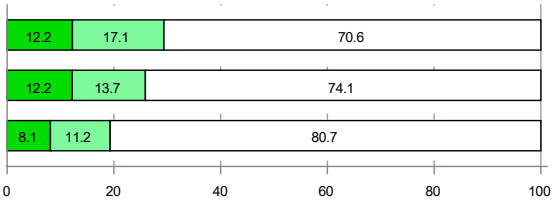
1. *Ich lerne gern im Sachunterricht.*
2. *Sachunterricht ist langweilig.*
3. *Ich mag Sachunterricht.*

Die Antworten der Schüler:innen wurden bei Aussage 1 und 3 mit 1 (*stimme überhaupt nicht zu*) bis 4 (*stimme völlig zu*) Punkten kodiert, bei Aussage 2 in umgekehrter Reihenfolge. Der Mittelwert über die drei Aussagen drückt aus, wie stark positiv die Einstellungen des Kindes zum Sachunterricht ausgeprägt sind. Die interne Konsistenz der Skala ist hoch (Cronbachs Alpha = .85). Im nächsten Schritt wurden die Schüler:innen anhand ihrer Mittelwerte in drei Gruppen mit *niedrig, mittel* und *hoch* ausgeprägten positiven Einstellungen zum Sachunterricht eingeteilt. Schüler:innen mit einem Mittelwert (M) von $1 \leq M < 2$ bildeten die Gruppe *niedrig,* mit einem Mittelwert von $2 \leq M < 3$ die Gruppe *mittel* und diejenigen mit einem Mittelwert von $3 \leq M \leq 4$ die Gruppe *hoch*.

In Abbildung 4.14 sind links die Mittelwerte und Standardfehler der Einstellungen der Grundschüler:innen aus Deutschland im aktuellen Studienzyklus sowie in den beiden ausgewählten vorherigen Studienzyklen dargestellt. Rechts sind in den horizontalen Stapelbalken die prozentualen Anteile (und deren Standardfehler) der Schüler:innen mit *niedrig*, *mittel* und *hoch* positiven Einstellungen gegenüber dem Sachunterricht dargestellt.

Abbildung 4.14: Mittlere positive Einstellung von Schüler:innen zum Sachunterricht sowie prozentuale Verteilungen nach *hoch*, *mittel* und *niedrig* positiver Einstellung – TIMSS 2007, 2019 und 2023 im Vergleich

	Mittlerer Skalenwert		Positive Einstellung zum Sachunterricht					
			niedrig		mittel		hoch	
	M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
2023	3.16	(0.02) ■	12.2	(0.5) ■	17.1	(0.6) ■	70.6	(0.8) ■
2019	3.22	(0.03) ■	12.2	(0.9) ■	13.7	(0.7) ▼	74.1	(1.3) ▲
2007	3.38	(0.02) ▲	8.1	(0.5) ▼	11.2	(0.6) ▼	80.7	(0.8) ▲



% der Schüler:innen mit niedriger positiver Einstellung zum Sachunterricht
% der Schüler:innen mit mittlerer positiver Einstellung zum Sachunterricht
% der Schüler:innen mit hoher positiver Einstellung zum Sachunterricht
▲ = Wert im Vergleich zu 2023 statistisch signifikant höher ($p \leq .05$)
▼ = Wert im Vergleich zu 2023 statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu 2023 nicht statistisch signifikant ($p > .05$)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Der Mittelwert der selbstberichteten Einstellungen zum Sachunterricht liegt bei 3.16 Punkten auf der vierstufigen Skala und ist damit sehr hoch. Mehr als zwei Drittel der Schüler:innen haben eine sehr positive Einstellung zum Sachunterricht. Im Vergleich zu TIMSS 2007 und 2019 ist der Mittelwert allerdings mit 0.22 bzw. 0.06 Punkten etwas niedriger. Die Veränderung zu 2007 ist dabei statistisch signifikant. Es handelt sich aber allenfalls um kleine Effekte (d = 0.20 bzw. 0.11).

Entsprechende Verschiebungen zeigen sich auch in den drei Gruppen zuungunsten der sehr positiven Gruppe.

Zur Erfassung des sachunterrichtsbezogenen Selbstkonzepts mussten die Schüler:innen die folgenden fünf Aussagen[4] mit den zuvor beschriebenen Abstufungen von *stimme überhaupt nicht zu* bis s*timme völlig zu* einschätzen:

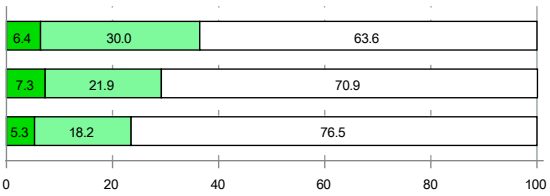
1. *Normalerweise bin ich gut im Sachunterricht.*
2. *Sachunterricht fällt mir schwerer als vielen meiner Mitschülerinnen und Mitschüler.*
3. *Ich bin einfach nicht gut im Sachunterricht.*
4. *Sachunterricht fällt mir leicht.*
5. *Ich finde Sachunterricht schwerer als alle anderen Fächer.*

Die Antworten der Schüler:innen wurden bei Aussage 1 und 4 mit 1 (*stimme überhaupt nicht zu*) bis 4 (*stimme völlig zu*) Punkten kodiert, bei den Aussagen 2 und 3 in umgekehrter Reihenfolge. Die interne Konsistenz dieser Skala ist mit Cronbachs Alpha = .78 etwas niedriger im Vergleich zu der der Einstellungen.

In Abbildung 4.15 sind analog zu Abbildung 4.14 die Mittelwerte und Standardfehler sowie die Verteilung auf die drei Gruppen von Schüler:innen mit einem *niedrigen*, *mittleren* und *hohen* sachunterrichtsbezogenem Selbstkonzept in den drei Studienzyklen dargestellt. Ähnlich wie die positiven Einstellungen ist auch das sachunterrichtsbezogene Selbstkonzept mit einem Mittelwert von 3.13 Punkten sehr hoch ausgeprägt. Knapp zwei Drittel der Schüler:innen schätzen ihre eigene Leistungsfähigkeit im Sachunterricht hoch ein. Gleichzeitig zeigt sich wie auch bei den Einstellungen eine Abnahme der Schüler:innen mit *hohem* sachunterrichtsbezogenen Selbstkonzept zugunsten der Gruppen *mittel* und *niedrig*. Die Unterschiede zwischen TIMSS 2023 und TIMSS 2019 betragen 0.09 Punkte bzw. 0.16 Punkte zu TIMSS 2007. Auch hierbei handelt es sich zwar um signifikante, aber lediglich sehr kleine Änderungen mit geringen Effekten (d = 0.12).

Abbildung 4.15: Mittleres sachunterrichtbezogenes Selbstkonzept von Schüler:innen sowie prozentuale Verteilungen nach *hohem*, *mittlerem* und *niedrigem* Selbstkonzept – TIMSS 2007, 2019, und 2023 im Vergleich

	Mittlerer Skalenwert		Sachunterrichtsbezogenes Selbstkonzept niedrig		mittel		hoch	
	M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
2023[A]	3.13	(0.01)	6.4	(0.3)	30.0	(0.7)	63.6	(0.8)
2019	3.22	(0.02)	7.3	(0.5)	21.9	(1.0)	70.9	(1.1)
2007	3.29	(0.01)	5.3	(0.4)	18.2	(0.6)	76.5	(0.8)



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen mit niedrigem sachunterrichtsbezogenen Selbstkonzept
% der Schüler:innen mit mittlerem sachunterrichtsbezogenen Selbstkonzept
% der Schüler:innen mit hohem sachunterrichtsbezogenen Selbstkonzept

A = Aufgrund veränderter Operationalisierung sind die Ergebnisse nur eingeschränkt mit vorherigen Zyklen vergleichbar.

IEA: Trends in International Mathematics and Science Study

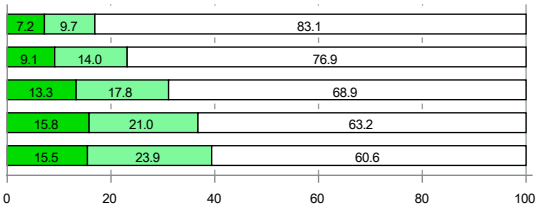
© TIMSS 2023

4 In den Erhebungsrunden vor TIMSS 2023 wurde statt der Aussagen 4 und 5 die Aussage *Ich lerne schnell im Sachunterricht* genutzt. Durch Veränderungen im internationalen Erhebungsinstrument wurde diese Aussage nicht mehr in TIMSS 2023 erhoben. Stattdessen wurden die Aussagen 4 und 5 ergänzt. Aufgrund der veränderten Zusammenstellung der Aussagen sind die Daten von TIMSS 2023 im Trend lediglich eingeschränkt vergleichbar.

Wie eingangs beschrieben besteht eine wechselseitige Beziehung zwischen Leistung und motivationalen Merkmalen. So sind verschiedene motivationale Merkmale wichtig für erfolgreiche Lernprozesse und umgekehrt wirkt sich eine günstige Leistungsentwicklung positiv auf motivationale Merkmale aus (Marsh et al., 2005). Entsprechend sind motivationale Merkmale wie Einstellungen, Interesse, Lernfreude sowie fachbezogene Selbstkonzepte und Selbstwirksamkeitserwartungen wichtige Ziele von Unterricht. Vor diesem Hintergrund wird analysiert, inwieweit die Einstellungen und das sachunterrichtsbezogene Selbstkonzept von Schüler:innen davon abhängig sind, welche Kompetenzstufe die Schüler:innen erreichen. Hierzu sind in den Abbildungen 4.16 und 4.17 die Mittelwerte und die beschriebene Einteilung in die drei Gruppen *niedrig, mittel* und *hoch* in Bezug auf die Einstellungen bzw. das fachbezogene Selbstkonzept für die Schüler:innen der Kompetenzstufen I (*rudimentär*) bis V (*fortgeschritten*) dargestellt.

Abbildung 4.16: Einstellungen zum Sachunterricht im Vergleich zur erreichten Kompetenzstufe

Kompetenzstufe	Mittlerer Skalenwert		Positive Einstellung zum Sachunterricht niedrig		mittel		hoch	
	M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
V	3.38	(0.04)	7.2	(1.0)	9.7	(1.6)	83.1	(1.6)
IV	3.26	(0.02)	9.1	(1.1)	14.0	(1.2)	76.9	(1.8)
III	3.13	(0.03)	13.3	(1.5)	17.8	(1.6)	68.9	(2.2)
II	3.03	(0.06)	15.8	(2.5)	21.0	(1.7)	63.2	(2.8)
I	2.93	(0.05)	15.5	(2.1)	23.9	(2.7)	60.6	(3.4)



% der Schüler:innen mit niedriger positiver Einstellung zum Sachunterricht
% der Schüler:innen mit mittlerer positiver Einstellung zum Sachunterricht
% der Schüler:innen mit hoher positiver Einstellung zum Sachunterricht

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

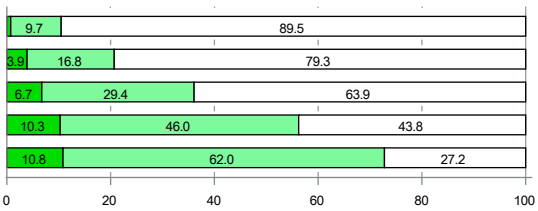
Aus Abbildung 4.16 lässt sich entnehmen, dass die positiven Einstellungen von leistungsstärkeren Schüler:innen im Vergleich zu leistungsschwächeren Schüler:innen auf den unteren Kompetenzstufen stärker ausgeprägt sind. So nimmt der Anteil der Schüler:innen der Gruppe *hoch* von der höchsten zur untersten Kompetenzstufe um ca. 20 Prozent ab. Gleichwohl zeigt sich, dass auch die leistungsschwächeren Schüler:innen immer noch relativ positive Einstellungen zum Fach Sachunterricht aufweisen.

Noch deutlicher ausgeprägt ist der Befund, wenn man das sachunterrichtsbezogene Selbstkonzept der Schüler:innen der verschiedenen Kompetenzstufen vergleicht (Abbildung 4.17). So nimmt hier der Mittelwert zwischen der Kompetenzstufe V (3.55) bis zur Kompetenzstufe I (2.64) um ca. eine Stufe auf der vierstufigen Skala ab. Gleichwohl liegt der Mittelwert auch auf der Kompetenzstufe I immer noch über dem theoretischen Mittelwert der Skala.

Zusammenfassend zeigt sich, dass die hier erhobenen motivationalen Merkmale der Schüler:innen hoch ausgeprägt sind und die Mehrheit, also ca. drei Viertel, günstige Voraussetzungen für das Lernen und die Weiterentwicklung von Interesse, Motivation und das Zutrauen in die eigenen Fähigkeiten im Kontext der Naturwissenschaften mitbringen. Gleichzeitig muss die sich abzeichnende, wenn auch nur geringe Abnahme der positiven Einstellungen und des sachunterrichtsbezogenen Selbstkonzepts in den bisherigen Studienzyklen aufmerksam beobachtet werden. So wäre es ungünstig, eine wachsende Gruppe von Schüler:innen zu haben, die dem Sachunterricht und damit auch den Naturwis-

Abbildung 4.17: Sachunterrichtsbezogenes Selbstkonzept im Vergleich zur erreichten Kompetenzstufe

Kompetenzstufe	Mittlerer Skalenwert		Sachunterrichtsbezogenes Selbstkonzept niedrig		mittel		hoch	
	M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
V	3.55	(0.03)	0.7	(0.4)	9.7	(1.6)	89.5	(1.5)
IV	3.35	(0.02)	3.9	(0.9)	16.8	(1.1)	79.3	(1.6)
III	3.11	(0.02)	6.7	(0.9)	29.4	(1.4)	63.9	(1.8)
II	2.85	(0.03)	10.3	(2.0)	46.0	(2.3)	43.8	(2.3)
I	2.64	(0.04)	10.8	(2.1)	62.0	(2.9)	27.2	(3.1)



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen mit niedrigem sachunterrichtsbezogenen Selbstkonzept
% der Schüler:innen mit mittlerem sachunterrichtsbezogenen Selbstkonzept
% der Schüler:innen mit hohem sachunterrichtsbezogenen Selbstkonzept

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

senschaften und dem eigenen Lernen gegenüber dort weniger positiv eingestellt ist. Dieser Trend betrifft erwartungskonform insbesondere die Gruppe der leistungsschwächeren Schüler:innen. Es ist entsprechend wichtig, insbesondere bei dieser Gruppe nicht nur die Wissensentwicklung, sondern auch die Entwicklung motivationaler Merkmale gezielt zu fördern. Dabei geht es nicht um die Entwicklung eines unrealistisch hohen sachunterrichtsbezogenen Selbstkonzepts, sondern darum, durch die Förderung von Wissen und Interesse, Motivation und Einstellungen die Entwicklung eines (realistischen) positiven sachunterrichtsbezogenen Selbstkonzepts zu unterstützen.

4.5.7 Umwelteinstellungen

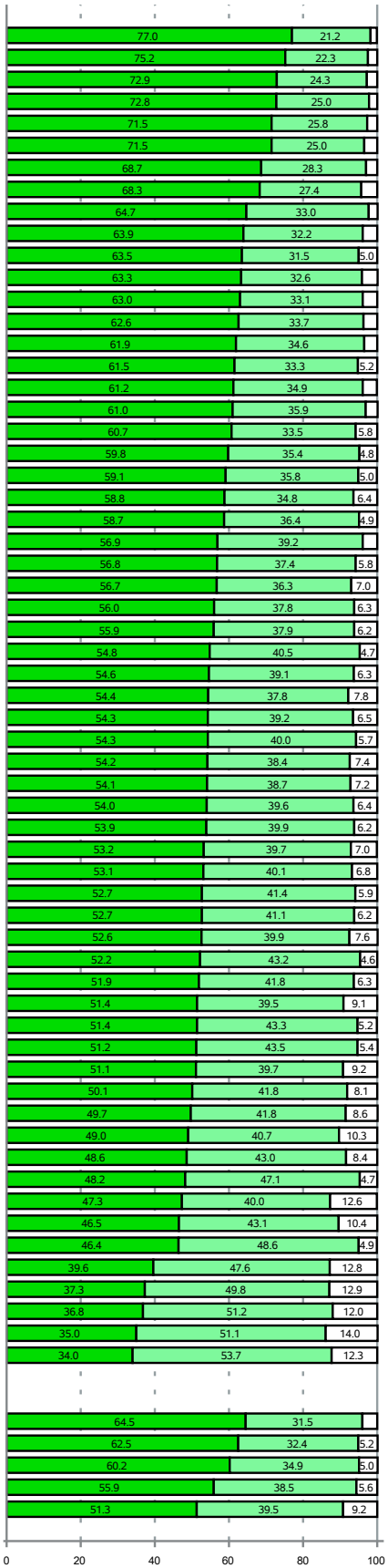
Neben den naturwissenschaftsspezifischen motivationalen Orientierungen wurden in TIMSS 2023 erstmalig auch Umwelteinstellungen erfasst. Die theoretische Grundlage hierfür liefert das zweidimensionale *Model of Ecological Values* (Maurer & Bogner, 2020; Wiseman & Bogner, 2003). Die zwei Dimensionen des Modells lassen sich als Präferenzen (1) für die Erhaltung und den Schutz der Natur und die Wertschätzung der Natur und (2) für die Naturnutzung (im Sinne von Naturausnutzung) beschreiben. Die beiden Dimensionen sind nicht unabhängig voneinander, sondern negativ miteinander korreliert (Milfont & Duckitt, 2004). Für TIMSS in der Grundschule wurde lediglich die Dimension Präferenzen für die Erhaltung der Natur und deren Wertschätzung eingesetzt. Hierzu mussten die Schüler:innen die folgenden fünf Aussagen mit den Abstufungen von *stimme überhaupt nicht zu* bis s*timme völlig zu* einschätzen:

1. *Es ist mir wichtig, dass Pflanzen und Tiere geschützt werden.*
2. *Es macht mich traurig, wenn die Natur zerstört wird.*
3. *Es macht mir Spaß herauszufinden, was für Pflanzen und Tiere in meiner Umgebung leben.*
4. *Ich bin gerne in der Natur (z. B. im Wald, im Park, am See).*
5. *Es sollte sehr wichtig sein, sich mit dem Klimawandel zu beschäftigen.*

Die Antworten der Schüler:innen wurden mit 1 (*stimme überhaupt nicht zu*) bis 4 (*stimme völlig zu*) Punkten kodiert. Die Ergebnisse zeigen, dass die Schüler:innen in Deutschland im Mittel eine hohe Zustimmung zu den Items

Abbildung 4.18: Umwelteinstellungen im internationalen Vergleich

Teilnehmer	hoch %	(SE)	mittel %	(SE)	niedrig %	(SE)
23 *Albanien*	77.0	(1.3)	21.2	(1.1)	1.7	(0.4)
123 *Türkei (5. Jgst.)*	75.2	(0.9)	22.3	(0.8)	2.5	(0.3)
Bulgarien	72.9	(1.3)	24.3	(1.1)	2.8	(0.5)
Portugal	72.8	(0.8)	25.0	(0.8)	2.2	(0.3)
23 *Rumänien*	71.5	(1.2)	25.8	(1.1)	2.7	(0.4)
Iran	71.5	(0.9)	25.0	(0.7)	3.5	(0.4)
2 *Zypern*	68.7	(0.9)	28.3	(0.8)	3.0	(0.2)
Nordmazedonien	68.3	(0.8)	27.4	(0.7)	4.3	(0.4)
2 *Georgien*	64.7	(1.2)	33.0	(1.2)	2.4	(0.3)
Irland	63.9	(1.0)	32.2	(0.9)	3.9	(0.4)
2 *Montenegro*	63.5	(0.7)	31.5	(0.7)	5.0	(0.3)
2 *Spanien*	63.3	(0.6)	32.6	(0.5)	4.2	(0.3)
2 *Serbien*	63.0	(1.2)	33.1	(1.1)	3.9	(0.4)
2 *Litauen*	62.6	(0.8)	33.7	(0.8)	3.7	(0.3)
2 *Kosovo*	61.9	(0.9)	34.6	(0.9)	3.5	(0.4)
23 *Chile*	61.5	(0.8)	33.3	(0.8)	5.2	(0.4)
2 *Italien*	61.2	(0.9)	34.9	(0.9)	3.8	(0.3)
2 *Bosnien und Herzegowina*	61.0	(1.4)	35.9	(1.3)	3.2	(0.4)
Marokko	60.7	(1.1)	33.5	(1.0)	5.8	(0.2)
23 *Kanada*	59.8	(0.5)	35.4	(0.5)	4.8	(0.1)
2 *Usbekistan*	59.1	(1.3)	35.8	(1.2)	5.0	(0.4)
Vereinigte Arabische Emirate (VAE)	58.8	(0.4)	34.8	(0.3)	6.4	(0.2)
2 *Ungarn*	58.7	(0.9)	36.4	(0.9)	4.9	(0.4)
2 *Slowenien*	56.9	(1.1)	39.2	(1.0)	3.9	(0.3)
VG EU	**56.8**	**(1.0)**	**37.4**	**(0.9)**	**5.8**	**(0.4)**
123 *England*	56.7	(1.3)	36.3	(1.1)	7.0	(0.6)
Internationaler Mittelwert	**56.0**	**(1.0)**	**37.8**	**(0.9)**	**6.3**	**(0.4)**
Deutschland	55.9	(0.8)	37.9	(0.7)	6.2	(0.3)
2 *Armenien*	54.8	(1.0)	40.5	(0.9)	4.7	(0.3)
VG OECD	**54.6**	**(1.0)**	**39.1**	**(0.9)**	**6.3**	**(0.4)**
Bahrain	54.4	(0.8)	37.8	(0.7)	7.8	(0.4)
23 *Belgien (Fläm. Gem.)*	54.3	(1.1)	39.2	(0.9)	6.5	(0.5)
Macau	54.3	(0.5)	40.0	(0.5)	5.7	(0.3)
2 *Belgien (Frz. Gem.)*	54.2	(1.0)	38.4	(0.9)	7.4	(0.4)
2 *Singapur*	54.1	(0.5)	38.7	(0.5)	7.2	(0.3)
Kuwait	54.0	(1.0)	39.6	(0.9)	6.4	(0.5)
2 *Frankreich*	53.9	(1.1)	39.9	(1.0)	6.2	(0.5)
23 *Brasilien*	53.2	(0.5)	39.7	(0.5)	7.0	(0.4)
123 *Neuseeland*	53.1	(0.8)	40.1	(0.8)	6.8	(0.5)
Aserbaidschan	52.7	(0.9)	41.4	(0.7)	5.9	(0.3)
Australien	52.7	(0.7)	41.1	(0.7)	6.2	(0.3)
Taiwan	52.6	(0.8)	39.9	(0.7)	7.6	(0.3)
Lettland	52.2	(1.3)	43.2	(1.2)	4.6	(0.5)
Slowakei	51.9	(1.2)	41.8	(1.1)	6.3	(0.5)
Katar	51.4	(0.6)	39.5	(0.5)	9.1	(0.3)
2 *Tschechien*	51.4	(0.8)	43.3	(0.7)	5.2	(0.3)
2 *Polen*	51.2	(1.0)	43.5	(0.9)	5.4	(0.4)
Oman	51.1	(0.7)	39.7	(0.6)	9.2	(0.3)
23 *USA*	50.1	(0.7)	41.8	(0.6)	8.1	(0.3)
3 *Hongkong*	49.7	(1.1)	41.8	(0.8)	8.6	(0.6)
Jordanien	49.0	(1.0)	40.7	(0.8)	10.3	(0.5)
Finnland	48.6	(1.0)	43.0	(0.8)	8.4	(0.5)
Japan	48.2	(1.2)	47.1	(1.1)	4.7	(0.4)
2 *Saudi-Arabien*	47.3	(0.9)	40.0	(0.7)	12.6	(0.5)
3 *Niederlande*	46.5	(1.2)	43.1	(1.2)	10.4	(0.7)
2 *Kasachstan*	46.4	(1.0)	48.6	(0.9)	4.9	(0.3)
2 *Schweden*	39.6	(0.9)	47.6	(0.9)	12.8	(0.6)
23 *Dänemark*	37.3	(1.0)	49.8	(0.9)	12.9	(0.8)
12 *Norwegen (5. Jgst.)*	36.8	(0.9)	51.2	(0.7)	12.0	(0.6)
1 *Südafrika (5. Jgst.)*	35.0	(0.8)	51.1	(0.6)	14.0	(0.4)
Republik Korea (Südkorea)	34.0	(0.9)	53.7	(0.9)	12.3	(0.6)
Benchmark-Teilnehmer						
Dubai, VAE	64.5	(0.6)	31.5	(0.6)	4.0	(0.2)
Schardscha, VAE	62.5	(0.7)	32.4	(0.6)	5.2	(0.3)
2 *Ontario, Kanada*	60.2	(0.7)	34.9	(0.7)	5.0	(0.3)
23 *Québec, Kanada*	55.9	(0.9)	38.5	(0.9)	5.6	(0.3)
Abu Dhabi, VAE	51.3	(0.7)	39.5	(0.4)	9.2	(0.3)



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen mit hoher nachhaltigkeitsbezogener Einstellung
% der Schüler:innen mit mittlerer nachhaltigkeitsbezogener Einstellung
% der Schüler:innen mit niedriger nachhaltigkeitsbezogener Einstellung

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

© TIMSS 2023

zeigen (M = 10.01). Die Skalenwerte wurden zudem in drei Kategorien unterteilt: Werte ab dem Wert 9.8 entsprechen einer sehr *hohen* Präferenz für den Schutz und die Wertschätzung, Werte zwischen 9.8 und 7.2 einer *mittleren* Präferenz und Werte unter 7.2 einer *geringeren* Präferenz. Abbildung 4.18 zeigt die Verteilung auf die Kategorien. Ungefähr die Hälfte der Schüler:innen in Deutschland zeigt eine *hohe* Präferenz zum Schutz der Natur und deren Wertschätzung, während nur 6 Prozent eine *gering* ausgeprägte Präferenz aufweisen. Die Verteilung auf die Kategorien ist sehr ähnlich zum internationalen Durchschnitt sowie zu den beiden Vergleichsgruppen EU und OECD.

Setzt man die Einstellungen mit den Testwerten der Nachhaltigkeitskompetenzen in Zusammenhang, ergibt sich folgendes Bild: Die Schüler:innen in Deutschland mit einer *hohen* Präferenz zum Schutz der Natur haben den höchsten Kompetenzwert (M = 516), während Schüler:innen mit einer *mittleren* oder *geringeren* Präferenz im Durchschnitt geringere Kompetenzwerte aufweisen (M = 490 bzw. 470). Dieses Muster zeigt sich auch in anderen Staaten. Mehr Wissen im Bereich der Naturwissenschaften scheint mit höherer Präferenz für den Schutz der Natur einherzugehen. Möglicherweise wird dieser Zusammenhang durch häusliche Aktivitäten und damit über den Bildungshintergrund vermittelt: Erziehungsberechtigte mit höherem Bildungshintergrund bieten mehr Aktivitäten in der Natur bzw. zum Thema Umweltschutz an, was sich in ausgeprägteren Präferenzen und höheren Kompetenzwerten widerspiegelt (Dominke & Steffensky, 2024; Junge et al., 2021).

4.6 Zusammenfassung

TIMSS 2023 stand wie schon die nationalen und internationalen Schulleistungsstudien der letzten Jahre (PISA 2022, IGLU 2021, IQB-Bildungstrends 2021 und 2022) im Schatten der COVID-19-Pandemie. Dabei zeigten alle Studien deutliche Kompetenzeinbußen in Deutsch und Mathematik gegenüber den Erhebungen vor der COVID-19-Pandemie. In den Naturwissenschaften fielen die Verluste etwas geringer aus (PISA 2022; Lewalter et al., 2023). TIMSS 2023 gibt erstmals Auskunft über potenzielle Pandemieeffekte auf naturwissenschaftliche Kompetenzen am Ende der 4. Jahrgangsstufe in der Grundschule, weshalb die Veränderungen zwischen TIMSS 2019 und TIMSS 2023 von besonderer Bedeutung sind. Gleichzeitig wurde mit Bezug auf das Jahr 2007, in dem sich Deutschland erstmalig an TIMSS-Grundschule beteiligt hat, die längerfristige Entwicklung der naturwissenschaftlichen Kompetenzen betrachtet. Im Folgenden werden die Ergebnisse zusammengefasst und vor allem aus einer nationalen Perspektive interpretiert.

Niveau und Streuung der naturwissenschaftlichen Kompetenz:

Gegenüber dem Studienzyklus 2019 (518 Punkte) ist der Mittelwert für Deutschland im Bereich der naturwissenschaftlichen Kompetenz um lediglich 3 auf 515 Punkte gesunken. Diese Differenz ist statistisch nicht signifikant, d. h., die im Jahr 2023 in Deutschland getesteten Schüler:innen erreichen dasselbe Niveau naturwissenschaftlicher Kompetenzen wie die im Jahr 2019 vor der COVID-19-Pandemie getesteten Schüler:innen. Dieser Befund ist bemerkenswert und gleichzeitig überraschend, bedenkt man, dass die in TIMSS 2023 getestete Kohorte in der 1. und 2. Jahrgangsstufe von den Schulschließungen mas-

siv betroffen war. Spezifisch für den Sachunterricht in Deutschland mag man an dieser Stelle einwenden, dass der Sachunterricht je nach Land der Bundesrepublik Deutschland erst in der 3. Klasse einsetzt, dort, wo er ab der 1. Klasse unterrichtet wird, die Zahl der Wochenstunden in den ersten beiden Schuljahren eher gering ist oder er teilweise in der Stundentafel in größeren Fächerverbünden integriert ist, was wiederum möglicherweise zu einer geringeren Berücksichtigung führt. Allerdings zeigt sich auch kein systematischer Effekt der COVID-19-Pandemie in den Vergleichsgruppen der EU- und OECD-Staaten. Hier ist die Veränderung der naturwissenschaftlichen Leistungen zwischen 2019 und 2023 unbedeutend. Über alle an TIMSS 2019 und TIMSS 2023 teilnehmenden Staaten bzw. Benchmark-Regionen hinweg weisen annähernd ebenso viele Staaten und Regionen einen Leistungsanstieg auf wie die, die einen Leistungsverlust verzeichnen. Auch aus den Daten im Bereich Mathematik in TIMSS 2023 (siehe Kapitel 3 in diesem Band) werden keine spezifischen Pandemieeffekte ersichtlich. Dies steht im Widerspruch zu anderen Studien, die auf Leistungsrückstände aufgrund von Schulschließungen verweisen, allerdings sind auch hier die Befunde nicht so eindeutig, wie man annehmen könnte. So zeigen sich Unterschiede in den Ergebnissen, die auf Spezifika der Erhebungszeiträume, der Fächer und der Länder bzw. Staaten zurückzuführen sind (Depping et al., 2021; Hammerstein et al., 2021; Schult et al., 2022).

Werden in Deutschland die Veränderungen der naturwissenschaftlichen Kompetenzen seit 2007 betrachtet, so zeigt sich ein Trend leicht sinkender Leistungen, gleichwohl liegt dieser nur bei einem Rückgang von 13 Punkten über einen Zeitraum von 16 Jahren. Festzuhalten bleibt demnach, dass sich die COVID-19-Pandemie nicht in merkbaren Kompetenzverlusten in Deutschland niedergeschlagen hat, wohl aber ein langfristiger Trend leicht zurückgehender Leistungen beobachtbar ist. Die Streuung der Leistungen hat in Deutschland in TIMSS 2023 weiter gegenüber TIMSS 2019 zugenommen. Dies ist vor allem dem Umstand geschuldet, dass sich der Anteil von Schüler:innen mit Bildungsrisiken u. a. durch Migrationsbewegungen vergrößert. Auch mögen hier wieder Effekte der COVID-19-Pandemie eine Rolle spielen, da Familien mit höherem sozioökonomischem Status während der COVID-19-Pandemie bessere Unterstützung leisten konnten als solche mit niedrigem sozioökonomischen Status (Lewalter et al., 2023). Welche Rolle der Lehrkräftemangel und die damit verbundenen Rekrutierungen von Quer- und Seiteneinsteigenden in den letzten Jahren spielen, lässt sich schwer abschätzen. Insgesamt erweist sich das System der Grundschule in Krisenzeiten als robust. Wie auch in den Jahren davor zeigen vor allem die asiatischen Staaten, dass man bei hohen durchschnittlichen Kompetenzen die Leistungsstreuung vergleichsweise klein halten kann. In diesen Staaten scheint die Förderung aller Leistungsgruppen besser zu gelingen als in Deutschland.

Kompetenzstufen:

Die Verteilung der Schüler:innen auf die fünf Kompetenzstufen zeigt, dass lediglich 9 Prozent der Schüler:innen die höchste Kompetenzstufe erreichen, dies ist ähnlich zu den Ergebnissen in Mathematik (siehe Kapitel 3 in diesem Band) und zu den beiden Vergleichsgruppen der EU- und OECD-Staaten. Die Ergebnisse der asiatischen Staaten, aber auch England und Australien verdeutlichen, dass Spitzenförderung noch besser gelingen kann. Gleichzeitig erreichen in Deutschland 9 Prozent der Schüler:innen nur die Kompetenzstufe I, was etwas mehr ist als in der Mathematik (6 %). Fasst man die beiden niedrigsten Kompetenz-

stufen zusammen, befindet sich hier ein Anteil von 30 Prozent. Das ist eine große Gruppe von Schüler:innen mit ungünstigen Voraussetzungen für den nachfolgenden naturwissenschaftlichen Unterricht. Die Ergebnisse aus PISA (Lewalter et al., 2023) zeigen einen ähnlichen Anteil von 23 Prozent auf den untersten Kompetenzstufen. Auch hier macht der Vergleich mit anderen Staaten deutlich, dass die Förderung der leistungsschwachen Schüler:innen in Deutschland nach wie vor herausfordernd ist.

Inhalts- und Anforderungsbereiche:

Die Rahmenkonzeption von TIMSS erlaubt die differenzierte Darstellung der Ergebnisse nach den drei Inhaltsbereichen *Biologie, Physik/Chemie* und *Geografie* und nach den drei kognitiven Anforderungsbereichen *Reproduzieren, Anwenden* und *Problemlösen*. Zwischen den drei Inhaltsbereichen zeigen sich in TIMSS 2023 für Deutschland keine substanziellen Unterschiede in den relativen Leistungen. Der Inhaltsbereich *Geografie* hat gegenüber den beiden anderen Inhaltsbereichen aufgeholt, sodass sich festhalten lässt, dass Ziele des Sachunterrichts in den drei Inhaltsbereichen gleichermaßen erreicht werden. Ob sich hierin auch eine vergleichbare Gewichtung der drei Inhaltsbereiche im Unterricht abbildet, kann lediglich vermutet werden, da TIMSS 2023 hierfür keine empirische Evidenz liefert. Schaut man sich die Entwicklung der drei Inhaltsbereiche seit 2007 an, so zeigt sich, dass die insgesamt sinkenden Leistungen in den letzten 16 Jahren in *Geografie* etwas schwächer ausfallen als in den anderen beiden Inhaltsbereichen. Für alle drei Inhaltsbereiche zeigt sich gegenüber 2019, dass die Leistungen hier weitgehend stabil geblieben sind bzw. die Veränderungen statistisch nicht signifikant sind.

Auch für die kognitiven Anforderungsbereiche zeigen sich in Deutschland in TIMSS 2023 kaum differenzielle Befunde. Zwar fällt der Mittelwert beim *Anwenden* mit 512 Punkten gegenüber dem Gesamtmittelwert in den Naturwissenschaften (515 Punkte) signifikant niedriger aus, die Differenz von drei Punkten ist aber praktisch unbedeutsam. Die fehlenden Unterschiede zwischen den drei Bereichen hatten sich auch schon in früheren Studienzyklen gezeigt. Festzuhalten ist daher, dass der Sachunterricht in Deutschland die drei kognitiven Anforderungsbereiche gleichermaßen bedient. Zu wünschen ist, dass er dies auch künftig leisten kann.

Einstellungen und Selbstkonzept:

Die sachunterrichtsbezogenen Einstellungen und das sachunterrichtsbezogene Selbstkonzept sind erfreulicherweise nach wie vor hoch ausgeprägt. Im Trend zeigt sich allerdings eine leichte Abnahme für beide Merkmale. Ob dies mit geänderten Inhalten und Anforderungen zusammenhängt, lässt sich aus den Daten nicht ableiten. Insbesondere Schüler:innen der unteren Kompetenzstufen haben weniger positive Einstellungen und ein geringeres sachunterrichtsbezogenes Selbstkonzept. Dies ist erwartbar, wenn man das Zusammenspiel zwischen Leistungen und motivationalen Orientierungen bedenkt. Es macht allerdings auch deutlich, dass insbesondere bei Schüler:innen der unteren Kompetenzstufen nicht nur Leistungen, sondern auch motivationale Orientierungen im Blick behalten werden sollten.

Nachhaltigkeitskompetenzen und Umwelteinstellungen:

Seit TIMSS 2023 wird eine Teilmenge der Naturwissenschaftsitems verwendet, um eine Subskala zu Nachhaltigkeitskompetenzen zu bilden. In TIMSS 2023 enthält diese Skala 44 Aufgaben aus dem bestehenden Itempool. Die Auswertungen dieser Subskala ergeben für Schüler:innen in Deutschland einen Testwert von 513 Punkten, der sich nicht signifikant vom Gesamtwert in den Naturwissenschaften (515 Punkte) unterscheidet. Dies gilt auch für die meisten an TIMSS teilnehmenden Staaten und Regionen mit Ausnahme der Spitzenreiter Singapur, Japan und Taiwan, in denen der Wert für die Nachhaltigkeitskompetenzen signifikant niedriger ausfällt als der Gesamtwert. Die Befunde in den übrigen Staaten deuten sehr stark darauf hin, dass naturwissenschaftliche Kompetenzen wichtige Grundlagen darstellen, um ein besseres Verständnis über den menschlichen Einfluss auf die Umwelt, über natürliche Ressourcen sowie über umweltverträgliche Handlungen zu erwerben.

Hinsichtlich der umweltbezogenen Einstellungen zeigt sich, dass die große Mehrzahl der Schüler:innen in Deutschland eine hohe Präferenz zum Schutz der Natur und deren Wertschätzung zeigt. Bemerkenswert ist dabei der positive Zusammenhang zwischen Nachhaltigkeitskompetenzen und Umwelteinstellungen: Je höher die Nachhaltigkeitskompetenzen ausgeprägt sind, desto höher ist auch die Präferenz zum Schutz der Natur und deren Wertschätzung. Diese Ergebnisse machen deutlich, wie wichtig naturwissenschaftliche Bildung im Bereich der Nachhaltigkeit für Umwelteinstellungen und in der Folge für umweltbewusstes Handeln ist. Bildung für nachhaltige Entwicklung bleibt demnach ein Auftrag für die naturwissenschaftlichen Fächer, aber auch ein überfachlicher Auftrag der Schule.

Literatur

Bayerisches Staatsministerium für Bildung und Kultus, Wissenschaft und Kunst (Hrsg.). (2014). *LehrplanPLUS. Lehrplan für die bayerische Grundschule*. https://www.lehrplanplus.bayern.de/schulart/grundschule

Blaseio, B. (2009). Neue Entwicklungstendenzen der Inhalte des Sachunterrichts. *Zeitschrift für Grundschulforschung*, *2*(1), 117–131.

Bos, W., Bonsen, M., Baumert, J., Prenzel, M., Selter, C., & Walther, G. (Hrsg.). (2008). *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Waxmann.

Bos, W., Wendt, H., Köller, O., & Selter, C. (Hrsg.). (2012). *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Waxmann. https://www.waxmann.com/buch2814.pdf

Böschl, F., Forbes, C., & Lange-Schubert, K. (2023). Investigating scientific modeling practices in U.S. and German elementary science classrooms: A comparative, cross-national video study. *Science Education,* 107, 368–400. https://doi.org/10.1002/SCE.21780

Bybee, R. & McCrae, B. (2011). Scientific Literacy and Student Attitudes: Perspectives from PISA 2006 science. *International Journal of Science Education*, *33*(1), 7–26. https://doi.org/10.1080/09500693.2010.518644

Costa, J., Alscher, P., & Thums, K. (2024). Global competences and education for sustainable development. A bibliometric analysis to situate the OECD global competences in the scientific discourse. *Zeitschrift für Erziehungswissenschaft*. https://doi.org/10.1007/s11618-024-01220-z

Depping, D., Lücken, M., Musekamp, F., & Thonke, F. (2021). Kompetenzstände Hamburger Schüler*innen vor und während der Corona-Pandemie. In D. Fickermann & B. Edelstein (Hrsg.), *Schule während der Corona-Pandemie. Neue Ergebnisse und*

Überblick über ein dynamisches Forschungsfeld (S. 51–79). Waxmann. https://doi.org/10.31244/9783830993315.03

Dominke, H., & Steffensky, M. (2024). The science-specific home learning environment of elementary school children – how are science experiences and science talk associated with the children's science achievement? *Research in Science & Technological Education*, 1–20. https://doi.org/10.1080/02635143.2024.2375512

GDSU (Gesellschaft für die Didaktik des Sachunterrichts). (2013). *Perspektivrahmen Sachunterricht*. Klinkhardt.

Hammerstein, S., König, C., Dreisörner, T., & Frey, A. (2021). Effects of COVID-19-related school closures on student achievement – A systematic review. *Frontiers in Psychology*, *12*, 746289. https://doi.org/10.3389/fpsyg.2021.746289

Harlen, W. (2015). *Working with big ideas*. Science Education Programme (SEP) of IAP.

Johnston, J. D. (2020). Climate change literacy to combat climate change and its impacts. In W. Leal Filho, A. M. Azul, L. Brandli, P. G. Özuyar & T. Wall (Hrsg.), *Climate Action* (S. 200–212). Springer. https://doi.org/10.1007/978-3-319-95885-9_31

Joint Research Centre. (2022). *GreenComp, the European sustainability competence framework*. European Commission. https://doi.org/10.2760/13286

Junge, K., Schmerse, D., Lankes, E.-M., Carstensen, C. H., & Steffensky, M. (2021). How the home learning environment contributes to children's early science knowledge — Associations with parental characteristics and science-related activities. *Early Childhood Research Quarterly, 56*, 294–305. https://doi.org/10.1016/j.ecresq.2021.04.004

Kleickmann, T., Brehl, T., Saß, S., Prenzel, M., & Köller, O. (2012). Naturwissenschaftliche Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 123–169). Waxmann. https://www.waxmann.com/buch2814.pdf

Kleickmann, T., Steffensky, M. & Praetorius, A.-K. (2020). Quality of teaching in elementary science education: More than three basic dimensions? In A.-K. Praetorius & E. Klieme (Hrsg.), *Empirische Forschung zu Unterrichtsqualität. Theoretische Grundfragen und quantitative Modellierungen* (S. 37–55). Beltz Juventa.

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2004). *Beschlüsse der Kultusministerkonferenz. Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss*. *Beschluss vom 16.12.2004*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Chemie.pdf

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik). (2024). *Weiterentwickelte Bildungsstandards in den Naturwissenschaften für das Fach Chemie (MSA)*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2024/2024_06_13-WeBiS_Chemie_MSA.pdf

Kriegbaum, K., Becker, N., & Spinath, B. (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review*, *25*, 120–148. https://doi.org/10.1016/j.edurev.2018.10.001

Lewalter, D., Diedrich, J., Goldhammer, F., Köller, O., & Reiss, K. (Hrsg.). (2023). *PISA 2022. Analyse der Bildungsergebnisse in Deutschland.* Waxmann. https://doi.org/10.31244/9783830998488

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416. https://doi.org/10.1111/j.1467-8624.2005.00853.x

Maurer, M., & Bogner, F. X. (2020). Modelling environmental literacy with environmental knowledge, values and (reported) behavior. *Studies in Educational Evaluation, 65*, 100863. https://doi.org/10.1016/j.stueduc.2020.100863

Milfont, T. L. & Duckitt, J. (2004). The structure of environmental attitudes: A first- and second-order confirmatory factor analysis. *Journal of Environmental Psychology, 24*(3), 289–303. https://doi.org/10.1016/j.jenvp.2004.09.001

Ministerium für Bildung, Wissenschaft und Kultur Mecklenburg-Vorpommern (Hrsg.). (2020). *Rahmenplan für die Primarstufe*. Schwerin.

Ministerium für Kultus, Jugend und Sport Baden-Württemberg (Hrsg.). (2016). *Bildungsplan der Grundschule Sachunterricht*.

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (Hrsg.). (2012). *Lehrpläne für die Primarstufe in Nordrhein-Westfalen*.

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen (Hrsg.). (2021). *Lehrpläne für die Primarstufe in Nordrhein-Westfalen.*

Möller, J. & Trautwein, U. (2009). Selbstkonzept. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 179–203). Springer. https://doi.org/10.1007/978-3-540-88573-3_8

Mullis, I. V. S., Martin, M. O. & von Davier, M. (Hrsg.). (2021). *TIMSS 2023 Assessment Frameworks*. TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2023/frameworks/pdf/T23_Frameworks.pdf

National Research Council. (2012). *A framework for K-12 science education. Practices, crosscutting concepts, and core ideas*. The National Academies Press.

Niedersächsischen Kultusministerium (Hrsg.). (2017). *Kerncurriculum für die Grundschule Schuljahrgänge 1–4 Sachunterricht*. https://cuvo.nibis.de/cuvo.php?p=search&k0_0=Fach&v0_0=Sachunterricht&

NOAA (The National Oceanic Atmospheric Administration of the U.S. Department of Commerce). (2014). *What is climate science literacy?* https://www.climate.gov/teaching/essential-principles-climate-literacy/what-climate-science-literacy

OECD. (2019). *PISA 2018 Assessment and Analytical Framework*. OECD. https://doi.org/10.1787/b25efab8-en

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*(9), 1049–1079. https://doi.org/10.1080/0950069032000032199

Owens, D. C. & Sadler, T. D. (2024). Socio-scientific issues instruction for scientific literacy: 5E Framing to enhance teaching practice. *School Science and Mathematics*, *124*(3), 203–210. https://doi.org/10.1111/ssm.12626

Robitzsch, A., Lüdtke, O., Schwippert, K., Goldhammer, F., Kroehne, U., & Köller, O. (2020). Leistungsveränderungen in TIMSS zwischen 2015 und 2019: Die Rolle des Testmediums und des methodischen Vorgehens bei der Trendschätzung. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern im internationalen Vergleich* (S. 169–186). Waxmann. https://doi.org/10.31244/9783830993193

Sächsisches Staatsministerium für Kultus (Hrsg.). (2019). *Lehrplan Grundschule Sachunterricht*.

Schmerse, D., Dominke, H., Mohr, J., & Steffensky, M. (2024). Children's understanding of scientific inquiry: The role of instructional support and comparison making. *Journal of Educational Psychology, 116*(2), 233–255. https://doi.org/10.1037/edu0000836

Schult, J., Mahler, N., Fauth, B., & Lindner, M. A. (2022). Did students learn less during the COVID-19 pandemic? Reading and mathematics competencies before and after the first pandemic wave. *School Effectiveness and School Improvement, 33*(4), 544–563. https://doi.org/10.1080/09243453.2022.2061014

Schwippert, K., Kasper, D., Köller, O., McElvany, N., Selter, C., & Steffensky, M. (Hrsg.). (2020). *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Waxmann. https://doi.org/10.31244/9783830993193

Seaton, M., Marsh, H. W. & Craven, R. G. (2009). Big-fish-little-pond effect. Generalizability and moderation – Two sides of the same coin. *American Educational Research Journal*, *47*(2), 390–433. https://doi.org/10.3102/0002831209350493

Selter, C., Walter, D., Heinze, A., Brandt, J., & Jentsch, A. (2020). Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern im internationalen Vergleich* (S. 57–113). Waxmann. https://doi.org/10.31244/9783830993193

Stanat, P., Schipolowski, S., Rjosk, C., Weirich, S., & Haag, N. (Hrsg.). (2017). *IQB-Bildungstrend 2016. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im zweiten Ländervergleich*. Waxmann. https://www.iqb.hu-berlin.de/bt/BT2016/Bericht/BT2016_Bericht.pdf

Stanat, P., Schipolowski, S., Schneider, R., Sachse, K. A., Weirich, S. & Henschel, S. (Hrsg.). (2022). *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich*. Waxmann https://doi.org/10.31244/9783830996064

Steffensky, M., Kleickmann, T., Kasper, D., & Köller, O. (2016). Naturwissenschaftliche Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015 Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 133–184). Waxmann. https://doi.org/10.25656/01:14022

Steffensky, M., Scholz, L. A., Kasper, D., & Köller, O. (2020). Naturwissenschaftliche Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter & M. Steffensky (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S 115–168). Waxmann. https://doi.org/10.31244/9783830993193

UN DESA (United Nations Department of Economic and Social Affairs). (2021). The *sustainable development goals report 2021*. https://unstats.un.org/sdgs/report/2021/The-Sustainable-Development-Goals-Report-2021.pdf

U.S. Global Change Research Program. (2009). *Climate literacy. Essential principles of climate science: a guide for individuals and communities.* https://www.climate.gov/teaching/essential-principles-climate-literacy/essential-principles-climate-literacy

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*(8), 1205–1228. https://doi.org/10.1177/0022022111428083

von Davier, M., Fishbein, B., & Kennedy, A. (Hrsg.). (2024). *TIMSS 2023 technical report (methods and procedures)*. TIMSS & PIRLS International Study Center.

von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U., & Yin, L. (2024). *TIMSS 2023 International Results in Mathematics and Science.* TIMSS & PIRLS International Study Center.

Wendt, H., Bos, W., Selter, C., Köller, O., Schwippert, K. & Kasper, D. (Hrsg.). (2016). *TIMSS 2015 Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Waxmann. https://doi.org/10.25656/01:14022

Wigfield, A., & Eccles, J. S. (2000). Expectancy – value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wiseman, M., & Bogner, F. X. (2003). A higher-order model of ecological values and its relationship to personality. *Personality and Individual Differences, 34*(5), 783–794. https://doi.org/10.1016/S0191-8869(02)00071-5

Wittwer, J., Saß, S., & Prenzel, M. (2008). Naturwissenschaftliche Kompetenz im internationalen Vergleich. Testkonzeption und Ergebnisse. In W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther (Hrsg.), *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 87–124). Waxmann.

Kapitel 5
Lehren und Lernen mit digitalen Medien in der Grundschule

Donieta Jusufi, Knut Schwippert, Birgit Eickelmann, Julia Gerick und Ulf Kröhne

5.1 Einleitung

Der *Trends in International Mathematics and Science Study* (TIMSS) im Jahr 2023 kommt hinsichtlich der Untersuchung digitaler Medien eine besondere Bedeutung zu. Erstmalig wird die Untersuchung dieses Bereiches im Rahmen eines Teilprojekts zur Digitalisierung in der Grundschule gesondert von der *Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland* (KMK) gefördert. Ausgangspunkt ist, dass dem Umgang mit digitalen Medien in Lehr- und Lernkontexten – auch aufgrund der Entwicklungen im Zuge der COVID-19-Pandemie – eine zunehmende Relevanz in Grundschulen beigemessen wird. Mit einer Ländervereinbarung der KMK wurde im Jahr 2020 erstmals für Deutschland festgehalten, dass Schüler:innen in einer zunehmend digitalisierten Welt von Beginn ihrer Schullaufbahn an dazu befähigt werden sollen, digitale Werkzeuge zu nutzen, mit ihnen selbstorganisiert zu lernen, sich kritisch mit ihnen auseinanderzusetzen und die Bereitschaft zu entwickeln, diese für ihr Lernen zu verwenden (KMK, 2020).

Mit der im Rahmen von TIMSS 2023 realisierten Erhebung des Lehrens und Lernens mit digitalen Medien im Mathematik- und naturwissenschaftlichen Sachunterricht an Grundschulen werden sowohl Fragen im internationalen Vergleich eingesetzt als auch spezifische Aspekte in Deutschland vertieft betrachtet. Ziel des Teilprojekts *TIMSS 2023 – Digitalisierung in der Grundschule*, dessen erste Ergebnisse im vorliegenden Kapitel vorgestellt werden, die im Zusammenhang mit der Gesamtstudie TIMSS 2023 zu sehen sind, ist es, zu untersuchen, wie digitale Medien in Mathematik und dem naturwissenschaftlichen Sachunterricht an Grundschulen in Deutschland und im internationalen Vergleich in schulischen Lehr- und Lernkontexten genutzt werden und welche Veränderungen in der Verfügbarkeit von häuslichen und schulischen digitalen Ressourcen im Trend seit 2007 bzw. im Vergleich zu TIMSS 2019 festzustellen sind. Dazu werden die Nut-

zungshäufigkeiten digitaler Medien aus verschiedenen Perspektiven abgebildet: Aus der Schülerperspektive wird die Nutzungshäufigkeit digitaler Medien durch die Schüler:innen für schulische Zwecke berichtet und in vertiefter Betrachtung die Ausdifferenzierung der verschiedenen Medien vorgenommen. Aus der Lehrkräfteperspektive wird die Nutzungshäufigkeit digitaler Medien im Fachunterricht durch Schüler:innen für verschiedene Lernangebote berichtet. Eine Erfassung digitaler Kompetenzen, wie sie in der *International Computer and Information Literacy Study* (ICILS) für die Sekundarstufe I realisiert wird (Eickelmann, Fröhlich et al., 2024), ist nicht Bestandteil des TIMSS-2023-Teilprojekts in Deutschland. Jedoch können über die international eingesetzten Schülerfragebögen von TIMSS 2023 Aspekte des selbsteingeschätzten Kompetenzerlebens im Umgang mit digitalen Medien einbezogen und hier sowohl für internationale Vergleiche als auch für national vertiefende Analysen genutzt werden. Weiterhin wird die Fortbildungsteilnahme der Mathematik- und Sachunterrichtslehrkräfte[1] für die *Integration von Informationstechnologien* im internationalen Vergleich und im Trend abgebildet.

Trends in der Nutzung digitaler Medien: Eine für die nachfolgenden Analysen besondere Herausforderung bei der Darstellung von Trends liegt in der Dynamik der Entwicklungen digitaler Medien und ihres Einsatzes für das schulische Lehren und Lernen in den vergangenen Jahren. Durch diese Dynamik wurde in den im Rahmen von TIMSS 2023 eingesetzten Hintergrundfragebögen und denen der vergangenen Zyklen teils nach unterschiedlichen digitalen Geräten (z. B. explizit Computer oder Tablets) gefragt. Die jeweiligen Fragen werden, um einen Trend seit 2007 abbilden zu können, trotz ihrer unterschiedlichen Schwerpunkte übergreifend betrachtet. Somit sind diese Trendbetrachtungen als generelle Entwicklung in Bezug auf die Nutzung digitaler Medien im Kontext des Lehrens und Lernens zu verstehen (Eickelmann et al., 2019; Schwippert et al., 2020). In TIMSS 2023 wird zudem primär die Formulierung *digitale Medien* genutzt, die neben technischen Geräten (z. B. Computer, Smartphones, Notebooks), auch schul- bzw. unterrichtsrelevante Software (bspw. Lernprogramme) umfasst.

(Nach-)Wirkungen der COVID-19-Pandemie: In TIMSS 2023 besteht die Möglichkeit, die Verfügbarkeit und Nutzung digitaler Medien in der Grundschule nach der COVID-19-Pandemie und der Rückkehr zum Regelbetrieb im Vergleich zum Zeitpunkt vor der Pandemie, wie er in TIMSS 2019 dokumentiert wurde, in den Blick zu nehmen. Hervorzuheben ist an dieser Stelle, dass TIMSS 2023 für den Grundschulbereich aktuelle, repräsentative Daten hervorbringt, die nicht mehr durch die besondere Umstrukturierung von Schule und Unterricht während der COVID-19-Pandemie verzerrt sind, da alle Grundschulen in Deutschland während der Haupterhebung im Frühjahr 2023 wieder vollständig zum Regelbetrieb übergegangen waren (für einen Überblick siehe Kapitel 2 in diesem Band).

Rahmenbedingungen zur Nutzung digitaler Medien im schulischen Kontext: Bildungspolitisch werden der Nutzung und dem Einsatz digitaler Medien zum schulischen Lehren und Lernen mit den Strategiepapieren zur *Bildung in der digitalen Welt* der KMK eine hohe Relevanz beigemessen (KMK, 2016, 2021). Diese Strategiepapiere dienen den Ländern der Bundesrepublik Deutschland als Basis, um

1 In diesem Kapitel wird aus Gründen der Einheitlichkeit von *Sachunterricht* und *Sachunterrichtslehrkräften* gesprochen, auch wenn es im internationalen Kontext keinen Sachunterricht, sondern naturwissenschaftlichen Unterricht gibt.

weitergehende Konzepte für die Veränderung von Curricula sowie für die Aus-, Fort- und Weiterbildung der Lehrkräfte zur Verfügung zu stellen. Kurz vor der Veröffentlichung der KMK-Strategie (KMK, 2016) wurde über die digitale Bildungsoffensive des Bundes der DigitalPakt Schule (BMBF, 2016) als bundesweites Ausstattungsprogramm angebahnt. Die Verfügbarkeit von digitalen Geräten wird in den genannten Strategiepapieren auch für den Grundschulbereich als die Grundlage für eine lernförderliche schulische Nutzung digitaler Medien angesehen. Eine zentrale Aufgabe, die dabei an Schulen herangetragen wird, ist das Lernen *mit* und *über* digitale Medien, welches viele Jahre nicht als fester Bestandteil im Unterricht an Grundschulen in Deutschland verankert war (Irion & Peschel, 2016). Dabei werden bereits nach Eickelmann und Vennemann (2014) digitale Kompetenzen in der Grundschule als Querschnittskompetenzen betrachtet. Nicht zuletzt hat schließlich die globale COVID-19-Pandemie auch in Deutschland als Beschleuniger dafür gesorgt, dass die Potenziale digitaler Medien für die Gestaltung schulischer Lehr-Lernprozesse und die zu dem Zeitpunkt bedeutsame Rolle eines zukunftsorientierten Lernens im schulischen Kontext in der Fläche auch für den Grundschulbereich verstärkt diskutiert wurden (Eickelmann & Gerick, 2020).

5.2 Relevante Merkmale im Zusammenhang mit digitalen Medien

Bildungsrelevante häusliche Ressourcen: Bei der Analyse sozialer Disparitäten werden in Schulleistungsstudien Informationen über schulische sowie häusliche bildungsrelevante Ressourcen insbesondere anhand des Bildungsabschlusses der Erziehungsberechtigten, deren beruflicher Stellung sowie allgemeiner Informationen über den sozioökonomischen Hintergrund der Familien und des heimischen Buchbesitzes betrachtet.[2] Welche Rolle darüber hinaus die Verfügbarkeit digitaler Medien im häuslichen Umfeld für die von Grundschulkindern gezeigten mathematischen und naturwissenschaftlichen Kompetenzen hat, wird in TIMSS 2023 in den Blick genommen. Die Ausstattung mit digitalen Medien im häuslichen Kontext erscheint nicht zuletzt aufgrund der allgemein in der Gesellschaft gemachten Erfahrungen von Lehrenden und Lernenden hinsichtlich der Einschränkungen während der COVID-19-Pandemie für Deutschland von hoher Relevanz (Züchner & Jäkel, 2021). Durch die Verfügbarkeit digitaler Medien im häuslichen Kontext konnte der Unterricht auf Distanz erst ermöglicht werden (Caruso & Bruns, 2021), diese war aber längst nicht für alle Schüler:innen gegeben, weshalb in diesem Kapitel die Verfügbarkeit bildungsrelevanter Ressourcen im Haushalt betrachtet wird.

Nutzungshäufigkeit digitaler Medien in Lehr- und Lernkontexten: Der Einsatz digitaler Medien in Schule und Unterricht birgt zahlreiche Potenziale, beispielsweise die Modernisierung des Unterrichts, die Entwicklung der Unterrichtsqualität oder aber die Unterstützung erweiterter, schülerorientierter Lernkulturen (Eickelmann & Gerick, 2020). Diese werden für Deutschland systematisch in dem KMK-Ergänzungspapier von 2021 zum Lehren und Lernen mit digitalen Medien aufgegriffen (KMK, 2021). In umfangreichen Metaanalysen konnte

2 Bildungsabschluss der Erziehungsberechtigten: *International Standard Classification of Education* (ISCED; UNESCO, 2003); berufliche Stellung: *International Standard Classification of Occupation* (ISCO; ILO, 2012) und *Erikson-Goldthorpe-Portocarero-Cassification* (EGP; Erikson & Goldthorpe, 2002; Erikson et al., 1979); Informationen über den sozioökonomischen Hintergrund der Familien: *Highest International Socio-Economic Index of Occupational Status* (HISEI; Ganzeboom et al., 1992).

gezeigt werden, dass kleine bis moderate Effekte zugunsten der Nutzung von Technologie gegenüber sogenannten ‚traditionellen' Unterrichtspraktiken ohne Technologie festgestellt wurden (Schaumburg, 2018; Tamim et al., 2011). Aktuellere Befunde aus internationalen Schulleistungsstudien, wie der *Internationalen-Grundschul-Leseuntersuchung* (IGLU, engl. *Progress in International Reading Literacy Study*; PIRLS; McElvany et al., 2023) 2021 geben für den Grundschulbereich nicht nur Hinweise darauf, dass die Ausstattung mit digitalen Medien in Deutschland im internationalen Vergleich sowohl im Allgemeinen (Anzahl der Computer pro Viertklässler:in) als auch in Bezug auf den Leseunterricht unterdurchschnittlich ist (Lorenz et al., 2023). Auch die Nutzungshäufigkeit digitaler Medien fiel sowohl im internationalen Vergleich insgesamt als auch differenziert nach unterschiedlichen Leseaktivitäten für Deutschland im Jahr 2021 gering aus. In diesem Kapitel werden die zuvor genannten Aspekte im Hinblick auf Zusammenhänge der in TIMSS 2023 erhobenen mathematischen und naturwissenschaftlichen Kompetenzen der Grundschulkinder untersucht. Dabei ist bereits an dieser Stelle darauf hinzuweisen, dass es sich nicht um Ursache-Wirkungsanalysen handelt und die Ergebnisse nicht mit dieser Perspektive interpretiert werden können.

Selbsteingeschätztes Kompetenzerleben aus Schülerperspektive: Der Einsatz von komplexen und lebensnahen Aufgaben im Kontext der Nutzung digitaler Medien im Fachunterricht kann eine besonders motivierende Wirkung auf Lernende haben und damit verbunden das Bedürfnis nach einer erfolgreichen Bearbeitung und somit eine Steigerung des Kompetenzerlebens bewirken (Herzig & Grafe, 2010). Vor diesem Hintergrund ist das selbsteingeschätzte Kompetenzerleben von Schüler:innen im Umgang mit digitalen Medien eine wichtige Information, die in TIMSS 2023 in Deutschland erhoben wurde. Die Befunde aus dem *Programme for International Student Assessment* (PISA) 2022 geben für die Sekundarstufe I zudem Hinweise darauf, dass Aspekte selbsteingeschätzter digitalisierungsbezogener Kompetenzen und der Lernmotivation begünstigende Faktoren für die Entwicklung von *Information and Communication Technology* (ICT) *Literacy* darstellen (Lewalter et al., 2023). Auch konnte im Rahmen von ICILS 2018 ein positiver Zusammenhang zwischen der Selbstwirksamkeit im Umgang mit digitalen Medien (basale Fähigkeiten) und computer- und informationsbezogenen Kompetenzen von Achtklässler:innen festgestellt werden (Gerick et al., 2019). Die Einschätzung von digitalisierungsbezogenen Kompetenzen aus Schülerperspektive wurde bisher in Schulleistungsstudien für den Grundschulbereich nicht in den Blick genommen.

Digitalisierungsbezogene Fortbildung von Lehrkräften: Neben der grundsätzlichen Verfügbarkeit digitaler Ressourcen lassen sich vor allem auch die Kompetenzen von Lehrkräften hinsichtlich des Umgangs und der didaktischen Einbettung digitaler Medien als ein Kriterium für ihre Implementation in den Unterricht identifizieren (Herzig, 2004; Lorenz, 2018). Lachner et al. (2020) betrachten den Einsatz digitaler Medien im Unterricht im Sinne eines *Angebots-Nutzungs-Modells,* wie es ursprünglich Helmke (2010) für den Kontext der Betrachtung von Unterrichtsqualität entwickelt hat. Die Autor:innen zeigen, dass der Einsatz digitaler Medien die Unterrichtsqualität (z. B. die kognitive Aktivierung, konstruktive Unterstützung und Klassenführung) erhöhen kann. Um den Schüler:innen das lernförderliche Lernen *mit* digitalen Medien ermöglichen zu können, benötigen Lehrkräfte insbesondere im Grundschulbereich eine wissenschaftliche und gleichsam praxisnahe Aus- und Fortbildung (Kunkel & Peschel, 2020;

Peschel, 2019). So stellen Fortbildungen zu digitalen Medien nach Drossel et al. (2019) professionelle Lerngelegenheiten für Lehrkräfte dar, die die Einbettung digitaler Medien im Unterricht begünstigen und für eine effiziente und lernförderliche Integration digitaler Medien in unterrichtliche Lehr- und Lernprozesse besonders relevant sind (Endberg et al., 2021; KMK, 2021; Scheiter, 2021).

Die Teilnahme von Lehrkräften an Fortbildungen ist Bestandteil ihrer Professionalisierung und wird als wichtiger Aspekt der beruflichen Tätigkeit gesehen. Bildungspolitisch ist diese in rechtlichen Rahmenvorgaben verankert und von der KMK geregelt (KMK, 2020). Im Jahresbericht der KMK (2023) wurden Konzepte für die Aus-, Fort- und Weiterbildung der Lehrkräfte festgehalten, die sich speziell auf den Ausbau von Angeboten für den Erwerb digitalisierungsbezogener und fachspezifischer Lehrkompetenzen beziehen. Die inhaltlichen Schwerpunkte liegen dabei auf dem Einsatz digitaler Medien im Fachunterricht, der Befassung mit digitalen Technologien und ihrer methodischen Verwendung im Unterricht sowie auf dem Kennenlernen des digitalen Medienangebots und der Bewertung unterrichtlicher Eignung.

Digitalisierungsbezogene Fortbildungen im Spezifischen sollen den Herausforderungen zunehmender Digitalisierung und gesellschaftlicher Veränderungen begegnen und zielen auf den Ausbau und die Vertiefung der Kenntnisse und Fertigkeiten von Lehrkräften beim Einsatz digitaler Medien im schulischen Kontext ab (Gerick et al., 2024). Der Anteil von Lehrkräften, die an digitalisierungsbezogenen Fortbildungen teilnehmen, war bundesweit jedoch zunächst sehr gering, wobei das in Deutschland begrenzt verfügbare Fortbildungsangebot zu berücksichtigen war (Gerick et al., 2019). Drossel et al. (2012) zeigten diesbezüglich bereits im Rahmen einer vertiefenden Analyse anhand der TIMSS-2011-Daten, dass Mathematiklehrkräfte in Deutschland zwar im Vergleich zu anderen europäischen teilnehmenden Staaten und Regionen häufiger an Fortbildungsveranstaltungen teilnahmen, die das Fach Mathematik, die Mathematikdidaktik oder Aspekte individueller Förderung betrafen. Hinsichtlich der Integration von Informationstechnologien in den Unterricht zeichnete sich jedoch fast keine Fortbildungsaktivität ab. Auch in TIMSS 2015 wurde festgestellt, dass kaum Fortbildungen hinsichtlich der Integration von Informationstechnologien in den Unterricht von den Lehrkräften besucht wurden. Mit 1.5 Prozent (international: 35.9 %) stellte Deutschland „im internationalen Vergleich das Schlusslicht“ (Porsch & Wendt, 2016, S. 195) dar. Auch im Rahmen von IGLU 2021 lag der Anteil der Schüler:innen, die von Lehrkräften unterrichtet wurden, die in den letzten zwei Jahren vor der Erhebung ein Fortbildungsangebot im Zusammenhang mit digitalen Kompetenzen besucht hatten, bei deutlich weniger als der Hälfte (41.6 %; McElvany et al., 2023).

Im Zuge der verschiedenen gesellschaftlichen Herausforderungen und der immer heterogener werdenden Zusammensetzung der Schülerschaft kommt es zunehmend zu Veränderungen der Anforderungen an Lehrkräfte (Kasper et al., 2020). Damit einher geht bereits seit Jahren ein Wandel des Professionalitätsprofils von Lehrkräften (Melzer & Hillenbrand, 2015), weshalb der Teilnahme an Fortbildungen eine besonders relevante Rolle zugesprochen werden kann. Auch wenn vielfach die Annahme formuliert wurde, dass die COVID-19-Pandemie wie ein „Digitalisierungsschub“ (Eickelmann & Gerick, 2020, S. 154) hätte wirken können, kamen vor allem große Lücken, Entwicklungs- und Unterstützungsbedarfe im Grundschulbereich zum Vorschein (Eickelmann & Drossel, 2020). Seither sind zahlreiche digitalisierungsbezogene Entwicklungen in den Schulen zu

beobachten, für die im Rahmen von TIMSS 2023 eine Bestandsaufnahme auf repräsentativer Datenbasis vorgenommen werden kann.

Im Folgenden werden sowohl Befunde von TIMSS 2023 im internationalen Vergleich als auch vertiefend für Deutschland präsentiert. Im internationalen Vergleich werden zunächst zentrale Befunde zur Verfügbarkeit und Nutzung digitaler Medien im schulischen und häuslichen Kontext vorgestellt (Abschnitt 5.3). Ebenso werden Befunde zur Nutzungshäufigkeit und den damit verbundenen Herausforderungen für den Einsatz digitalisierungsbezogener schulischer Lernangebote im Mathematik- und naturwissenschaftlichen Sachunterricht der Grundschule (Abschnitt 5.4) sowie zu dem von den Schüler:innen selbsteingeschätzten Kompetenzerleben hinsichtlich der Nutzung digitaler Medien (Abschnitt 5.5) dargelegt. Für Deutschland wird anschließend vertiefend auf die Nutzung digitaler Medien in unterschiedlichen Lernkontexten in der Grundschule eingegangen (Abschnitt 5.6). Im letzten Abschnitt werden Befunde zu der Teilnahme der Lehrkräfte an digitalisierungsbezogenen Fortbildungen im internationalen Vergleich berichtet (Abschnitt 5.7). Die zentralen Befunde dieses Kapitels werden abschließend zusammenfassend diskutiert (Abschnitt 5.8).

5.3 Verfügbarkeit schulischer und häuslicher digitaler Ressourcen im internationalen Vergleich

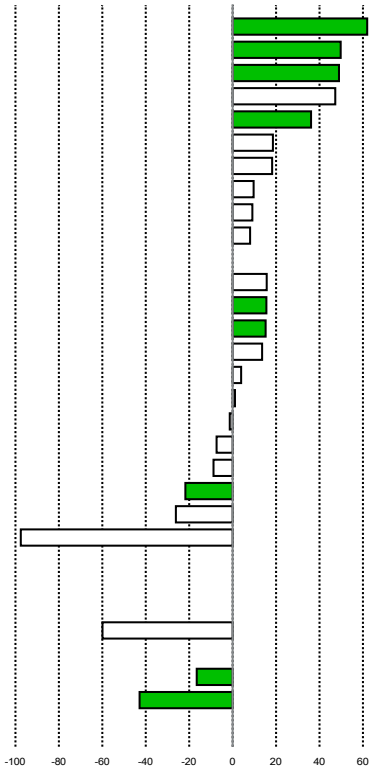
Um abbilden zu können, inwiefern die Digitalisierung an Grundschulen in Deutschland voranschreitet, wird im Rahmen von TIMSS 2023 zunächst die Ausstattung sowohl im schulischen als auch im häuslichen Kontext erhoben. Im Einzelnen wird dabei gefragt, inwiefern Schüler:innen digitale Geräte im schulischen Kontext (Anzahl der verfügbaren Computer, inklusive Tablets für den Unterricht pro Viertklässler:in) und verschiedene digitale Ressourcen (Computer oder Laptop; Tablet oder E-Reader; zuverlässiger Internetzugang) im häuslichen Kontext zur Verfügung stehen.

5.3.1 Schulische Ausstattung mit digitalen Medien im Trend

Im Rahmen von TIMSS 2023 wurden die Schulleitungen dazu befragt, wie viele digitale Geräte (Computer inklusive Tablets) in den Schulen zur Verfügung stehen, die von den Viertklässler:innen zu Unterrichtszwecken genutzt werden können. Werden diese Angaben in Relation zu den aktuellen Schülerzahlen für die vierte Jahrgangsstufe in den Schulen gestellt, ergibt sich ein Prozentwert, der die Verfügbarkeit von Computern (inklusive Tablets) für die Viertklässler:innen beschreibt. Ein Wert von 100 Prozent gibt an, dass jedes Kind einen Computer oder ein Tablet nutzen kann und entspräche einem Verhältnis von 1:1. Werte unter 100 Prozent zeigen, dass es in den Schulen weniger Computer oder Tablets als Schüler:innen gibt. Werte über 100 Prozent verweisen darauf, dass es mehr digitale Geräte als Viertklässler:innen an den Schulen gibt. Abbildung 5.1 zeigt die Ergebnisse zum mittleren Verhältnis von Viertklässler:innen pro Computer (inklusive Tablets) in Deutschland im internationalen Vergleich. Da die hier ausgewerteten Fragen auch bereits in TIMSS 2007 und 2019 eingesetzt wurden, ist es möglich, einen Vergleich der Ausstattung mit digitalen Medien im Trend und damit insbesondere auch vor und nach der COVID-19-Pandemie vorzunehmen.

Abbildung 5.1: Mittleres Verhältnis von Computern (inklusive Tablets) zu Viertklässler:innen an Grundschulen in Deutschland im internationalen Vergleich und im Trend (Angaben aus dem Schulleitungsfragebogen in Prozent, gewichtet auf die Schülerpopulation)

Fußnoten 2007	2019	2023	T	Teilnehmer	2007 %	2007 (SE)	2019 %	2019 (SE)	2023 %	2023 (SE)	Differenz A $\%_{23} - \%_{07}$	Differenz (SE)	$\%_{23} - \%_{19}$	(SE)
				Deutschland	28.5	(1.6)	54.0	(7.1)	116	(17.1)	87.4	(17.2)	62.0	(18.5)
		2		Tschechien	67.1	(3.9)	115	(8.8)	165	(8.5)	98.0	(9.4)	49.7	(12.2)
		2		Ungarn	58.6	(5.1)	81.4	(6.5)	130	(17.0)	71.8	(17.7)	49.0	(18.1)
2 3	2 3	2 3		USA	96.7	(15.2)	127	(14.9)	174	(49.2)	77.3	(51.4)	47.3	(51.3)
2	2	2	8	Litauen	53.4	(8.6)	107	(7.8)	143	(9.7)	90.0	(13.0)	36.1	(12.4)
2 3	3	3	8	Hongkong	123	(66.1)	112	(8.9)	131	(10.4)	7.3	(66.9)	18.5	(13.7)
				VG EU	**62.2**	**(5.6)**	**98.6**	**(11.1)**	**117**	**(12.0)**	**54.5**	**(13.3)**	**18.2**	**(16.3)**
				VG OECD	**72.2**	**(9.5)**	**112**	**(19.7)**	**121**	**(17.2)**	**49.2**	**(19.7)**	**9.7**	**(26.1)**
				Internationaler Mittelwert	**66.6**	**(17.3)**	**104**	**(17.6)**	**113**	**(15.2)**	**46.5**	**(23.1)**	**9.1**	**(23.2)**
3	3	3	8	Niederlande	66.7	(6.7)	106	(18.7)	114	(8.1)	47.6	(10.5)	8.0	(20.4)
				Australien	120	(22.2)	117	(8.1)	132	(7.8)	12.0	(23.6)	15.7	(11.2)
2		2	8	Italien	24.9	(1.8)	42.4	(2.9)	58.0	(3.8)	33.0	(4.2)	15.5	(4.8)
	2	2	8	Singapur	50.6	(2.4)	110	(4.5)	125	(4.2)	74.6	(4.8)	15.2	(6.1)
	3		8	Japan	71.1	(5.1)	97.5	(10.2)	111	(6.9)	40.0	(8.5)	13.6	(12.2)
				Taiwan	60.1	(7.1)	103	(16.3)	107	(10.0)	46.6	(12.2)	3.9	(19.1)
		4	8	Iran	2.0	(0.6)	5.7	(1.3)	6.8	(1.0)	4.8	(1.2)	1.1	(1.6)
2	2	2		Georgien	18.2	(4.0)	81.3	(7.1)	80.0	(4.5)	61.8	(6.1)	-1.3	(8.3)
1 3	1 2	1 2 3	8	England	90.2	(13.9)	148	(12.4)	140	(17.5)	50.0	(22.3)	-7.4	(21.4)
	2			Slowakei	61.7	(4.2)	124	(10.4)	115	(18.9)	53.2	(19.4)	-8.8	(21.5)
3	3	2 3	8	Dänemark	152	(9.1)	107	(8.7)	85.4	(3.9)	-66	(9.9)	-22	(9.5)
	2	2		Schweden	47.3	(4.2)	150	(18.2)	123	(10.1)	76.0	(11.0)	-26	(20.8)
1 2	1 2	1 2 3	8	Neuseeland	71.9	(6.8)	188	(61.5)	90.1	(4.5)	18.2	(8.1)	-98	(61.6)
				Benchmark-Teilnehmer										
2		2	8	Ontario, Kanada	98.5	(6.2)	180	(80.2)	120	(3.5)	21.9	(7.2)	-60	(80.2)
2 3 6	2		8	Dubai, VAE	89.6	(21.3)	97.7	(5.0)	81.1	(3.1)	-8.5	(21.5)	-17	(5.8)
2	3	2 3	8	Québec, Kanada	73.0	(9.8)	157	(15.4)	114	(5.0)	40.7	(11.0)	-43	(16.1)



Keine statistisch signifikante Veränderung zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Veränderung zwischen 2019 und 2023 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten in Mathematik
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

Die in Abbildung 5.1 dargestellten Ergebnisse zeigen zunächst, dass die Verfügbarkeit von digitalen Geräten für die Viertklässler:innen in 2023 in Deutschland (116 %) als sehr gut bezeichnet werden kann. Dieses Verhältnis unterscheidet sich nicht statistisch signifikant vom internationalen Mittelwert (113 %) sowie den Vergleichsgruppen EU (117 %) und OECD (121 %). Grundschulleitungen in Deutschland geben somit für 2023 an, dass es mehr digitale Geräte (Computer inklusive Tablets) als Viertklässler:innen an den Schulen gibt. Dass diese Werte zum Teil auch über 100 Prozent liegen, ist dadurch zu erklären, dass die erfragte Anzahl von Computern und Tablets möglicherweise nicht exklusiv für Viertklässler:innen betrachtet wurde, sondern auch solche Geräte einbezogen wurden, die den Schüler:innen der anderen Jahrgänge an der jeweiligen Schule zur Verfügung stehen. Mit Blick auf den Vergleich von 2019 und 2023 ist für Deutschland im internationalen Vergleich der deutlichste Anstieg des Ausstattungsverhältnisses mit digitalen Geräten (Computer inklusive Tablets) zu verzeichnen (62 %). Diese damit in 2023 deutlich verbesserte Ausstattungssituation mit digitalen Medien setzt die bereits seit 2007 zu beobachtende Erweiterung der IT-Ausstattung fort und sollte in Deutschland im Zusammenhang mit den länderspezifischen Förderprogrammen hinsichtlich der Verwaltungsvereinbarung zum DigitalPakt Schule, die im Kontext der COVID-19-Pandemie nochmals verstärkt wurde, gesehen werden, die als Flächenprogramm zu einer erheblichen Ausweitung der schulischen IT-Ausstattung geführt haben.

5.3.2 Verfügbarkeit digitaler Ressourcen im häuslichen Kontext

Neben der Ausstattung mit digitalen Medien im schulischen Kontext wird in TIMSS 2023 auch die Ausstattung mit digitalen Medien im häuslichen Kontext der Viertklässler:innen betrachtet. Dazu wurden die Erziehungsberechtigten gefragt, ob verschiedene digitale Medien oder Ressourcen im Haushalt vorhanden sind. Tabelle 5.1 zeigt die Verfügbarkeit digitaler Ressourcen im häuslichen Kontext im internationalen Vergleich. Die Tabelle ist absteigend nach der prozentualen Angabe der Verfügbarkeit für *Computer oder Laptop* sortiert.

Im Vergleich mit den internationalen Mittelwerten liegt in Deutschland die Verfügbarkeit digitaler Ressourcen statistisch signifikant höher: So geben 94.2 Prozent der Erziehungsberechtigten in Deutschland an, dass *Computer oder Laptops* zur Verfügung stehen (international: 82.8 %). Hinsichtlich der Verfügbarkeit von *Tablets oder E-Readern* bejahen 86.8 Prozent (international: 68.2 %) der Erziehungsberechtigten die Verfügbarkeit. Einen *zuverlässigen Internetzugang* geben 97.5 Prozent der Erziehungsberechtigten (international: 90.5 %) an. Das generelle Ausstattungsniveau in den Vergleichsgruppen der EU und der OECD liegt für alle drei Kategorien über den internationalen Vergleichswerten, aber immer noch statistisch signifikant unterhalb der für Deutschland ermittelten Ausstattungsquoten.

Insgesamt kann konstatiert werden, dass die durchschnittliche Verfügbarkeit digitaler Ressourcen im häuslichen Kontext in Deutschland vergleichsweise hoch ist. Auf Basis der vorliegenden Datengrundlage kann jedoch im internationalen Vergleich nicht geklärt werden, ob die digitalen Ressourcen auch durch die Viertklässler:innen (zum Lernen) genutzt werden können und ob sich beispielsweise Unterschiede ergäben, wenn die Schüler:innen differenziert nach der sozialen Herkunft betrachtet würden.

Tabelle 5.1: Häusliche Ausstattung mit digitalen Ressourcen in Deutschland im internationalen Vergleich (Angaben der Erziehungsberechtigten)

	Computer oder Laptop		Tablet oder E-Reader		zuverlässiger Internetzugang	
Teilnehmer	%	(*SE*)	%	(*SE*)	%	(*SE*)
2 *Slowenien*	98.2	(0.2)	67.1	(0.9)	97.5	(0.2)
2 *Polen*	97.5	(0.3)	65.0	(0.8)	98.9	(0.2)
2 *Tschechien*	97.3	(0.2)	73.7	(0.6)	97.3	(0.2)
2 *Litauen*	96.6	(0.3)	76.0	(0.7)	98.0	(0.1)
2 *Singapur*	96.6	(0.2)	79.7	(0.4)	97.9	(0.1)
2 3 *Dänemark*	96.4	(0.2)	89.0	(0.6)	96.6	(0.3)
2 3 *Belgien (Fläm. Gem.)*	96.3	(0.4)	89.0	(0.6)	97.1	(0.2)
2 3 *Kanada*	95.7	(0.4)	89.7	(0.4)	99.0	(0.1)
2 *Schweden*	95.5	(0.5)	87.8	(0.8)	97.8	(0.3)
Finnland	95.3	(0.4)	79.6	(0.6)	95.9	(0.2)
Lettland	95.0	(0.6)	67.7	(1.0)	97.2	(0.4)
Deutschland	94.2	(0.3)	86.8	(0.6)	97.5	(0.3)
Republik Korea (Südkorea)	93.9	(0.5)	77.4	(1.8)	91.8	(0.4)
Macau	93.7	(0.2)	88.6	(0.3)	96.0	(0.2)
3 *Hongkong*	93.4	(0.6)	94.0	(0.5)	96.6	(0.3)
Portugal	93.4	(0.6)	76.8	(0.8)	97.8	(0.2)
1 2 *Norwegen (5. Jgst.)*	93.3	(0.7)	93.4	(0.6)	98.5	(0.3)
Vereinigte Arabische Emirate (VAE)	92.9	(0.2)	82.3	(0.4)	91.0	(0.2)
2 *Georgien*	92.3	(0.6)	39.5	(1.2)	86.0	(0.6)
VG EU	**92.1**	**(0.8)**	**76.3**	**(0.9)**	**95.8**	**(0.5)**
Irland	91.5	(0.6)	88.5	(0.7)	95.3	(0.4)
VG OECD	**91.4**	**(0.7)**	**76.1**	**(0.9)**	**95.7**	**(0.4)**
Katar	90.9	(0.6)	71.8	(0.9)	94.6	(0.4)
2 *Frankreich*	90.5	(0.6)	71.4	(1.0)	95.7	(0.4)
2 *Belgien (Frz. Gem.)*	90.4	(0.6)	74.0	(0.9)	98.5	(0.2)
Slowakei	89.6	(1.1)	64.7	(1.2)	93.0	(0.8)
2 *Ungarn*	89.4	(1.3)	65.5	(1.7)	95.9	(0.6)
2 *Spanien*	88.2	(0.7)	81.9	(0.7)	94.9	(0.2)
2 *Italien*	88.1	(0.6)	79.2	(0.8)	90.1	(0.6)
Kuwait	88.1	(0.7)	74.6	(1.0)	89.4	(0.6)
Taiwan	88.0	(0.4)	78.3	(0.5)	91.2	(0.3)
Bahrain	87.4	(0.5)	65.8	(0.6)	92.7	(0.3)
2 *Serbien*	87.2	(0.9)	58.5	(1.1)	92.3	(0.8)
2 *Zypern*	86.0	(0.6)	85.0	(0.6)	95.1	(0.3)
2 *Bosnien und Herzegowina*	84.0	(1.0)	55.0	(1.1)	88.8	(1.1)
Japan	83.9	(0.9)	78.1	(0.8)	97.7	(0.3)
2 3 *Rumänien*	83.1	(1.7)	67.8	(1.5)	93.4	(1.0)
Internationaler Mittelwert	**82.8**	**(0.9)**	**68.2**	**(1.0)**	**90.5**	**(0.7)**
Bulgarien	82.0	(2.0)	65.1	(1.5)	88.4	(1.4)
Nordmazedonien	77.8	(1.0)	62.4	(0.9)	87.7	(0.8)
2 *Montenegro*	77.2	(0.6)	47.2	(0.8)	88.1	(0.4)
2 3 *Chile*	75.2	(1.3)	51.2	(1.1)	88.3	(0.6)
Oman	72.8	(0.7)	61.5	(0.8)	89.1	(0.5)
2 *Saudi-Arabien*	72.3	(0.9)	61.2	(0.9)	88.7	(0.5)
2 *Kosovo*	72.2	(1.3)	59.5	(1.6)	88.6	(0.8)
2 *Armenien*	70.8	(1.2)	47.0	(1.2)	90.9	(0.6)
2 *Kasachstan*	70.4	(1.0)	44.6	(1.3)	79.1	(1.0)
1 2 3 *Türkei (5. Jgst.)*	66.0	(1.5)	62.8	(1.1)	87.4	(1.0)
2 3 *Albanien*	61.7	(2.8)	51.3	(3.2)	85.5	(1.9)
Jordanien	50.3	(1.4)	36.9	(0.9)	81.9	(0.6)
2 3 *Brasilien*	47.2	(1.3)	36.7	(1.2)	86.9	(0.4)
Aserbaidschan	46.5	(1.3)	24.7	(0.9)	70.0	(1.3)
Iran	42.9	(1.0)	35.5	(0.8)	74.2	(0.9)
1 *Südafrika (5. Jgst.)*	42.8	(1.2)	37.9	(0.9)	47.8	(1.1)
2 *Usbekistan*	38.9	(1.4)	29.1	(1.2)	67.4	(1.4)
Marokko	36.7	(1.3)	31.7	(1.1)	50.5	(1.4)
Benchmark-Teilnehmer						
Dubai, VAE	94.8	(0.5)	90.2	(0.5)	95.0	(0.4)
2 3 *Québec, Kanada*	94.5	(0.4)	93.7	(0.5)	99.7	(0.1)
Abu Dhabi, VAE	94.0	(0.2)	80.6	(0.7)	90.7	(0.3)
Schardscha, VAE	91.8	(0.3)	80.8	(1.4)	91.4	(0.4)

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

© TIMSS 2023

5.4 Digitalisierungsbezogene Lernangebote im Fachunterricht im internationalen Vergleich

Nach der Betrachtung der allgemeinen Verfügbarkeit digitaler Medien in der Schule und zu Hause wird im Folgenden auf den fachspezifischen Kontext fokussiert. Dabei wird aus der Perspektive der Lehrkräfte zunächst betrachtet, inwiefern digitale Medien für den Unterricht im Mathematik- und naturwissenschaftlichen Sachunterricht verfügbar sind und mit welcher Häufigkeit diese für verschiedene unterrichtliche Tätigkeiten genutzt werden. Daran anschließend werden die in TIMSS 2023 ebenfalls erhobenen Hindernisse für den Einsatz digitaler Medien im Fachunterricht im internationalen Vergleich betrachtet. Abschließend werden Ergebnisse zu der von den Viertklässler:innen selbsteingeschätzten Kompetenz im Umgang mit digitalen Medien berichtet.

5.4.1 Verfügbarkeit und Nutzungshäufigkeit digitaler Medien im Fachunterricht

Tabelle 5.2 zeigt die Ergebnisse zur Verfügbarkeit und Einsatzhäufigkeit digitaler Medien für verschiedene Tätigkeiten im Mathematikunterricht im internationalen Vergleich aus Perspektive der Lehrkräfte, gewichtet auf die von ihnen unterrichteten Viertklässler:innen. Die Tabelle ist nach dem Kriterium *Verfügbarkeit digitaler Medien* absteigend sortiert. Mit Blick auf die Verfügbarkeit digitaler Medien im Mathematikunterricht finden sich in der ersten Spalte der Tabelle 5.2 die Anteile der Viertklässler:innen, die von Lehrkräften unterrichtet werden, die angeben, dass digitale Geräte (inklusive Computer, Tablets oder Smartphones) für den Mathematikunterricht in der TIMSS-Klasse zur Verfügung stehen.

Es zeigt sich, dass fast zwei Drittel (63.6 %) der Viertklässler:innen in Deutschland von Lehrkräften unterrichtet werden, die angeben, dass digitale Geräte für ihren Mathematikunterricht verfügbar sind. Dieser Anteil für Deutschland ist signifikant höher als der internationale Mittelwert (48.2 %) und der Wert in den Vergleichsgruppe EU (54.0 %) sowie vergleichbar mit dem Anteil der Vergleichsgruppe OECD (61.1 %). Die höchsten Anteile, die signifikant über dem Wert für Deutschland liegen, weisen neben dem Spitzenreiter Japan (97.5 %) insbesondere auch die skandinavischen Staaten Norwegen (95.6 %), Dänemark (94.5 %) und Schweden (92.6 %) auf. Die weiteren Angaben in Tabelle 5.2 zur Nutzung digitaler Medien im Hinblick auf unterschiedliche Tätigkeiten im Mathematikunterricht beziehen sich jeweils auf die Anteile von Klassen, in denen entsprechend auch digitale Medien verfügbar sind. Die angegebenen Prozentwerte geben dabei die Anteile der Viertklässler:innen an, die von Mathematiklehrkräften unterrichtet werden, die angeben, digitale Medien *mindestens einmal pro Woche* für verschiedene Tätigkeiten im Unterricht zu nutzen.

Zunächst zeigt sich für Deutschland, dass etwa ein Drittel (33.8 %) der Viertklässler:innen in Klassen, in denen digitale Medien für den Mathematikunterricht verfügbar sind, ein Unterrichtsangebot erhält, in dem nach Angaben der Lehrkräfte digitale Medien zum Üben von Aufgaben *mindestens einmal pro Woche* eingesetzt werden. Dieser Anteil unterscheidet sich nicht statistisch signifikant von den entsprechenden Anteilen für die Vergleichsgruppen der EU (32.3 %) und der OECD (37.0 %) und ist signifikant kleiner als der internationale Vergleichswert (43.4 %). In zwölf Staaten liegt der Anteil der Viertklässler:innen, die *mindestens einmal pro Woche* Aufgaben mit digitalen Medien im Mathematikunter-

Tabelle 5.2: Verfügbarkeit und Einsatzhäufigkeit digitaler Medien im Mathematikunterricht für verschiedene Tätigkeiten in Deutschland im internationalen Vergleich (Kategorie *mindestens einmal pro Woche*, Angaben der Lehrkräfte gewichtet auf die Schülerpopulation)

	Verfügbarkeit digitaler Medien		Aufgaben üben		Komplexe Aufgaben lösen		Grafiken/ Tabellen u. a. erstellen		Math. Spiele spielen		Lehrbuch lesen/ Lernvideos ansehen		Tests/ Klassen-arbeiten schreiben	
Teilnehmer	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
Japan	97.5	(1.2)	29.2	(3.7)	16.6	(2.9)	3.8	(1.5)	18.0	(2.7)	27.9	(3.8)	1.7	(1.1)
1 2 *Norwegen (5. Jgst.)*	95.6	(1.5)	69.8	(3.3)	54.8	(3.4)	10.8	(2.6)	36.1	(3.6)	57.4	(3.7)	2.6	(1.0)
2 3 *Dänemark*	94.5	(1.8)	41.5	(4.0)	22.8	(3.4)	17.5	(3.2)	20.4	(3.5)	25.4	(3.0)	4.6	(2.0)
2 *Schweden*	92.6	(1.8)	42.5	(3.9)	20.2	(3.1)	1.7	(0.6)	51.6	(4.4)	37.1	(4.0)	0.4	(0.4)
1 2 3 *Neuseeland*	89.7	(2.7)	67.0	(3.8)	38.0	(4.0)	9.4	(2.0)	73.2	(3.5)	37.6	(4.0)	4.1	(1.3)
Republik Korea (Südkorea)	88.3	(2.8)	11.4	(2.1)	12.0	(2.3)	5.7	(1.8)	8.1	(2.0)	16.6	(2.9)	4.8	(1.8)
2 3 *USA*	85.6	(1.6)	83.1	(1.6)	62.2	(2.0)	22.7	(1.8)	78.9	(1.5)	47.9	(2.0)	33.8	(2.1)
2 *Litauen*	83.6	(2.2)	48.9	(3.5)	36.1	(3.1)	9.5	(2.0)	36.5	(2.7)	44.5	(3.5)	9.8	(1.7)
3 *Hongkong*	83.3	(3.1)	36.5	(5.5)	30.4	(4.1)	10.5	(2.9)	27.9	(4.6)	20.3	(4.4)	9.6	(3.0)
3 *Niederlande*	83.2	(3.6)	73.1	(4.8)	41.5	(5.0)	14.0	(3.3)	53.5	(6.2)	33.2	(4.8)	12.8	(3.1)
Australien	76.6	(3.0)	61.2	(3.7)	34.1	(3.3)	15.3	(2.8)	57.4	(3.5)	34.4	(3.7)	19.8	(3.2)
Finnland	75.9	(2.9)	27.3	(3.4)	18.0	(2.7)	0.0	(0.0)	18.0	(2.8)	23.6	(2.9)	0.0	(0.0)
Lettland	75.6	(3.0)	30.5	(4.2)	15.2	(2.6)	5.5	(1.8)	10.7	(3.1)	26.0	(3.4)	3.2	(1.8)
2 3 *Belgien (Fläm. Gem.)*	66.9	(3.5)	25.6	(4.6)	9.5	(2.9)	3.2	(1.4)	24.0	(4.1)	15.4	(3.8)	1.6	(0.8)
2 *Ungarn*	65.6	(3.7)	29.7	(4.5)	10.4	(2.6)	5.8	(2.1)	27.5	(4.1)	21.8	(3.9)	3.6	(2.5)
Vereinigte Arab. Emirate (VAE)	63.8	(1.1)	79.7	(1.0)	69.8	(1.3)	33.9	(1.6)	71.2	(1.3)	71.4	(1.0)	54.2	(1.3)
Deutschland	63.6	(2.7)	33.8	(3.5)	8.1	(1.9)	1.6	(1.1)	11.6	(2.0)	10.1	(1.9)	0.7	(0.7)
2 *Kasachstan*	63.3	(2.7)	75.4	(2.9)	67.2	(3.4)	53.8	(4.1)	64.0	(3.9)	73.3	(3.5)	64.5	(3.6)
VG OECD	**61.1**	**(3.0)**	**37.0**	**(4.2)**	**22.1**	**(3.5)**	**11.0**	**(2.6)**	**34.8**	**(4.3)**	**31.2**	**(4.4)**	**9.2**	**(2.5)**
2 *Singapur*	60.6	(2.0)	23.1	(2.3)	14.0	(2.0)	4.0	(0.9)	19.3	(1.7)	18.5	(2.4)	7.0	(1.8)
2 3 *Kanada*	60.2	(1.7)	48.8	(2.6)	31.3	(2.1)	10.4	(1.3)	57.6	(2.3)	32.1	(2.2)	8.7	(1.8)
Macau	57.1	(0.1)	36.6	(0.0)	19.8	(0.0)	6.7	(0.0)	16.9	(0.0)	36.1	(0.0)	7.4	(0.0)
1 2 3 *Türkei (5. Jgst.)*	56.4	(4.3)	70.5	(5.2)	59.2	(5.5)	64.2	(5.1)	56.1	(5.8)	66.6	(5.0)	62.7	(5.1)
Portugal	55.9	(3.4)	24.7	(4.0)	14.1	(3.0)	12.8	(2.8)	36.3	(4.1)	51.7	(3.6)	6.6	(1.9)
Taiwan	54.7	(2.7)	37.0	(3.3)	24.7	(2.9)	14.7	(2.2)	17.8	(2.1)	51.7	(4.2)	11.7	(2.6)
2 *Tschechien*	54.6	(2.6)	24.8	(2.8)	9.8	(1.4)	2.5	(0.9)	12.1	(1.8)	17.0	(2.7)	6.2	(1.7)
VG EU	**54.0**	**(3.1)**	**32.3**	**(4.6)**	**18.7**	**(3.9)**	**9.3**	**(2.9)**	**29.5**	**(4.5)**	**32.5**	**(4.7)**	**7.2**	**(2.8)**
2 *Georgien*	51.8	(3.5)	35.3	(5.0)	19.0	(4.3)	19.4	(4.2)	18.8	(3.8)	23.8	(4.2)	14.5	(3.7)
Internationaler Mittelwert	**48.2**	**(2.8)**	**43.4**	**(6.2)**	**31.7**	**(5.6)**	**19.0**	**(5.2)**	**38.9**	**(6.2)**	**43.1**	**(6.4)**	**20.3**	**(5.2)**
Bulgarien	46.3	(4.1)	38.7	(5.4)	32.5	(5.8)	18.8	(4.8)	40.6	(5.5)	74.2	(4.7)	32.5	(5.8)
Irland	45.5	(4.0)	25.5	(5.7)	12.9	(4.5)	0.0	(0.0)	46.1	(6.2)	31.8	(5.4)	2.2	(1.3)
Slowakei	44.7	(3.1)	57.5	(4.7)	19.5	(4.3)	5.1	(1.8)	34.2	(5.1)	30.1	(5.1)	11.4	(3.2)
2 3 *Rumänien*	44.6	(4.4)	51.5	(5.7)	37.3	(6.9)	27.9	(6.2)	51.5	(6.9)	70.5	(5.8)	19.1	(4.6)
2 *Spanien*	44.2	(2.3)	37.2	(2.7)	28.3	(2.7)	13.5	(2.0)	37.5	(2.9)	36.9	(2.5)	3.6	(1.1)
2 *Polen*	44.1	(3.5)	24.7	(4.9)	13.2	(3.9)	7.6	(2.7)	9.9	(3.2)	44.5	(5.2)	1.6	(1.1)
1 2 3 *England*	42.8	(4.4)	40.7	(6.4)	24.5	(6.3)	12.8	(5.2)	53.3	(7.7)	28.8	(6.0)	10.2	(3.6)
2 3 *Chile*	37.1	(3.6)	17.2	(4.2)	13.3	(3.7)	8.7	(2.8)	36.4	(6.3)	43.6	(6.9)	8.9	(3.6)
Katar	35.0	(2.2)	74.7	(4.0)	61.6	(3.8)	36.2	(4.0)	66.2	(4.0)	64.8	(2.9)	52.6	(3.7)
2 *Italien*	33.7	(3.5)	33.1	(5.2)	17.2	(4.1)	14.4	(4.0)	42.3	(5.9)	42.9	(6.1)	8.1	(4.7)
Bahrain	33.2	(2.7)	71.2	(4.9)	59.7	(5.1)	35.7	(6.5)	59.7	(5.6)	55.9	(5.7)	24.7	(4.7)
2 *Usbekistan*	33.0	(3.5)	78.2	(5.3)	60.3	(6.4)	58.1	(6.3)	74.5	(5.4)	78.6	(4.2)	73.2	(5.7)
Iran	29.6	(2.3)	45.2	(4.1)	46.2	(4.3)	35.2	(4.1)	49.4	(4.3)	68.3	(3.7)	28.3	(4.2)
Kuwait	26.8	(3.8)	47.1	(8.2)	46.4	(8.2)	20.3	(6.7)	49.1	(8.3)	62.4	(8.0)	20.8	(6.9)
2 *Serbien*	25.9	(3.2)	43.3	(7.8)	27.2	(6.6)	8.3	(4.7)	18.7	(5.8)	68.0	(7.8)	8.3	(3.8)
2 3 *Brasilien*	24.9	(2.1)	32.3	(3.9)	20.8	(3.6)	17.3	(4.0)	34.6	(4.7)	53.8	(6.1)	13.7	(2.3)
Jordanien	24.7	(2.7)	43.8	(5.6)	37.7	(5.5)	30.8	(5.4)	34.4	(5.7)	43.3	(5.7)	41.8	(5.8)
2 *Saudi-Arabien*	23.8	(2.6)	68.6	(5.3)	61.2	(5.6)	49.8	(5.7)	60.2	(5.7)	80.9	(3.6)	77.5	(4.3)
Aserbaidschan	23.8	(3.0)	75.3	(5.6)	59.6	(6.9)	57.0	(6.7)	39.6	(6.7)	61.1	(7.2)	51.0	(7.2)
2 *Slowenien*	19.9	(3.0)	13.1	(6.0)	7.3	(3.9)	5.1	(3.5)	10.1	(4.3)	27.7	(7.9)	3.1	(3.0)
2 *Zypern*	19.2	(2.5)	15.2	(6.3)	17.8	(6.6)	0.0	(0.0)	23.5	(4.9)	31.6	(7.5)	6.7	(3.8)
Nordmazedonien	18.5	(2.9)	42.6	(10.3)	26.5	(7.2)	10.3	(4.9)	22.3	(8.0)	34.2	(9.2)	8.1	(3.5)
2 *Belgien (Frz. Gem.)*	17.8	(2.2)	10.2	(5.3)	0.8	(0.8)	1.0	(1.0)	42.3	(6.6)	10.0	(4.5)	6.0	(3.1)
2 *Frankreich*	15.5	(2.6)	2.0	(2.0)	0.0	(0.0)	0.0	(0.0)	9.9	(3.9)	8.1	(4.6)	0.0	(0.0)
2 *Kosovo*	12.5	(2.7)	26.8	(10.8)	28.3	(10.0)	15.9	(9.1)	40.6	(13.6)	55.4	(11.1)	21.6	(11.4)
Oman	11.6	(1.5)	51.2	(6.7)	41.2	(6.8)	30.5	(6.5)	39.4	(6.8)	60.2	(7.9)	30.9	(7.1)
1 *Südafrika (5. Jgst.)*	10.2	(2.0)	50.1	(9.0)	51.3	(9.3)	20.8	(10.1)	52.5	(9.5)	44.2	(9.5)	25.0	(9.2)
2 3 *Albanien*	10.1	(2.9)	27.2	(12.5)	13.4	(9.2)	13.8	(9.4)	19.7	(10.8)	53.0	(15.2)	13.8	(9.4)
2 *Montenegro*	8.4	(1.4)	7.6	(1.3)	7.5	(1.3)	6.0	(1.0)	6.7	(1.1)	14.1	(2.4)	27.3	(11.0)
2 *Bosnien und Herzegowina*	6.9	(2.2)	16.5	(12.7)	13.8	(12.7)	13.6	(12.7)	43.0	(16.6)	55.3	(17.3)	13.6	(12.7)
2 Armenien	5.8	(1.6)	34.0	(15.9)	18.3	(11.6)	0.0	(0.0)	15.8	(10.9)	40.4	(13.3)	25.1	(12.6)
Marokko	3.7	(1.2)	41.5	(20.6)	39.1	(20.3)	48.3	(21.8)	41.7	(20.6)	43.2	(20.8)	30.3	(15.6)
Benchmark-Teilnehmer														
Dubai, VAE	90.7	(1.0)	77.4	(1.0)	70.0	(1.7)	31.0	(3.1)	71.3	(2.1)	69.7	(1.7)	54.7	(1.8)
2 *Ontario, Kanada*	76.4	(2.5)	54.5	(3.6)	40.2	(3.4)	12.3	(1.7)	57.9	(3.1)	35.3	(3.1)	10.2	(2.8)
Abu Dhabi, VAE	55.6	(1.2)	78.8	(2.0)	68.7	(2.5)	30.3	(1.9)	66.4	(2.9)	66.0	(2.1)	50.9	(2.9)
Schardscha, VAE	51.4	(4.5)	85.2	(3.2)	71.3	(3.0)	40.3	(3.7)	77.3	(2.8)	76.9	(3.2)	57.5	(3.3)
2 3 *Québec, Kanada*	40.5	(3.2)	26.6	(4.4)	10.9	(2.3)	6.0	(2.3)	48.2	(4.8)	29.0	(4.1)	10.0	(3.9)

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

© TIMSS 2023

richt üben, mit über zwei Dritteln deutlich und statistisch signifikant über dem Anteil von Deutschland. Hierzu zählen z. B. Norwegen, Neuseeland, die Niederlande oder die USA. Für die Häufigkeit des Einsatzes digitaler Medien für die weiteren betrachteten Tätigkeiten im Mathematikunterricht fallen die Anteile in Deutschland deutlich geringer aus. So liegt der Anteil der Schüler:innen für das Lösen komplexer Aufgaben bei 8.1 Prozent (international: 31.7 %), das Spielen von mathematischen Spielen bei 11.6 Prozent (international: 38.9 %) und das Lesen eines Lehrbuchs mittels digitaler Medien bzw. das Ansehen von Lernvideos bei 10.1 Prozent (international: 43.1 %). Diese Anteile für Deutschland liegen signifikant und deutlich unter den internationalen Vergleichswerten und auch unterhalb derer von den Vergleichsgruppen der EU und der OECD.

Eine noch größere Diskrepanz zeigt sich bei der Nutzung digitaler Medien für die Erstellung von Grafiken oder Tabellen sowie weiteren Datendarstellungen in Deutschland: Lediglich 1.6 Prozent der Viertklässler:innen in Deutschland werden von Mathematiklehrkräften unterrichtet, die angeben, *mindestens einmal pro Woche* digitale Medien für diesen Zweck im Fachunterricht einzusetzen (international: 19.0 %). Der Anteil für das Schreiben von Tests bzw. Klassenarbeiten liegt in Deutschland bei nur 0.7 Prozent (international: 20.3 %) und zeigt, dass nicht viele Lehrkräfte digitale Medien zur Leistungsüberprüfung in Grundschulen in Deutschland heranziehen. Einige Staaten, die im oberen Teil der Tabelle 5.2 dargestellt sind, weisen durchgängig deutlich und signifikant höhere Werte auf als Deutschland, beispielsweise die USA, Norwegen und die Niederlande. Der Vergleich zeigt, dass digitale Medien im Mathematikunterricht für die abgefragten Tätigkeiten dort wesentlich häufiger eingesetzt werden.

Nach der Betrachtung des Mathematikunterrichts wird im Folgenden die *Verfügbarkeit digitaler Medien* sowie die Nutzungshäufigkeit für verschiedene Tätigkeiten im naturwissenschaftlichen Sachunterricht vorgestellt. Die Verfügbarkeit digitaler Medien im Sachunterricht ist in der linken Spalte der Tabelle 5.3 in Prozent von Viertklässler:innen angegeben, die von Lehrkräften unterrichtet werden, die angeben, dass im Sachunterricht digitale Geräte (inklusive Computer, Tablets oder Smartphones) zur Verfügung stehen.

Es zeigt sich, dass über vier Fünftel (82.6 %) der Viertklässler:innen in Deutschland von Lehrkräften unterrichtet werden, die angeben, dass digitale Geräte für den Sachunterricht verfügbar sind. Dieser Anteil für Deutschland ist signifikant höher als der internationale Mittelwert (60.5 %) sowie die Werte in den Vergleichsgruppen der EU (66.6 %) sowie der OECD (73.7 %). Auch hier liegen allerdings insbesondere in den skandinavischen Staaten Norwegen, Dänemark und Schweden, aber auch in Finnland und den USA die Anteile der Verfügbarkeit digitaler Medien für den naturwissenschaftlichen Sachunterricht signifikant über dem für Deutschland.

Die weiteren Angaben in Tabelle 5.3 zur Nutzung digitaler Medien im Hinblick auf unterschiedliche Tätigkeiten im Sachunterricht beziehen sich – wie bereits in der Betrachtung für den Mathematikunterricht – jeweils nur auf die Anteile von Klassen, in denen entsprechend digitale Medien auch verfügbar sind. Die Prozentwerte spiegeln die Anteile der Viertklässler:innen wider, die von Sachunterrichtslehrkräften unterrichtet werden, die angeben, digitale Medien *mindestens einmal pro Woche* für verschiedene Tätigkeiten im Unterricht zu nutzen. Dabei zeigt sich, dass 15.0 Prozent der Viertklässler:innen, in deren Klassen digitale Medien für den Sachunterricht zur Verfügung stehen, *mindestens einmal pro Woche* mittels digitaler Medien *Lehrbücher lesen* oder *Lernvideos anschauen* (15.0 %). Dieser Anteil ist in Deutschland signifikant niedriger verglichen mit

Tabelle 5.3: Verfügbarkeit und Einsatzhäufigkeit digitaler Medien im naturwissenschaftlichen Sachunterricht für verschiedene Tätigkeiten in Deutschland im internationalen Vergleich (Kategorie *mindestens einmal pro Woche*, Angaben der Lehrkräfte gewichtet auf die Schülerpopulation)

Teilnehmer	Verfügbarkeit digitaler Medien		Virtuelle Experimente durchführen		Komplexe Aufgaben lösen		Grafiken/ Tabellen u. a. erstellen		Naturwiss. Spiele spielen		Lehrbuch lesen/ Lernvideos ansehen		Tests/ Klassenarbeiten schreiben	
	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
Japan	100.0	(0.0)	4.2	(1.4)	6.6	(2.2)	7.4	(2.3)	3.6	(1.2)	29.7	(4.0)	4.2	(1.7)
1 2 *Norwegen (5. Jgst.)*	98.1	(1.3)	0.9	(0.6)	30.9	(3.5)	3.0	(1.4)	2.0	(1.1)	38.4	(3.5)	0.3	(0.3)
2 3 *Dänemark*	96.9	(1.4)	0.7	(0.7)	12.9	(2.6)	0.3	(0.3)	0.5	(0.5)	26.6	(4.1)	0.6	(0.6)
2 *Schweden*	96.6	(1.5)	3.9	(1.5)	15.3	(3.4)	0.6	(0.5)	7.2	(2.3)	43.4	(3.8)	0.0	(0.0)
Finnland	96.5	(1.0)	0.0	(0.0)	13.6	(1.8)	1.3	(0.5)	1.7	(0.4)	38.7	(3.0)	0.0	(0.0)
Republik Korea (Südkorea)	94.6	(1.8)	10.6	(2.4)	24.7	(3.5)	8.4	(2.2)	3.8	(1.5)	31.2	(4.0)	2.4	(1.3)
2 3 *USA*	94.4	(1.1)	20.7	(1.8)	21.7	(1.6)	8.9	(0.9)	20.0	(1.9)	43.9	(2.2)	15.7	(1.6)
2 3 *Belgien (Fläm. Gem.)*	93.8	(1.9)	2.0	(0.7)	4.1	(1.4)	1.3	(0.5)	3.6	(1.0)	15.8	(2.5)	2.1	(0.9)
1 2 3 *Neuseeland*	92.9	(1.9)	2.5	(0.8)	3.8	(1.6)	4.5	(1.7)	4.1	(1.3)	8.0	(2.0)	0.7	(0.4)
3 *Hongkong*	92.1	(2.3)	8.4	(2.5)	22.7	(4.3)	7.9	(2.3)	7.5	(2.3)	37.6	(4.7)	10.6	(3.0)
Lettland	91.6	(1.4)	6.4	(1.5)	16.2	(2.9)	1.3	(0.9)	4.2	(1.5)	58.6	(4.1)	5.5	(1.8)
2 *Litauen*	88.1	(1.9)	10.3	(1.8)	16.8	(2.0)	6.4	(1.4)	17.0	(2.4)	55.5	(3.6)	9.0	(1.6)
3 *Niederlande*	87.3	(3.5)	2.4	(1.4)	9.4	(3.1)	1.6	(1.1)	4.0	(1.9)	24.6	(5.0)	3.9	(2.0)
2 *Singapur*	85.3	(1.2)	5.5	(1.0)	3.1	(0.5)	3.1	(0.6)	4.7	(1.0)	21.5	(1.7)	2.0	(0.7)
Australien	84.9	(2.5)	5.1	(1.2)	13.1	(2.1)	8.8	(2.0)	6.7	(1.3)	19.2	(2.8)	3.8	(1.0)
Deutschland	82.6	(2.1)	1.2	(0.8)	2.2	(1.1)	1.3	(0.7)	3.1	(1.4)	15.0	(2.2)	0.0	(0.0)
Macau	79.3	(0.0)	10.1	(0.0)	21.0	(0.0)	4.6	(0.0)	7.4	(0.0)	47.7	(0.1)	7.4	(0.0)
2 3 *Kanada*	77.2	(1.3)	12.2	(1.1)	14.4	(1.1)	5.4	(0.9)	9.0	(1.0)	29.7	(1.6)	3.2	(1.1)
2 *Ungarn*	76.5	(3.7)	3.9	(1.5)	9.5	(2.5)	1.6	(0.9)	24.2	(3.9)	39.6	(4.3)	1.8	(1.0)
1 2 3 *Türkei (5. Jgst.)*	75.5	(3.7)	58.6	(5.0)	46.3	(4.4)	49.5	(5.0)	46.5	(4.8)	76.5	(4.0)	57.4	(4.7)
2 *Kasachstan*	75.3	(2.8)	32.9	(3.5)	65.4	(3.2)	49.6	(3.2)	50.0	(3.3)	77.4	(3.2)	60.7	(3.6)
VG OECD	**73.7**	**(2.8)**	**7.4**	**(1.9)**	**11.5**	**(2.4)**	**5.8**	**(1.8)**	**10.2**	**(2.3)**	**33.7**	**(3.7)**	**6.7**	**(2.0)**
1 2 3 *England*	73.3	(4.3)	3.0	(1.7)	5.5	(2.6)	1.9	(1.8)	1.9	(1.2)	13.9	(3.6)	0.3	(0.3)
Vereinigte Arab. Emirate (VAE)	72.3	(0.8)	48.3	(1.3)	59.1	(1.2)	33.9	(1.3)	52.0	(1.6)	74.3	(1.3)	49.1	(1.4)
Taiwan	71.3	(2.6)	9.7	(2.2)	13.5	(2.4)	7.4	(2.1)	8.3	(2.2)	47.4	(3.5)	7.5	(2.0)
2 *Georgien*	71.1	(3.1)	6.3	(1.9)	15.6	(3.3)	10.7	(2.9)	7.3	(2.2)	34.5	(3.6)	13.8	(3.3)
2 *Spanien*	67.1	(2.2)	7.4	(1.5)	17.3	(1.7)	5.0	(1.1)	19.7	(2.0)	45.0	(2.4)	3.1	(0.9)
VG EU	**66.6**	**(3.0)**	**6.4**	**(2.3)**	**11.0**	**(2.7)**	**3.9**	**(1.6)**	**11.0**	**(2.7)**	**36.9**	**(4.1)**	**5.5**	**(2.1)**
Irland	64.9	(3.7)	0.3	(0.3)	2.1	(1.5)	0.0	(0.0)	1.7	(1.0)	8.3	(2.9)	0.0	(0.0)
Portugal	64.6	(3.6)	12.5	(2.6)	15.1	(2.2)	9.6	(1.9)	26.2	(3.6)	46.2	(4.2)	8.4	(2.0)
Bulgarien	61.3	(3.8)	14.1	(3.7)	24.1	(4.4)	11.3	(3.3)	19.6	(4.0)	71.1	(4.1)	22.8	(3.7)
Internationaler Mittelwert	**60.5**	**(2.8)**	**18.3**	**(4.2)**	**23.2**	**(4.5)**	**14.3**	**(3.8)**	**21.3**	**(4.4)**	**47.8**	**(5.2)**	**18.6**	**(4.3)**
2 *Polen*	58.9	(4.1)	5.1	(2.5)	6.3	(2.5)	1.3	(0.9)	6.3	(2.2)	43.4	(4.6)	0.6	(0.6)
2 3 *Rumänien*	56.9	(5.1)	37.7	(5.9)	29.9	(5.4)	16.5	(3.8)	44.2	(6.5)	69.7	(6.5)	13.1	(4.3)
Slowakei	54.9	(3.0)	5.0	(1.5)	16.3	(2.8)	2.6	(1.1)	16.4	(3.5)	39.1	(4.1)	7.3	(2.3)
Bahrain	54.9	(2.4)	43.6	(5.2)	37.6	(4.5)	22.1	(4.6)	57.1	(4.7)	81.0	(3.4)	38.8	(3.5)
Katar	54.5	(3.0)	39.7	(3.5)	40.1	(3.2)	28.3	(2.8)	54.9	(3.2)	73.3	(3.0)	48.0	(3.9)
Iran	53.6	(2.9)	43.3	(3.5)	32.2	(3.3)	31.7	(3.1)	46.2	(3.4)	65.5	(3.4)	41.3	(3.7)
2 *Tschechien*	53.5	(2.8)	0.9	(0.6)	4.5	(1.4)	1.2	(0.4)	3.6	(1.2)	23.3	(3.0)	2.6	(0.7)
2 *Usbekistan*	53.3	(4.1)	30.9	(5.0)	54.4	(5.5)	41.9	(5.3)	61.9	(5.4)	81.8	(4.5)	54.1	(5.6)
2 3 *Chile*	52.5	(3.7)	9.3	(3.4)	5.8	(2.9)	7.1	(3.2)	19.4	(4.2)	54.3	(5.1)	13.7	(4.6)
Kuwait	51.1	(3.9)	53.7	(5.9)	33.2	(5.2)	20.7	(4.8)	62.1	(5.8)	86.3	(4.0)	25.7	(4.9)
2 *Italien*	48.2	(3.3)	10.5	(3.2)	12.4	(3.2)	11.0	(3.2)	20.8	(3.5)	54.4	(5.0)	6.0	(2.1)
2 *Serbien*	45.9	(3.3)	10.8	(3.7)	17.8	(3.9)	4.0	(2.6)	16.3	(4.1)	58.9	(5.7)	7.6	(2.9)
2 *Slowenien*	36.8	(3.6)	3.5	(2.0)	3.4	(2.0)	1.5	(1.6)	5.7	(2.5)	50.2	(5.6)	8.2	(3.4)
Aserbaidschan	34.6	(3.2)	31.9	(5.6)	56.4	(5.2)	37.5	(4.8)	27.7	(4.6)	66.6	(6.1)	45.0	(5.6)
2 3 *Brasilien*	32.6	(2.4)	14.7	(1.9)	32.1	(2.9)	22.5	(2.1)	25.1	(2.4)	59.9	(5.1)	19.0	(1.9)
2 *Frankreich*	30.2	(3.4)	0.5	(0.0)	0.5	(0.0)	0.0	(0.0)	2.1	(1.4)	5.7	(2.8)	0.0	(0.0)
2 *Belgien (Frz. Gem.)*	30.1	(3.2)	3.0	(1.8)	1.9	(1.3)	1.0	(1.0)	2.6	(1.6)	3.8	(2.3)	1.3	(1.2)
Oman	29.1	(2.2)	46.3	(4.3)	30.3	(4.5)	35.1	(4.5)	50.0	(5.3)	83.9	(2.7)	29.1	(4.2)
2 *Saudi-Arabien*	27.8	(2.5)	57.5	(6.3)	64.7	(6.1)	41.5	(6.4)	57.7	(6.5)	87.6	(3.7)	67.9	(5.6)
2 *Zypern*	27.2	(2.7)	9.1	(4.3)	8.1	(3.8)	1.4	(1.4)	6.7	(3.6)	33.1	(6.6)	3.5	(2.5)
Jordanien	24.7	(2.5)	45.0	(6.3)	40.1	(5.4)	23.9	(4.5)	42.0	(6.4)	62.8	(5.8)	41.7	(8.1)
Nordmazedonien	22.3	(2.9)	12.6	(6.8)	28.6	(8.6)	23.9	(6.3)	17.7	(7.2)	56.6	(9.6)	10.2	(5.0)
2 *Kosovo*	17.6	(3.0)	26.7	(10.7)	27.5	(9.7)	24.3	(9.0)	30.3	(10.1)	48.8	(10.0)	19.8	(8.6)
2 *Bosnien und Herzegowina*	17.1	(4.4)	0.0	(0.0)	29.1	(12.3)	0.0	(0.0)	23.3	(11.9)	54.3	(14.0)	10.3	(7.2)
2 *Armenien*	16.8	(3.3)	4.0	(4.0)	13.2	(6.3)	4.6	(4.6)	16.9	(7.9)	36.9	(10.4)	24.7	(10.3)
2 *Montenegro*	16.5	(1.8)	4.0	(0.4)	4.7	(0.5)	4.1	(0.5)	8.4	(3.7)	26.9	(5.7)	4.3	(0.4)
2 3 *Albanien*	15.2	(3.4)	42.1	(13.9)	34.3	(12.0)	31.9	(12.1)	27.7	(11.2)	58.1	(14.8)	25.0	(12.6)
1 *Südafrika (5. Jgst.)*	14.5	(2.3)	17.1	(7.0)	21.9	(6.9)	4.2	(0.7)	7.4	(3.2)	55.4	(8.4)	19.2	(6.1)
Marokko	6.4	(1.4)	45.6	(13.4)	35.4	(13.3)	35.4	(13.3)	35.6	(13.6)	88.6	(9.0)	52.3	(13.4)
Benchmark-Teilnehmer														
Dubai, VAE	91.4	(0.6)	46.4	(2.2)	65.1	(1.6)	37.0	(2.1)	50.9	(2.6)	70.5	(1.5)	44.1	(2.2)
2 *Ontario, Kanada*	86.8	(1.8)	13.8	(2.0)	20.8	(2.1)	7.3	(1.5)	12.4	(1.8)	38.8	(2.7)	6.1	(2.4)
Schardscha, VAE	70.1	(2.7)	66.8	(2.8)	74.9	(2.9)	48.8	(4.6)	65.3	(3.0)	90.7	(1.7)	62.1	(3.1)
Abu Dhabi, VAE	65.0	(1.5)	39.9	(2.9)	44.3	(3.0)	22.7	(1.6)	46.3	(3.0)	66.5	(3.6)	48.6	(2.8)
2 3 *Québec, Kanada*	58.0	(3.6)	6.6	(2.8)	5.3	(2.0)	4.4	(2.1)	3.5	(0.4)	14.7	(1.9)	0.6	(0.0)

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

© TIMSS 2023

den entsprechenden Anteilen für die Vergleichsgruppen der EU (36.9 %), der OECD (33.7 %) und dem internationalen Vergleichswert (47.8 %). Hinsichtlich des fachspezifischen Aspekts *virtuelle Experimente durchführen* liegt der internationale Mittelwert bei 18.3 Prozent, während in Deutschland der Anteil lediglich 1.2 Prozent beträgt. Dieser Anteil ist in Deutschland signifikant niedriger als in den Vergleichsgruppen der EU (6.4 %) und der OECD (7.4 %).

Werden die beiden in TIMSS untersuchten Domänen Mathematik und Naturwissenschaften hinsichtlich der Nutzungshäufigkeit digitaler Medien für unterschiedliche Aspekte verglichen, zeigen sich für die Tätigkeit *Lehrbücher lesen* bzw. *Lernvideos ansehen* vergleichbar hohe Werte (Mathematik: 43.1 %; Sachunterricht: 47.8 %). In der Gegenüberstellung mit den internationalen Mittelwerten und denen der Vergleichsgruppen der EU und der OECD liegen die Werte für die übrigen Tätigkeiten für Deutschland mit Werten zwischen 0.0 Prozent (*Tests bzw. Klassenarbeiten schreiben*; international: 18.6 %; OECD: 6.7 %; EU: 5.5 %) und 3.1 Prozent (*naturwissenschaftliche Spiele spielen*; international: 21.3 %; OECD: 10.2 %; EU: 11.0 %) bedeutend niedriger. Die Staaten, die im oberen Teil der Tabelle 5.3 für die Naturwissenschaften dargestellt sind, weisen zwar durchgängig deutlich höhere Werte als Deutschland auf, dennoch fallen die Anteile insgesamt niedriger aus als für Mathematik. Der Vergleich zwischen den beiden Domänen Mathematik und Naturwissenschaften zeigt, dass digitale Medien, auch wenn diese zur Verfügung stehen, in Deutschland in beiden Domänen sehr selten genutzt werden.

5.4.2 Herausforderungen bei der Nutzung digitaler Medien im Fachunterricht

Um mögliche Hindernisse identifizieren zu können, die Lehrkräfte davon abhalten, digitale Medien im Fachunterricht zu nutzen, wurden drei mögliche Herausforderungen bei der Nutzung digitaler Medien im Unterricht abgefragt. Diese sind *fehlendes Wissen* darüber, inwiefern sich das Lernen mit digitalen Geräten verbessern lässt, *unzureichender Zugang* hinsichtlich digitaler Geräte sowie *mangelnde Unterstützung* durch die Schule im Hinblick auf die Nutzung von Technologien.

Tabelle 5.4 sind die Angaben der Mathematiklehrkräfte – gewichtet auf die Viertklässler:innen – zu entnehmen, bei denen die angegebenen Aspekte für die Lehrkraft *kein Hindernis* darstellen, um digitale Medien im Mathematikunterricht zu nutzen. Der Anteil der Viertklässler:innen in Deutschland, der von Mathematiklehrkräften unterrichtet wird, für die *fehlendes Wissen*, inwiefern sich das Lernen mit digitalen Geräten verbessern lässt, *kein Hindernis* für den Einsatz digitaler Medien im Mathematikunterricht darstellt, liegt bei 43.2 Prozent. Hinsichtlich des *unzureichenden Zugangs* zu digitalen Geräten liegt der Anteil bei etwa einem Drittel (33.4 %). Damit liegen die Anteile in Deutschland signifikant niedriger als im internationalen Durchschnitt (57.7 % bzw. 38.1 %) sowie in den Vergleichsgruppen der EU (61.3 % bzw. 38.0 %) und der OECD (59.0 % bzw. 42.7 %), was bedeutet, dass in Deutschland signifikant mehr Lehrkräfte *fehlendes Wissen* als Hindernis zur Nutzung digitaler Medien sehen. Weiterhin zeigt sich, dass der Anteil der Viertklässler:innen in Deutschland, der von Mathematiklehrkräften unterrichtet wird, die angeben, dass *mangelnde Unterstützung* durch die Schule zur Nutzung digitaler Technologien kein Hindernis für den Einsatz digitaler Medien im Mathematikunterricht darstellt, bei 50.9 Prozent und damit

Tabelle 5.4: Herausforderungen beim Einbezug digitaler Medien im Mathematikunterricht in Deutschland im internationalen Vergleich (Angaben der Lehrkräfte in der Kategorie *kein Hindernis*, gewichtet auf die Schülerpopulation)

	Fehlendes Wissen		Unzureichender Zugang		Mangelnde Unterstützung	
Teilnehmer	%	(*SE*)	%	(*SE*)	%	(*SE*)
2 *Polen*	79.5	(2.9)	23.3	(3.2)	47.7	(3.3)
Katar	78.4	(2.3)	45.2	(2.7)	72.8	(2.3)
Slowakei	77.3	(2.4)	34.1	(3.3)	51.4	(3.2)
Vereinigte Arabische Emirate (VAE)	77.3	(0.9)	58.0	(0.9)	74.3	(0.9)
Lettland	74.8	(3.1)	31.9	(3.6)	51.5	(3.8)
Bulgarien	74.6	(3.5)	38.8	(4.0)	72.7	(3.8)
3 *Niederlande*	74.4	(5.2)	72.2	(4.4)	74.9	(4.3)
2 *Singapur*	74.2	(1.8)	66.9	(1.8)	81.2	(1.6)
Bahrain	73.9	(3.0)	33.8	(2.6)	65.2	(2.3)
2 3 *USA*	73.6	(1.8)	85.7	(1.7)	78.3	(1.7)
Nordmazedonien	70.8	(2.4)	33.8	(3.2)	34.6	(3.0)
2 *Italien*	70.4	(2.8)	22.7	(2.9)	37.3	(3.4)
2 *Litauen*	69.6	(2.7)	31.1	(2.6)	56.3	(2.7)
2 *Bosnien und Herzegowina*	68.8	(3.7)	11.4	(2.8)	9.1	(2.6)
2 *Schweden*	67.8	(3.4)	88.9	(2.1)	79.7	(3.0)
2 *Ungarn*	67.4	(3.6)	32.6	(3.9)	46.1	(3.9)
1 2 3 *Neuseeland*	64.5	(4.2)	55.3	(3.9)	73.3	(3.3)
1 2 3 *England*	64.3	(3.8)	42.8	(4.6)	62.4	(4.3)
2 *Spanien*	64.3	(1.6)	41.2	(2.2)	45.3	(2.1)
2 *Montenegro*	63.7	(2.7)	17.4	(1.8)	27.7	(2.3)
2 3 *Dänemark*	63.5	(4.2)	78.2	(3.4)	71.4	(3.8)
Kuwait	61.8	(4.1)	22.4	(3.4)	31.0	(4.1)
VG EU	**61.3**	**(3.4)**	**38.0**	**(3.2)**	**50.6**	**(3.4)**
2 3 *Rumänien*	61.2	(4.7)	26.6	(4.2)	38.8	(4.6)
2 *Slowenien*	60.7	(3.0)	36.4	(3.3)	47.1	(3.6)
2 *Tschechien*	60.6	(2.5)	37.6	(2.5)	64.6	(2.8)
2 3 *Belgien (Fläm. Gem.)*	60.4	(4.0)	48.5	(4.0)	50.0	(4.1)
1 *Südafrika (5. Jgst.)*	60.3	(2.7)	27.4	(2.5)	35.7	(3.1)
Marokko	60.2	(2.9)	33.9	(2.7)	22.5	(2.6)
Iran	60.2	(2.8)	18.0	(2.2)	35.8	(2.6)
VG OECD	**59.0**	**(3.3)**	**42.7**	**(3.2)**	**53.3**	**(3.4)**
Internationaler Mittelwert	**57.7**	**(3.2)**	**38.1**	**(3.0)**	**47.8**	**(3.2)**
2 3 *Chile*	57.3	(3.8)	24.9	(3.3)	29.9	(3.3)
Finnland	57.3	(3.2)	58.4	(3.5)	70.7	(3.1)
2 *Saudi-Arabien*	56.0	(3.0)	11.3	(1.9)	30.3	(2.8)
1 2 *Norwegen (5. Jgst.)*	55.9	(3.5)	76.2	(3.0)	64.2	(3.9)
Macau	55.7	(0.0)	34.3	(0.0)	30.5	(0.0)
2 3 *Kanada*	55.6	(1.6)	42.3	(1.9)	48.8	(1.7)
2 *Serbien*	54.7	(3.3)	27.1	(3.3)	39.9	(3.6)
Portugal	54.6	(2.7)	39.5	(3.1)	29.6	(3.0)
Australien	53.2	(3.2)	63.3	(3.1)	64.9	(3.3)
3 *Hongkong*	51.8	(5.3)	62.3	(4.5)	64.7	(4.2)
2 *Georgien*	50.8	(3.3)	26.3	(2.5)	37.1	(3.3)
Republik Korea (Südkorea)	49.2	(4.2)	68.1	(3.3)	62.7	(3.7)
Aserbaidschan	46.8	(3.5)	22.2	(2.8)	34.9	(3.1)
2 *Frankreich*	45.8	(3.8)	13.5	(2.4)	33.5	(3.1)
2 *Kasachstan*	45.4	(3.1)	34.6	(2.9)	37.8	(3.0)
2 *Zypern*	44.7	(3.1)	12.1	(1.6)	20.4	(2.4)
Deutschland	43.2	(2.9)	33.4	(2.5)	50.9	(3.0)
Taiwan	42.8	(3.1)	50.9	(3.1)	58.2	(3.1)
2 *Usbekistan*	42.5	(3.9)	29.6	(3.2)	37.5	(3.3)
Irland	42.3	(3.9)	19.7	(3.4)	48.0	(4.0)
2 *Kosovo*	41.0	(3.6)	33.3	(3.7)	25.1	(3.6)
2 3 *Brasilien*	39.6	(2.2)	23.1	(1.6)	26.5	(1.6)
Jordanien	37.8	(3.0)	13.1	(2.2)	25.4	(3.2)
1 2 3 *Türkei (5. Jgst.)*	37.3	(3.7)	21.9	(2.6)	43.4	(3.8)
Oman	34.8	(2.2)	15.5	(1.4)	25.7	(1.9)
2 *Belgien (Frz. Gem.)*	33.6	(2.7)	16.0	(2.6)	25.4	(2.9)
Japan	28.1	(3.7)	33.2	(3.9)	47.6	(4.0)
2 *Armenien*	27.7	(3.5)	10.1	(2.4)	24.8	(3.9)
2 3 *Albanien*	25.1	(5.0)	21.9	(4.1)	22.6	(4.8)
Benchmark-Teilnehmer						
Schardscha, VAE	81.9	(1.8)	55.9	(2.9)	73.6	(2.5)
Dubai, VAE	78.5	(1.6)	77.3	(1.6)	80.2	(1.3)
Abu Dhabi, VA*E*	76.7	(1.7)	48.9	(1.8)	70.8	(1.8)
2 *Ontario, Kanada*	56.4	(2.4)	46.2	(3.0)	47.8	(2.9)
2 3 *Québec, Kanada*	53.9	(2.7)	34.8	(3.1)	39.2	(3.4)

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

© TIMSS 2023

Tabelle 5.5: Herausforderungen beim Einbezug digitaler Medien im naturwissenschaftlichen Sachunterricht in Deutschland im internationalen Vergleich (Angaben der Lehrkräfte in der Kategorie *kein Hindernis*, gewichtet auf die Viertklässler:innen)

	Fehlendes Wissen		Unzureichender Zugang		Mangelnde Unterstützung	
Teilnehmer	%	(*SE*)	%	(*SE*)	%	(*SE*)
Lettland	83.6	(2.5)	42.6	(3.7)	55.0	(3.8)
Slowakei	82.1	(2.7)	33.6	(3.6)	54.5	(3.4)
Bulgarien	79.3	(3.1)	41.3	(4.1)	73.1	(3.8)
Bahrain	77.6	(2.1)	37.9	(2.8)	66.1	(2.9)
Vereinigte Arabische Emirate (VAE)	76.6	(0.8)	56.9	(0.9)	73.3	(1.0)
Nordmazedonien	76.5	(2.4)	38.0	(3.5)	36.6	(3.2)
2 *Singapur*	74.3	(1.8)	66.9	(1.7)	79.2	(1.7)
2 *Italien*	70.2	(2.7)	23.5	(3.0)	36.5	(3.4)
Katar	69.6	(2.6)	35.4	(2.5)	73.4	(2.4)
2 *Polen*	69.4	(3.1)	21.0	(3.3)	39.6	(3.9)
2 *Litauen*	68.9	(2.8)	35.2	(3.0)	55.8	(2.9)
2 *Tschechien*	68.5	(2.4)	42.7	(2.7)	67.2	(2.1)
2 3 *USA*	67.9	(2.0)	86.9	(1.5)	78.8	(1.7)
2 *Montenegro*	67.7	(2.6)	13.2	(2.0)	26.4	(2.1)
Kuwait	66.8	(3.8)	27.4	(3.5)	48.8	(4.0)
3 *Niederlande*	66.6	(5.1)	71.0	(4.8)	73.3	(3.9)
Marokko	64.1	(2.8)	36.7	(3.0)	23.2	(2.6)
2 *Ungarn*	63.5	(3.7)	36.1	(3.6)	47.6	(4.1)
2 *Spanien*	62.3	(1.8)	37.7	(1.9)	41.8	(2.1)
2 *Bosnien und Herzegowina*	62.2	(5.1)	11.9	(2.6)	9.3	(2.5)
2 *Serbien*	60.7	(3.4)	30.2	(3.4)	38.9	(3.9)
VG EU	**60.2**	**(3.3)**	**39.7**	**(3.3)**	**50.4**	**(3.5)**
2 3 *Dänemark*	60.1	(3.8)	70.7	(4.0)	74.0	(3.9)
2 *Saudi-Arabien*	59.8	(3.3)	10.6	(1.5)	34.4	(2.9)
2 *Schweden*	58.9	(3.9)	89.7	(1.9)	77.4	(3.4)
Oman	58.7	(2.5)	17.8	(1.3)	31.3	(2.2)
Iran	58.4	(2.6)	13.0	(2.2)	32.3	(2.8)
2 *Slowenien*	58.1	(3.0)	35.7	(3.3)	47.2	(3.6)
2 3 *Rumänien*	57.6	(4.5)	27.5	(4.6)	40.0	(5.3)
Internationaler Mittelwert	**57.2**	**(3.3)**	**38.8**	**(3.1)**	**48.0**	**(3.3)**
Portugal	56.7	(3.1)	41.1	(3.4)	28.8	(2.9)
1 *Südafrika (5. Jgst.)*	56.4	(2.8)	22.7	(2.0)	31.6	(3.0)
VG OECD	**56.2**	**(3.4)**	**44.3**	**(3.3)**	**52.7**	**(3.5)**
2 3 *Belgien (Fläm. Gem.)*	56.1	(4.1)	54.0	(3.9)	50.6	(4.3)
2 3 *Kanada*	55.4	(1.8)	47.0	(1.8)	50.8	(1.7)
Macau	54.6	(0.1)	39.2	(0.1)	26.0	(0.0)
2 *Zypern*	54.6	(3.0)	14.7	(2.2)	25.4	(2.8)
2 *Kasachstan*	54.6	(3.0)	33.8	(3.1)	39.3	(3.4)
2 3 *Chile*	54.4	(3.8)	20.9	(3.3)	29.7	(4.0)
Finnland	53.6	(3.3)	61.1	(3.3)	67.7	(3.1)
Australien	53.1	(3.2)	61.3	(3.5)	65.6	(3.3)
2 3 *Brasilien*	52.5	(2.1)	27.9	(1.6)	25.3	(2.0)
Taiwan	52.3	(3.1)	54.8	(3.2)	70.0	(2.9)
3 *Hongkong*	49.1	(4.8)	59.6	(5.5)	64.2	(4.4)
2 *Kosovo*	48.9	(4.5)	35.4	(4.2)	26.1	(3.8)
1 2 *Norwegen (5. Jgst.)*	47.9	(4.4)	72.2	(3.3)	65.4	(3.6)
1 2 3 *England*	46.4	(4.9)	40.6	(5.1)	58.8	(4.7)
1 2 3 *Neuseeland*	46.3	(4.2)	56.4	(3.7)	70.4	(3.7)
2 *Georgien*	45.6	(3.5)	23.4	(2.5)	39.4	(3.5)
Republik Korea (Südkorea)	45.0	(4.0)	65.1	(3.4)	62.7	(4.0)
Japan	43.8	(4.4)	51.5	(4.3)	50.7	(4.2)
Deutschland	43.3	(2.5)	41.4	(2.8)	50.6	(3.0)
Aserbaidschan	42.3	(3.5)	19.9	(2.9)	31.9	(3.1)
2 *Frankreich*	42.0	(3.9)	11.5	(2.3)	32.1	(3.2)
2 *Usbekistan*	40.6	(3.7)	23.0	(3.5)	32.7	(3.5)
Jordanien	34.7	(3.3)	15.6	(1.8)	29.8	(3.4)
2 *Belgien (Frz. Gem.)*	34.6	(3.0)	16.9	(2.1)	24.3	(2.4)
Irland	33.4	(2.9)	23.9	(3.5)	47.4	(4.1)
1 2 3 *Türkei (5. Jgst.)*	28.6	(3.9)	22.4	(3.4)	35.8	(3.7)
2 *Armenien*	25.1	(3.9)	11.6	(2.7)	16.4	(3.0)
2 3 *Albanien*	17.3	(4.4)	13.8	(4.1)	24.6	(5.1)
Benchmark-Teilnehmer						
Schardscha, VAE	80.6	(2.3)	60.5	(3.3)	70.1	(3.0)
Dubai, VAE	80.2	(1.4)	78.4	(1.2)	85.7	(1.3)
Abu Dhabi, VAE	76.5	(1.3)	46.0	(1.6)	68.4	(1.7)
2 *Ontario, Kanada*	54.9	(3.2)	48.2	(3.1)	46.7	(3.1)
2 3 *Québec, Kanada*	54.7	(3.1)	40.3	(3.2)	47.0	(3.3)

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

 © TIMSS 2023

im Bereich des internationalen Mittelwertes (47.8 %) liegt. Auch die Sachunterrichtslehrkräfte wurden nach Hindernissen des Einsatzes digitaler Medien in ihrem Fachunterricht gefragt.

Tabelle 5.5 zeigt die Ergebnisse und fokussiert erneut die Antwortkategorie *kein Hindernis*. Der Anteil der Viertklässler:innen in Deutschland, der von Sachunterrichtslehrkräften unterrichtet wird, die angeben, dass *fehlendes Wissen*, wie sich das Lernen mit digitalen Geräten verbessern lässt, *kein Hindernis* für den Einsatz digitaler Medien im Sachunterricht darstellt, liegt bei 43.3 Prozent. Damit liegen die Anteile in Deutschland signifikant niedriger als im internationalen Durchschnitt (57.2 %), in der Vergleichsgruppe OECD (56.2 %) und in der EU (60.2 %). Richtet man den Blick auf die Aspekte *unzureichender Zugang* zu digitalen Geräten (Deutschland: 41.4 %) und *mangelnde Unterstützung* durch die Schule (Deutschland: 50.6 %) zeigen sich keine statistisch auffälligen Unterschiede im internationalen Vergleich oder in Bezug auf die Vergleichsgruppe der OECD bzw. der EU.

Im internationalen Vergleich zeigt sich weiterhin, dass es Staaten gibt, in denen der Anteil der Viertklässler:innen, die von Lehrkräften unterrichtet werden, die angeben, dass sie in Bezug auf die drei betrachteten Aspekte keine Hindernisse für den Einsatz digitaler Medien im Sachunterricht sehen, signifikant höher ausfällt als in Deutschland. Dazu zählen beispielsweise die Nachbarstaaten von Deutschland Niederlande und Dänemark.

5.5 Selbsteingeschätzte Kompetenzen der Schüler:innen im Umgang mit digitalen Medien im internationalen Vergleich

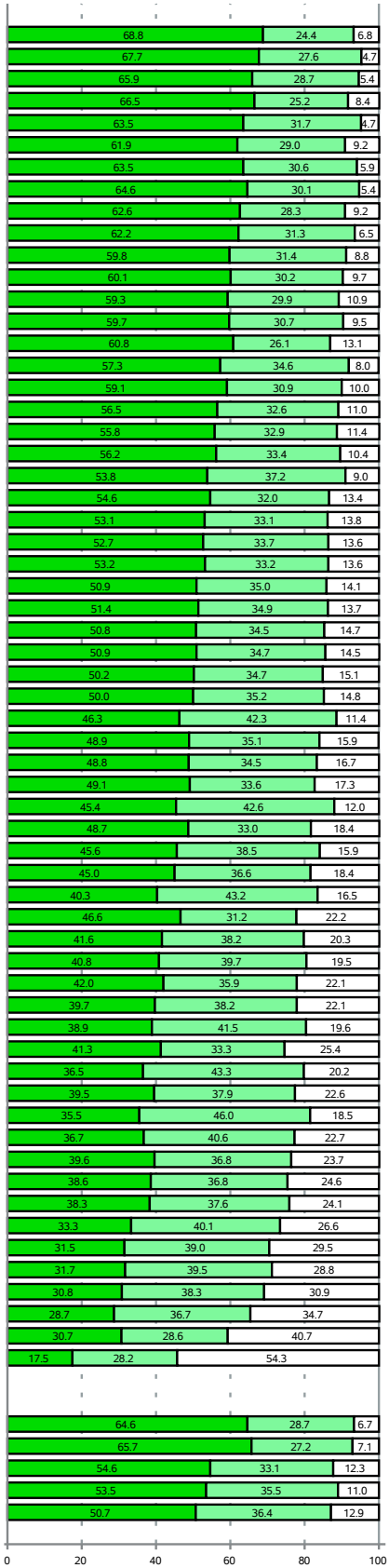
Im Rahmen von TIMSS 2023 wurde erstmalig danach gefragt, wie Viertklässler:innen ihre eigenen Kompetenzen im Umgang mit digitalen Medien einschätzen. Hierzu wurden den Viertklässler:innen drei Fähigkeiten – einen Text schreiben und bearbeiten können, Vorträge für die Schule vorbereiten können sowie Tabellen, Diagramme und Grafiken erstellen können – genannt. Mit einem vierstufigen Antwortformat *stark*, *einigermaßen*, *wenig* oder *überhaupt nicht* sollten die Viertklässler:innen einschätzen, inwiefern sie sich in der Lage fühlen, diese Tätigkeiten am Computer, Tablet oder Smartphone auszuführen.

Für die Darstellung in Abbildung 5.2 wurde aus den drei Einzelitems eine Mittelwertskala gebildet, die im Sinne einer selbst eingeschätzten Zustimmung bezogen auf verschiedene Fertigkeiten mit digitalen Medien einen Wertebereich von 1 (entspricht der Antwortkategorie *überhaupt nicht*) bis 4 (entspricht der Antwortkategorie *stark*) mit einem theoretischen Mittelwert von 2.5 aufweist. Anschließend wurde diese Skala auf drei Stufen des Kompetenzerlebens rekodiert (selbstberichtetes Kompetenzerleben: *gering:* 1 bis 2; *mittel*: > 2 bis 3; *stark:* > 3 bis 4).

Mit einem Mittelwert von 2.86 liegen die selbsteingeschätzten Kompetenzen der Viertklässler:innen in Deutschland im Umgang mit digitalen Medien leicht über dem theoretischen Skalenmittelwert. Der Wert fällt jedoch statistisch signifikant niedriger aus als der internationale Mittelwert (M = 3.06), der Mittelwert der Vergleichsgruppe EU (M = 3.07) und der OECD (M = 3.10). Damit schätzen Viertklässler:innen in Deutschland ihre Kompetenzen signifikant niedriger ein als im OECD- und EU-Vergleich. Im internationalen Vergleich zeigt sich, dass Viertklässler:innen beispielsweise in Finnland, Norwegen, Schweden und Dänemark ihre eigenen Kompetenzen im Umgang mit digitalen Medien

Abbildung 5.2: Mittlere selbsteingeschätzte Kompetenzen im Umgang mit digitalen Medien in Deutschland im internationalen Vergleich (Angaben der Viertklässler:innen)

Teilnehmer [A]	M	(SE)	hoch %	(SE)	mittel %	(SE)	niedrig %	(SE)
Taiwan	3.43	(0.01)	68.8	(0.6)	24.4	(0.6)	6.8	(0.4)
Finnland	3.37	(0.01)	67.7	(0.8)	27.6	(0.7)	4.7	(0.4)
12 *Norwegen (5. Jgst.)*	3.35	(0.01)	65.9	(1.0)	28.7	(1.0)	5.4	(0.5)
123 *Türkei (5. Jgst.)*	3.34	(0.02)	66.5	(1.1)	25.2	(1.0)	8.4	(0.5)
2 *Schweden*	3.34	(0.02)	63.5	(1.4)	31.7	(1.1)	4.7	(0.5)
3 *Hongkong*	3.33	(0.02)	61.9	(1.1)	29.0	(0.9)	9.2	(0.5)
3 *Niederlande*	3.33	(0.02)	63.5	(1.1)	30.6	(0.9)	5.9	(0.6)
23 *Dänemark*	3.32	(0.02)	64.6	(1.1)	30.1	(0.9)	5.4	(0.4)
Australien	3.29	(0.02)	62.6	(0.9)	28.3	(0.7)	9.2	(0.5)
Lettland	3.29	(0.01)	62.2	(1.0)	31.3	(0.9)	6.5	(0.6)
23 *Kanada*	3.27	(0.01)	59.8	(0.6)	31.4	(0.5)	8.8	(0.3)
123 *England*	3.26	(0.02)	60.1	(1.2)	30.2	(1.0)	9.7	(0.6)
2 *Singapur*	3.25	(0.01)	59.3	(0.7)	29.9	(0.5)	10.9	(0.4)
2 *Litauen*	3.24	(0.01)	59.7	(1.0)	30.7	(0.9)	9.5	(0.5)
23 *Rumänien*	3.24	(0.03)	60.8	(1.7)	26.1	(1.3)	13.1	(1.0)
2 *Polen*	3.22	(0.01)	57.3	(0.9)	34.6	(0.9)	8.0	(0.5)
Portugal	3.22	(0.02)	59.1	(1.1)	30.9	(0.9)	10.0	(0.6)
Vereinigte Arabische Emirate (VAE)	3.20	(0.01)	56.5	(0.4)	32.6	(0.3)	11.0	(0.3)
Bahrain	3.19	(0.02)	55.8	(0.9)	32.9	(0.9)	11.4	(0.6)
2 *Georgien*	3.19	(0.02)	56.2	(1.1)	33.4	(0.9)	10.4	(0.7)
23 *Belgien (Fläm. Gem.)*	3.17	(0.02)	53.8	(1.4)	37.2	(0.9)	9.0	(1.0)
123 *Neuseeland*	3.15	(0.02)	54.6	(1.2)	32.0	(0.9)	13.4	(0.8)
23 *Chile*	3.15	(0.01)	53.1	(0.8)	33.1	(0.8)	13.8	(0.6)
Katar	3.14	(0.01)	52.7	(0.8)	33.7	(0.7)	13.6	(0.4)
Macau	3.13	(0.01)	53.2	(0.5)	33.2	(0.5)	13.6	(0.3)
2 *Saudi-Arabien*	3.12	(0.01)	50.9	(0.9)	35.0	(0.8)	14.1	(0.6)
VG OECD	**3.10**	**(0.02)**	**51.4**	**(1.1)**	**34.9**	**(0.9)**	**13.7**	**(0.8)**
2 *Spanien*	3.09	(0.01)	50.8	(0.6)	34.5	(0.4)	14.7	(0.4)
Irland	3.08	(0.02)	50.9	(1.3)	34.7	(0.9)	14.5	(1.0)
23 *USA*	3.08	(0.01)	50.2	(0.9)	34.7	(0.6)	15.1	(0.5)
VG EU	**3.07**	**(0.02)**	**50.0**	**(1.1)**	**35.2**	**(0.9)**	**14.8**	**(0.9)**
2 *Bosnien und Herzegowina*	3.07	(0.01)	46.3	(1.0)	42.3	(1.0)	11.4	(0.7)
2 *Serbien*	3.06	(0.02)	48.9	(1.3)	35.1	(1.0)	15.9	(1.0)
Internationaler Mittelwert	**3.06**	**(0.02)**	**48.8**	**(1.1)**	**34.5**	**(0.9)**	**16.7**	**(0.8)**
Nordmazedonien	3.06	(0.02)	49.1	(1.2)	33.6	(0.8)	17.3	(0.8)
2 *Slowenien*	3.05	(0.02)	45.4	(1.0)	42.6	(0.8)	12.0	(0.6)
23 *Albanien*	3.04	(0.04)	48.7	(1.9)	33.0	(1.6)	18.4	(1.6)
Oman	3.02	(0.01)	45.6	(0.7)	38.5	(0.6)	15.9	(0.5)
Kuwait	2.99	(0.02)	45.0	(1.3)	36.6	(1.0)	18.4	(1.0)
Japan	2.97	(0.02)	40.3	(1.4)	43.2	(1.0)	16.5	(0.9)
Bulgarien	2.94	(0.03)	46.6	(1.4)	31.2	(1.0)	22.2	(1.5)
2 *Kosovo*	2.92	(0.02)	41.6	(1.0)	38.2	(1.1)	20.3	(0.9)
Slowakei	2.92	(0.03)	40.8	(1.6)	39.7	(1.1)	19.5	(1.6)
2 *Montenegro*	2.92	(0.01)	42.0	(0.8)	35.9	(0.7)	22.1	(0.7)
Iran	2.91	(0.02)	39.7	(1.2)	38.2	(1.0)	22.1	(0.8)
2 *Kasachstan*	2.89	(0.02)	38.9	(1.2)	41.5	(0.8)	19.6	(0.8)
2 *Zypern*	2.87	(0.02)	41.3	(1.1)	33.3	(0.8)	25.4	(1.0)
2 *Frankreich*	2.87	(0.02)	36.5	(1.0)	43.3	(1.0)	20.2	(0.9)
Deutschland	2.86	(0.01)	39.5	(0.8)	37.9	(0.6)	22.6	(0.7)
2 *Tschechien*	2.86	(0.01)	35.5	(0.8)	46.0	(0.6)	18.5	(0.6)
2 *Belgien (Frz. Gem.)*	2.85	(0.01)	36.7	(0.8)	40.6	(0.8)	22.7	(0.6)
23 *Brasilien*	2.85	(0.01)	39.6	(0.6)	36.8	(0.4)	23.7	(0.6)
Republik Korea (Südkorea)	2.83	(0.02)	38.6	(1.1)	36.8	(0.8)	24.6	(0.9)
Aserbaidschan	2.83	(0.02)	38.3	(0.8)	37.6	(0.9)	24.1	(0.9)
2 *Ungarn*	2.75	(0.03)	33.3	(1.0)	40.1	(1.0)	26.6	(1.2)
3 *Italien*	2.70	(0.02)	31.5	(1.1)	39.0	(1.0)	29.5	(1.1)
1 *Südafrika (5. Jgst.)*	2.69	(0.02)	31.7	(0.9)	39.5	(0.6)	28.8	(0.9)
2 *Usbekistan*	2.67	(0.02)	30.8	(1.2)	38.3	(1.2)	30.9	(1.1)
2 *Armenien*	2.61	(0.02)	28.7	(0.9)	36.7	(0.8)	34.7	(0.9)
Jordanien	2.54	(0.04)	30.7	(1.3)	28.6	(1.1)	40.7	(2.0)
Marokko	2.14	(0.03)	17.5	(0.7)	28.2	(1.1)	54.3	(1.5)
Benchmark-Teilnehmer								
2 *Ontario, Kanada*	3.35	(0.01)	64.6	(0.9)	28.7	(0.7)	6.7	(0.5)
Dubai, VAE	3.35	(0.01)	65.7	(0.7)	27.2	(0.6)	7.1	(0.3)
Schardscha, VAE	3.17	(0.01)	54.6	(0.8)	33.1	(0.7)	12.3	(0.5)
23 *Québec, Kanada*	3.16	(0.02)	53.5	(0.9)	35.5	(0.6)	11.0	(0.6)
Abu Dhabi, VAE	3.11	(0.01)	50.7	(0.8)	36.4	(0.6)	12.9	(0.5)



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen mit hohem selbsteingeschätzten Kompetenzerleben im Umgang mit digitalen Medien
% der Schüler:innen mit mittlerem selbsteingeschätzten Kompetenzerleben im Umgang mit digitalen Medien
% der Schüler:innen mit niedrigem selbsteingeschätzten Kompetenzerleben im Umgang mit digitalen Medien

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Die drei Kategorien des Kompetenzerlebens ergeben sich aus der Rekodierung der Mittelwerte: hoch = 1 bis 2; mittel > 2 bis 3; niedrig > 3 bis 4.

 © TIMSS 2023

signifikant höher einschätzen als in Deutschland. In diesen Staaten schätzen rund zwei Drittel der Viertklässler:innen ihre Kompetenzen als *stark* ein (Werte von > 3 bis 4), während dieser Anteil in Deutschland nur bei knapp 40 Prozent liegt.

Die bisherigen Darstellungen zur Verfügbarkeit und Nutzung digitaler Medien durch die Schüler:innen im häuslichen und schulischen Umfeld unterliegen der Annahme, dass diese auch bedeutsam für den Kompetenzerwerb in den betrachteten Domänen Mathematik und Naturwissenschaften sind. In den Tabellen 5.6 (für den Mathematikunterricht) und 5.7 (für den naturwissenschaftlichen Sachunterricht) wird diese Annahme mithilfe von Regressionsmodellen überprüft. Dabei werden in sogenannten Mehrebenenmodellen zunächst für die individuellen Schüler:innen relevante Merkmale in die Berechnungen aufgenommen und diese dann um Informationen zur Verfügbarkeit digitaler Medien in den Schulen und deren Nutzung im Unterricht erweitert. Für die Modellierung wird zunächst eine nähere Betrachtung der Zusammenhänge von individuellen Voraussetzungen der Schüler:innen vorgenommen, auf die durch Lehrkräfte bzw. Schulleitungen kein bis wenig Einfluss genommen werden kann. Anschließend wird schrittweise überprüft, welche Zusammenhänge die schulischen und unterrichtlichen Merkmale, bezogen auf digitale Medien, mit den von den Schüler:innen erreichten Kompetenzstände haben. Mit diesen Modellen wird eine Exploration der aktuellen Praxis anhand der TIMSS-2023-Daten vorgenommen. Es sei bereits an dieser Stelle darauf hingewiesen, dass es zukünftig vertiefende Analysen, auch unter Berücksichtigung weiterer Datenquellen, braucht, um diese ersten explorativen Befunde zu prüfen und einzuordnen.

Im Modell 1 (Tabelle 5.6) sind als Kontrollvariable zunächst die häuslichen Lernbedingungen anhand des *Home Resources for Learning* (HRL)-Indizes[3] berücksichtigt, um Effekte des sozialen Hintergrundes empirisch zu kontrollieren und so mögliche Effekte digitalisierungsbezogener Merkmale deutlicher und bereinigter darstellen zu können. Hierbei zeigt sich, dass es einen signifikanten positiven Effekt des häuslichen Lernumfeldes auf die in TIMSS erfassten Mathematikkompetenzen gibt. Dabei beträgt der durchschnittliche Kompetenzunterschied von Schüler:innen, die nur über eine sehr geringe Ausstattung verfügen (die untersten 25 % des HRL-Indizes) und denen, die die umfangreichste Ausstattung haben (die obersten 25 % des HRL-Indizes) 21 Punkte, was knapp einem halben Schuljahr Lernfortschritt entspricht (Wendt et al., 2017).

Im Modell 2 werden zusätzlich zum HRL-Index auch die Verfügbarkeit digitaler Medien im häuslichen Kontext und die von den Schüler:innen selbst wahrgenommenen Kompetenzen im Umgang mit digitalen Medien für schulische Zwecke berücksichtigt. Hierbei zeigt sich, dass neben dem allgemeinen Lernumfeld (HRL-Index) die Verfügbarkeit von digitalen Medien zu Hause signifikant mit dem Kompetenzstand der Schüler:innen zusammenhängt. Im hypothetischen Vergleich von Schüler:innen, die in Haushalten leben, in denen es keine digitalen Medien gibt, und Schüler:innen, die in Haushalten leben, in denen alle in TIMSS 2023 abgefragten digitalen Medien (Computer, Laptop, Smartphone,

3 Der HRL-Index basiert auf fünf Variablen, die den sozioökonomischen Status abbilden. Es handelt sich um folgende Angaben: (1) *heimischer Buchbesitz* (Schülerangabe), (2) *heimischer Kinderbuchbesitz* (Elternangabe), (3) *Anzahl bildungsrelevanter Ressourcen* (Schülerangabe), (4) *höchster Bildungsabschluss* eines Elternteils (Elternangabe), (5) *höchster Berufsabschluss* eines Elternteils (Elternangabe). Skalenwerte unter 7.4 verweisen auf eine geringe Ausstattung; Werte über 11.8 auf eine reichhaltige Ausstattung (Mullis et al., 2020). Der Mittelwert dieses Indizes liegt in Deutschland bei 10.8 mit einer Standardabweichung von 1.7.

Tabelle 5.6: Mehrebenenmodell zur Vorhersage der mathematischen Leistungen von Schüler:innen

	Mathematik							
	Modell 1		Modell 2		Modell 3		Modell 4	
	b	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)
Konstante	310.1	(12.6)**	266.3	(15.6)**	266.2	(15.8)**	260.5	(16.9)**
Schülerebene								
Home Ressources for Learning (HRL)	21.0	(1.1)**	19.7	(1.1)**	20.3	(1.1)**	20.6	(1.1)**
Häusliche Verfügbarkeit digitaler Medien			49.8	(11.2)**	50.5	(11.7)**	50.3	(12.6)**
Kompetenzerleben im Umgang mit digitalen Medien			3.9	(1.6)*	2.6	(1.6)^{ns}	2.9	(1.7)^{ns}
Schulebene								
Verfügbarkeit Computer/Tablets[A]					-2.0	(2.1)^{ns}	-1.9	(2.3)^{ns}
Nutzungshäufigkeit digitaler Medien[B]							-10.8	(13.8)^{ns}
Wissen über Lernen mit digitalen Medien[C]							4.3	(6.2)^{ns}
Zugang zu digitalen Geräten[C]							5.9	(4.9)^{ns}
Unterstützung durch die Schule[C]							-1.9	(6.1)^{ns}
Erklärte Varianzanteile								
Zwischen den Schulen (21.1%)	0.0		0.0		18.2		17.0	
Innerhalb der Schulen (78.8%)	10.0		11.2		11.2		11.2	
Gesamt	7.9		8.8		12.7		12.4	

b = Regressionsgewicht (unstandardisiert)
Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p \leq .05$); ** = signifikant ($p \leq .01$)
A = Computer/Tablets geteilt durch die Anzahl der Viertklässler:innen
B = Mittelwert von sechs Variablen (*Aufgaben üben; Komplexe Aufgaben lösen; Grafiken/Tabellen u. a. erstellen; Mathematische Spiele spielen; Lehrbuch lesen/Lernvideos ansehen; Tests/Klassenarbeiten schreiben;* 1 = *mindestens einmal pro Woche;* 0 = *seltener*)
C = Kodierung jeweils 1 = *überhaupt nicht eingeschränkt;* 0 = *etwas oder sehr eingeschränkt*

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

E-Reader und zuverlässiger Internetzugang) vollständig verfügbar sind, beträgt der mittlere Unterschied in den mathematischen Kompetenzen 49.8 Punkte. Dies entspricht einem möglichen Unterschied von bis zu zwei Schuljahren (Wendt et al., 2017). Das selbstberichtete Kompetenzerleben von Viertklässler:innen im Umgang mit digitalen Medien steht ebenso in einem positiven Zusammenhang mit den mathematischen Kompetenzen, dieser ist jedoch mit ca. 4 Punkten vernachlässigbar gering.

Die Berücksichtigung der Information zur Verfügbarkeit digitaler Medien, die die Schüler:innen der vierten Klassen für Unterrichtszwecke nutzen können (Modell 3), weist keinen empirischen Zusammenhang mit Kompetenzunterschieden zwischen den Schüler:innen in Mathematik auf (siehe Tabelle 5.6). Die *Nutzungshäufigkeit digitaler Medien* im Mathematikunterricht und die Herausforderungen aus Perspektive der Lehrkräfte stehen ebenfalls in keinem empirisch signifikanten Zusammenhang mit den Mathematikkompetenzen der Viertklässler:innen (Modell 4). Anzumerken ist an dieser Stelle, dass die Angaben zur Nutzungshäufigkeit keine Aussage über die Qualität der Nutzung im Unterricht zulassen und somit zukünftigen differenzierten Untersuchungen vorbehalten bleiben.

Bei der Analyse von Zusammenhängen individueller Schülermerkmale und unterrichts- bzw. schulbezogener Merkmale in Bezug auf digitale Medien und den naturwissenschaftlichen Kompetenzständen der Viertklässler:innen zeigen sich, wie in Tabelle 5.7 dargestellt, folgende Befunde: Statistisch auffällige posi-

Tabelle 5.7: Mehrebenenmodell zur Vorhersage der naturwissenschaftlichen Leistungen von Schüler:innen

	Naturwissenschaften							
	Modell 1		Modell 2		Modell 3		Modell 4	
	b	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)
Konstante	252.6	(15.4)**	214.1	(17.6)**	210.8	(18.4)**	215.9	(19.9)**
Schülerebene								
Home Ressources for Learning (HRL)	25.8	(1.3)**	24.7	(1.3)**	25.4	(1.3)**	25.8	(1.4)**
Häusliche Verfügbarkeit digitaler Medien			47.1	(10.8)**	46.6	(11.1)**	38.9	(12.0)**
Kompetenzerleben im Umgang mit digitalen Medien			2.6	(2.3)^{ns}	1.8	(2.4)^{ns}	1.9	(2.7)^{ns}
Schulebene								
Verfügbarkeit Computer/Tablets^{A}					-0.4	(2.5)^{ns}	-1.1	(2.7)^{ns}
Nutzungshäufigkeit digitaler Medien^{B}							-54.8	(25.8)*
Wissen über Lernen mit digitalen Medien^{C}							3.5	(6.3)^{ns}
Zugang zu digitalen Geräten^{C}							-0.1	(6.0)^{ns}
Unterstützung durch die Schule^{C}							-6.8	(7.9)^{ns}
Erklärte Varianzanteile								
Zwischen den Schulen (24.3%)	0.0		0.0		17.0		19.1	
Innerhalb der Schulen (75.6%)	17.5		18.0		18.0		18.0	
Gesamt	13.2		13.6		17.8		18.3	

b = Regressionsgewicht (unstandardisiert)
Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p \leq .05$); ** = signifikant ($p \leq .01$)
A = Computer/Tablets geteilt durch die Anzahl der Viertklässler:innen
B = Mittelwert von sechs Variablen (*Virtuelle Experimente durchführen; Komplexe Aufgaben lösen; Grafiken/Tabellen u. a. erstellen; Naturwissenschaftliche Spiele spielen; Lehrbuch lesen/Lernvideos ansehen; Tests/Klassenarbeiten schreiben;* 1 = *mindestens einmal pro Woche;* 0 = *seltener*)
C = Kodierung jeweils 1 = *überhaupt nicht eingeschränkt;* 0 = *etwas oder sehr eingeschränkt*

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

tive Zusammenhänge weisen die heimische Ausstattung mit bildungsrelevanten Gütern (HRL-Index; Modell 1) und die durchschnittliche häusliche Verfügbarkeit von digitalen Medien auf (Modell 2). Werden wiederum die erwarteten Kompetenzen von Schüler:innen mit geringer häuslicher Lernausstattung (die niedrigsten 25 % des HRL-Index) mit denen von Schüler:innen mit besserer häuslicher Ausstattung (die obersten 25 % des HRL-Index) verglichen, geht dies mit einem Unterschied in den naturwissenschaftlichen Kompetenzen einher, der dem Lernfortschritt von rund einem halben Lernjahr entspricht. Die Berücksichtigung des HRL-Indizes führt dazu, dass die weiteren auf digitale Medien bezogenen Erklärungsvariablen unabhängig von diesen allgemeinen häuslichen Ausstattungsmerkmalen betrachtet werden können. Werden Kinder, die in Haushalten leben, die über keine digitalen Medien verfügen, und solche, die alle in TIMSS 2023 abgefragten digitalen Medien (Computer, Laptop, Smartphone, E-Reader und zuverlässiger Internetzugang) zur Verfügung haben, verglichen, entspricht der beobachtbare Unterschied wiederum einem Äquivalent von rund einem Lernjahr. Die selbsteingeschätzten Kompetenzen der Schüler:innen im Umgang mit digitalen Medien stehen in keinem signifikanten Zusammenhang mit den naturwissenschaftlichen Kompetenzen.

Wie bereits für den Mathematikunterricht beschrieben stehen auch im naturwissenschaftlichen Sachunterricht die Verfügbarkeit digitaler Medien für den Gebrauch durch Viertklässler:innen sowie mögliche Hinderungsgründe in

Bezug auf die Nutzung nicht im Zusammenhang mit den naturwissenschaftlichen Kompetenzen der Schüler:innen. Dieser Befund überrascht insofern nicht, da die unterrichtliche Nutzung verschiedener digitaler Medien generell auf einem niedrigen Niveau liegt (Tabelle 5.3). Auf den ersten Blick überraschend ist der signifikante negative Effekt der *Nutzungshäufigkeit digitaler Medien* im Unterricht für verschiedene Tätigkeiten auf den Kompetenzstand der Viertklässler:innen. Mit 54.8 Leistungspunkten Unterschied zwischen den Gruppen ist dieser Wert zunächst als auffällig groß zu charakterisieren. Jedoch stellt dieser Wert den maximalen Unterschied möglicher durchschnittlicher Kompetenzen von Schüler:innen dar, die einerseits von Lehrkräften unterrichtet werden, die keines der abgefragten digitalen Medien *mindestens einmal pro Woche* nutzen, und andererseits solchen Schüler:innen, für die Lehrkräfte angeben, alle Medienarten *mindestens einmal wöchentlich* im Fachunterricht einzusetzen. Der zusätzliche Einsatz eines der digitalen Medien in der Woche entspricht damit einem durchschnittlichen Kompetenzunterschied von rund 8 Punkten. Das negative Vorzeichen verweist darauf, dass eine höhere Nutzung mit geringeren naturwissenschaftlichen Kompetenzen einhergeht. Da es sich bei TIMSS-Daten um Querschnittsdaten handelt, kann dieser Befund allerdings genauso darauf hindeuten, dass bei geringeren naturwissenschaftlichen Kompetenzen häufiger digitale Medien zum Einsatz kommen. Es kann vermutet werden, dass der Einsatz digitaler Medien im Sachunterricht eher zur spezifischen Förderung genutzt wird, ihre Nutzung aber nicht in der Breite lernwirksam im Unterricht implementiert ist.

5.6 Nutzung digitaler Medien in der Schule und im Unterricht in Deutschland

Nach den Ergebnissen zur schulischen und häuslichen Ausstattung mit digitalen Medien werden in diesem Abschnitt die Ergebnisse zu der Nutzung digitaler Medien berichtet. Dazu wurden in TIMSS 2023 über die international erfragten Informationen hinaus zusätzlich national vertiefende Fragen zur schulischen und unterrichtsbezogenen Nutzung digitaler Medien eingesetzt. Dabei wird die Nutzungshäufigkeit im Fachunterricht und im häuslichen Kontext zu schulischen Zwecken sowohl aus der Perspektive der Lehrkräfte als auch aus Schülerperspektive betrachtet. Vor dem Hintergrund der besonderen Herausforderungen für das Lehren und Lernen während der COVID-19-Pandemie wird zudem berichtet, in welchem zeitlichen Rahmen der Unterricht an die veränderten Bedingungen hinsichtlich der Nutzung digitaler Medien durch die Lehrkraft angepasst wurde.

5.6.1 Nutzung digitaler Medien für schulische Zwecke

Um einen Eindruck zu erhalten, wo bzw. wie häufig Viertklässler:innen in Deutschland digitale Medien für schulische Zwecke nutzen, sind in Tabelle 5.8 Informationen zur Nutzungshäufigkeit dargestellt. Es zeigt sich, dass zwar statistisch signifikant aber nur geringfügig mehr Grundschüler:innen angeben, digitale Medien mindestens *jeden Tag* außerhalb der Schule für schulische Zwecke (z. B. für Hausaufgaben) zu nutzen (28.5 %) als innerhalb der Schule (25.8 %). Ungefähr ein Drittel der Grundschüler:innen berichtet, dass sie digitale Medien *wöchentlich* bzw. *monatlich* in der Schule (35.6 %) und außerhalb der Schule für schulische Zwecke (30.6 %) nutzen. Zudem zeigt sich ein hoher Anteil von

Tabelle 5.8: Nutzungshäufigkeit digitaler Medien durch Schüler:innen für schulische Zwecke

	Nutzungsort digitaler Medien durch Schüler:innen											
	mehrmals am Tag		jeden Tag		mind. einmal pro Woche		mind. einmal pro Monat		weniger als einmal pro Monat		nie	
	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
in der Schule für schulische Zwecke (z.B. für den Unterricht)	17.1	(0.7)	8.7	(0.4)	21.9	(0.7)	13.7	(0.5)	10.9	(0.5)	27.7	(0.8)
außerhalb der Schule für schulische Zwecke (z.B. Hausaufgaben)	14.4	(0.5)	14.1	(0.5)	18.6	(0.5)	12.0	(0.5)	12.9	(0.5)	28.0	(0.7)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Schüler:innen, die angeben, sehr selten (*weniger als einmal im Monat*) oder sogar *nie* digitale Medien in der Schule (38.6 %) oder außerhalb der Schule für schulische Zwecke (40.9 %) zu nutzen.

In einer zusätzlichen Frage wurden die Schüler:innen nach der Nutzungshäufigkeit verschiedener digitaler Medien gefragt. Hierbei zeigt sich, dass gut zwei Drittel (68.8 %) der Viertklässler:innen in Deutschland angeben, zu Hause ein Tablet oder Laptop *mindestens einmal pro Woche* für schulische Zwecke zu nutzen (Tabelle 5.9). Hingegen gibt nur rund ein Drittel (35.2 %) der Schüler:innen an, im Unterricht *mindestens einmal pro Woche* ein Tablet oder Laptop für schulische Zwecke zu nutzen. Allerdings zeigen die Befunde auch, dass es einen gewissen Anteil an Viertklässler:innen in Deutschland gibt, die berichten, dass an ihrer Schule entweder keine Tablets oder Laptops für den Unterricht zur Verfügung stehen (17.0 %) oder sie diese *nie* nutzen (19.3 %). Diese Anteile fallen für den häuslichen Kontext deutlich geringer aus. Diesbezüglich gibt jeweils weniger als ein Zehntel der Schüler:innen an, keine entsprechenden Geräte für schulische Zwecke zur Verfügung zu haben (8.3 %) bzw. diese *nie* zu nutzen (9.1 %).

Tabelle 5.9: Nutzungshäufigkeit verschiedener digitaler Medien durch Viertklässler:innen für die Schule (im Unterricht oder zu Hause)

	Nutzungshäufigkeit digitaler Medien durch Schüler:innen											
	täglich		mehrmals pro Woche		etwa einmal pro Woche		weniger als einmal pro Woche		nie		gibt es nicht/habe ich nicht	
	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
Tablet oder Laptop im Unterricht	7.2	(0.5)	10.4	(0.7)	17.6	(0.8)	28.4	(1.0)	19.3	(0.8)	17.0	(0.8)
Tablet oder Laptop zu Hause	33.6	(0.7)	22.7	(0.6)	12.5	(0.4)	13.8	(0.5)	9.1	(0.3)	8.3	(0.4)
Lernplattform	12.7	(0.5)	16.6	(0.6)	18.1	(0.6)	23.2	(0.6)	15.5	(0.6)	13.8	(0.5)
Smartboard	22.4	(1.1)	10.9	(0.5)	8.2	(0.4)	9.6	(0.4)	18.2	(0.7)	30.7	(0.8)

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Weiterhin wird in Tabelle 5.9 die Nutzungshäufigkeit von Lernplattformen sowie Smartboards (interaktiven Whiteboards) betrachtet. Mehr als 40 Prozent der Viertklässler:innen berichten, dass sie Lernplattformen (47.4 %) mindestens *etwa einmal pro Woche* für schulische Zwecke nutzen. Auch hinsichtlich der Nutzung von Smartboards geben 41.5 Prozent der Viertklässler:innen an, diese mindestens *etwa einmal pro Woche* für schulische Zwecke zu nutzen. Allerdings sei auch hier darauf hingewiesen, dass ein nicht unerheblicher Anteil an Schüler:innen angibt, *nie* Lernplattformen (15.5 %) bzw. Smartboards (18.2 %) zu nutzen oder keine Lernplattformen (13.8 %) bzw. Smartboards (30.7 %) für schulische Zwecke zur Verfügung zu haben.

In der Zusammenschau der Befunde der Tabellen 5.8 und 5.9 bestätigt sich der Eindruck, dass die Nutzung digitaler Medien insbesondere in den Schulen, aber auch zu Hause für schulische Zwecke in Deutschland nicht regelmäßig stattfindet. Dies korrespondiert mit den für Mathematik (Tabelle 5.2) und den für Naturwissenschaften (Tabelle 5.3) festgestellten Befunden, dass digitale Medien in Deutschland im internationalen Vergleich, aber auch in der Gegenüberstellung mit den Vergleichsgruppen der OECD und EU, seltener genutzt werden.

Die berichteten Nutzungshäufigkeiten unterschiedlicher digitaler Medien für schulische Zwecke geben zudem zwar Auskunft über die Quantität der Nutzung durch die Schüler:innen, nicht jedoch über die Lernangebote, in denen digitale Medien im Fachunterricht zum Einsatz kommen. Daher wurden in TIMSS 2023 die Grundschullehrkräfte ergänzend gefragt, wie häufig aus ihrer Sicht Viertklässler:innen Lernangebote im Mathematik- bzw. im naturwissenschaftlichen Sachunterricht nutzen, in denen digitale Medien eingesetzt werden. Um den Spezifika der beiden Fächer Rechnung zu tragen, wurden jeweils auch domänenspezifische Unterrichtsaktivitäten mit digitalen Medien erfragt.

Tabelle 5.10: Nutzungshäufigkeit digitaler Medien im Mathematikunterricht durch Schüler:innen für verschiedene Lernangebote

	Nutzungshäufigkeit digitaler Medien im Mathematikunterricht durch Schüler:innen							
	in jeder oder fast jeder Stunde		**ein- bis zweimal pro Woche**		**einmal oder zweimal pro Monat**		**nie oder fast nie**	
	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
individuelles Bearbeiten von Lernmaterialien im eigenen Arbeitstempo	10.8	(1.9)	25.1	(2.7)	29.9	(2.6)	34.2	(2.8)
Fertigkeiten und Prozeduren üben	7.9	(1.8)	30.0	(3.0)	34.1	(2.6)	28.0	(2.7)
nach Ideen und Informationen suchen	1.7	(0.8)	9.9	(2.0)	34.2	(2.9)	54.1	(2.9)
Lernvideos ansehen	1.5	(0.7)	12.3	(2.1)	44.5	(3.0)	41.7	(3.3)
sich mit mathematischen Grundsätzen und Konzepten befassen	0.9	(0.6)	7.2	(1.5)	20.9	(2.2)	71.0	(2.4)
Reflexion ihrer Lernerfahrungen (z.B. mit einem digitalen Lernjournal oder Lerntagebuch)	0.4	(0.0)	0.7	(0.6)	7.3	(1.8)	91.5	(1.9)
eigenständiges Arbeiten und Planen im Rahmen von Projektarbeit	0.0	(0.0)	1.0	(0.6)	15.6	(2.1)	83.3	(2.0)

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

© TIMSS 2023

Tabelle 5.11: Nutzungshäufigkeit digitaler Medien im naturwissenschaftlichen Sachunterricht durch Schüler:innen für verschiedene Lernangebote

	Nutzungshäufigkeit digitaler Medien im naturwissenschaftlichen Sachunterricht durch Schüler:innen							
	in jeder oder fast jeder Stunde		ein- bis zweimal pro Woche		einmal oder zweimal pro Monat		nie oder fast nie	
	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
nach Ideen und Informationen suchen	8.6	(1.7)	34.3	(3.2)	47.9	(3.1)	9.2	(1.8)
individuelles Bearbeiten von Lernmaterialien im eigenen Arbeitstempo	7.0	(1.7)	15.9	(2.2)	40.4	(3.4)	36.8	(3.0)
eigenständiges Arbeiten und Planen im Rahmen von Projektarbeit	2.7	(1.0)	12.9	(1.9)	44.4	(3.4)	40.0	(3.0)
naturwissenschaftliche Verfahren oder Experimente durchführen	1.8	(0.7)	3.7	(1.3)	28.6	(2.5)	65.9	(2.7)
Naturphänomene anhand von Simulationen erforschen	1.6	(0.4)	3.8	(1.5)	18.6	(2.5)	76.0	(2.7)
Fähigkeiten und Verfahren üben	1.5	(0.8)	13.6	(2.5)	36.2	(3.0)	48.6	(3.0)
Lernvideos ansehen	1.4	(0.8)	28.6	(3.0)	55.2	(3.4)	14.7	(2.3)
Reflexion ihrer Lernerfahrungen (z.B. mit einem digitalen Lernjournal oder Lerntagebuch)	1.2	(0.6)	2.1	(0.7)	12.0	(2.3)	84.8	(2.4)

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

In den Tabellen 5.10 und 5.11 ist die Nutzungshäufigkeit digitaler Medien durch die Schüler:innen nach Angabe der Lehrkräfte für den Mathematik- bzw. naturwissenschaftlichen Sachunterricht angegeben. In beiden Domänen wird neben der Informationsbeschaffung (*nach Ideen und Informationen suchen*) auch danach gefragt, wie oft Lernvideos angesehen oder Lernmaterialien im eigenen Arbeitstempo bearbeitet bzw. Fertigkeiten und Prozeduren geübt werden. Darüber hinaus wurden die Lehrkräfte gefragt, wie häufig die Viertklässler:innen mit digitalen Medien selbstständig arbeiten bzw. diese zur Reflexion des eigenen Lernens nutzen. Domänenspezifisch wird das *Befassen mit mathematischen Grundsätzen und Konzepten* bzw. das *Durchführen von naturwissenschaftlichen Verfahren oder Experimenten* und das *Erforschen von Naturphänomenen anhand von Simulationen* im naturwissenschaftlichen Sachunterricht erfragt.

Werden die für den Mathematikunterricht möglichen Einsatzoptionen digitaler Medien betrachtet (siehe Tabelle 5.10), so zeigt sich, dass nach Einschätzung der Lehrkräfte rund ein Drittel (35.9 %) der Schüler:innen Lernmaterialien *mindestens einmal pro Woche* individuell im eigenen Arbeitstempo bearbeiten bzw. 37.9 Prozent der Schüler:innen Fertigkeiten und Prozeduren üben. Deutlich seltener sehen Schüler:innen Lernvideos (mindestens wöchentlich: 13.8 %) bzw. suchen in Mathematik nach Ideen und Informationen (mindestens wöchentlich: 11.6 %). Nach Angabe der Lehrkräfte befassen sich 8.1 Prozent der Schüler:innen *mindestens einmal pro Woche* mit mathematischen Grundsätzen und Konzepten.

Jeweils nur ca. 1 Prozent der Schüler:innen reflektiert *mindestens einmal pro Woche* die eigenen Lernerfahrungen anhand eines digitalen Lernjournals oder Lerntagebuchs oder nutzt digitale Medien zum eigenständigen Arbeiten und Planen im Rahmen von Projektarbeit in Mathematik.

Bezogen auf den naturwissenschaftlichen Sachunterricht zeigt sich im Vergleich zum Mathematikunterricht eine deutlich abweichende Lehrkräfteeinschätzung der Nutzungshäufigkeit digitaler Medien für verschiedene Unterrichtsaktivitäten durch die Schüler:innen. Tabelle 5.11 ist zu entnehmen, dass rund ein Viertel (22.9 %) der Schüler:innen von Lehrkräften unterrichtet wird, die *mindestens einmal pro Woche Lernmaterialien individuell im eigenen Arbeitstempo bearbeiten* bzw. 15.1 Prozent der Schüler:innen *Fähigkeiten und Verfahren üben* lassen – was deutlich unter den Anteilen der entsprechenden Nutzung in Mathematik liegt (*Fertigkeiten und Prozeduren üben*: 37.9 %).

Nach Einschätzung der Lehrkräfte sehen 30.0 Prozent der Schüler:innen im naturwissenschaftlichen Sachunterricht *mindestens wöchentlich* Lernvideos und 42.9 Prozent der Schüler:innen suchen *mindestens einmal wöchentlich* im naturwissenschaftlichen Sachunterricht *nach Ideen und Informationen*, was beides wiederum öfter genutzt wird als im Mathematikunterricht (13.8 % bzw. 11.6 %). Wie auch in Mathematik (1.1 %) nutzen Schüler:innen im naturwissenschaftlichen Sachunterricht relativ selten digitale Medien, um ihre eigenen Lernerfahrungen beispielsweise anhand eines digitalen Lernjournals oder Lerntagebuchs zu reflektieren (3.3 %). Spezifisch auf den naturwissenschaftlichen Sachunterricht bezogen zeigt sich, dass ein relativ kleiner Anteil der Grundschulkinder (5.4 %) *mindestens einmal pro Woche* Naturphänomene anhand von Simulationen erforscht bzw. sich mit naturwissenschaftlichen Verfahren oder Experimenten befasst (5.5 %).

Aufgrund der nicht vergleichbaren curricularen Strukturen des Mathematikunterrichts (stark sequenziell strukturiert) und des Sachunterrichts (gering sequenziell strukturiert; Blömeke et al., 2011) lässt sich aus den unterschiedlichen Nutzungspraktiken digitaler Medien in den beiden betrachteten Domänen jedoch kein qualitativer Unterschied ableiten, da ein Mehr oder Weniger an der Nutzung der spezifischen digitalen Medien eine Frage der didaktischen Passung zu den domänenspezifischen Unterrichtsinhalten darstellt (für eine vertiefende Erfassung der curricularen Strukturen des Fachunterrichts siehe in diesem Band Kapitel 3 und 4).

5.6.2 Unterrichtsanpassung im Zuge der COVID-19-Pandemie

Im Rahmen von TIMSS 2023 wurde von den Lehrkräften in Deutschland erfragt, wie schnell der Unterricht an die pandemiebedingten Veränderungen im Hinblick auf die Nutzung digitaler Medien angepasst werden konnte. Tabelle 5.12 bildet die Einschätzungen der Grundschullehrkräfte hierzu ab. Zunächst zeigt sich, dass zu Beginn der COVID-19-Pandemie nach Angaben der Lehrkräfte lediglich deutlich weniger als ein Fünftel (17.1 %) der Viertklässler:innen in Deutschland von Lehrkräften unterrichtet wurde, für die die benötigte Hardware sofort verfügbar bzw. diese bereits auf die neuen Bedürfnisse angepasst war. Für weniger als 10 Prozent der Schüler:innen berichten Lehrkräfte, dass Software (9.6 %), elektronische Lernmaterialien (z. B. Arbeitsblätter: 8.8 %) oder elektronische Lehrmaterialien (z. B. Tafelbilder: 6.3 %) sofort verfügbar oder bereits angepasst waren. Diese Befunde weisen darauf hin und bestätigen rückblickend nochmals, dass Lehrkräfte an Grundschulen in Deutschland hinsichtlich der digitalisierungsbezogenen Voraussetzungen für einen Unterricht unter Pandemiebedingungen kaum hinreichend ausgestattet waren. Weiterhin zeigt sich jedoch auch, dass es vielfach gelungen ist, zeitnah entsprechende Voraussetzungen zu schaffen. Für

Tabelle 5.12: Dauer der Anpassung von digitalisierungsbezogenen Voraussetzungen und Grundlagen für den Unterricht bedingt durch die COVID-19-Pandemie

	Dauer der Unterrichtsanpassung bedingt durch die COVID-19-Pandemie											
	sofort/ war bereits angepasst		innerhalb von 1 Woche		innerhalb von 1 Monat		innerhalb von 2 bis 3 Monaten		innerhalb von 4 bis 6 Monaten		mehr als 6 Monate		konnte/ wollte ich nicht anpassen	
	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
Hardware	17.1	(2.3)	22.0	(2.5)	15.4	(2.0)	13.4	(2.2)	4.9	(1.3)	6.7	(1.5)	20.6	(2.3)
Software	9.6	(1.9)	21.5	(2.5)	20.3	(2.4)	16.4	(1.9)	7.7	(1.5)	5.4	(1.5)	19.2	(2.3)
elektronische Lernmaterialien (z.B. Arbeitsblätter)	8.8	(2.1)	26.9	(2.9)	24.2	(2.6)	9.6	(1.4)	4.1	(0.9)	4.4	(1.5)	21.9	(2.6)
elektronische Lehrmaterialien (z.B. Tafelbilder, Poster)	6.3	(1.6)	21.1	(2.7)	27.1	(2.8)	10.3	(1.7)	4.8	(1.1)	4.1	(1.3)	26.4	(2.7)

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

fast zwei Fünftel (37.4 %) der Schüler:innen in Deutschland geben Lehrkräfte an, dass es maximal einen Monat gedauert hat, entsprechende Voraussetzungen bezüglich der Hardware zu ermöglichen. Bei der Software sind es 41.8 Prozent, bei den elektronischen Lernmaterialien sogar 51.1 Prozent und bei den elektronischen Lehrmaterialien 48.2 Prozent. Insgesamt zeigt sich hinsichtlich der Unterrichtsumstellung, dass notwendige Anpassungen der unterschiedlichen Aspekte bereits zu guten Teilen *innerhalb von einer Woche* (rund 21 bis 27 %) spätestens aber *innerhalb eines Monats* (ca. 15 bis 27 %) realisiert wurden. Für ca. 10 bis 16 Prozent der Schüler:innen wurden die abgefragten Aspekte erst mittelfristig (*innerhalb von zwei bis drei Monaten*) angepasst. Längere Umstellungen in den Schulen (länger als vier Monate) betrafen rund 10 Prozent der Schüler:innen. Die Ergebnisse zeigen jedoch auch, dass etwa ein Fünftel der Schüler:innen eine Grundschule besuchten, an der die Lehrkräfte angeben, dass Hardware (20.6 %), Software (19.2 %) sowie elektronische Lernmaterialien (21.9 %) nicht angepasst werden konnten bzw. sie dies nicht wollten. Bei diesen auf den ersten Blick recht hohen Anteilen muss an dieser Stelle einschränkend darauf hingewiesen werden, dass gerade mit Blick auf digitalisierungsbezogene Voraussetzungen auf Ebene der einzelnen Lehrkräfte nur eingeschränkt Handlungsspielraum bestand, da das gewählte und ermöglichte Vorgehen auch mit der jeweiligen Strategie und den finanziellen Ressourcen der Schule sowie des Schulträgers zusammenhing. Bei den elektronischen Lehrmaterialien geben 26.4 Prozent der Lehrkräfte an, dass sie diese durch die COVID-19-Pandemie notwendigen Unterrichtsanpassungen selbst nicht vornehmen konnten bzw. wollten. Hier könnte vermutet werden, dass für die Lehrkräfte die Unterstützung und Begleitung der Lernprozesse der Kinder während der COVID-19-Pandemie im Fokus stand und weniger die Weiterentwicklung von eigenen Lehrmaterialien.

5.7 Bedarf und Teilnahme von Lehrkräften an Fortbildungen zu digitalen Medien im internationalen Vergleich und im Trend

Im Rahmen von TIMSS 2023 wurden die Grundschullehrkräfte der an TIMSS teilnehmenden Schüler:innen gefragt, ob sie in den letzten zwei Jahren vor der Erhebung, d. h. im Zeitraum von Frühjahr 2021 bis zum Frühjahr 2023, an verschiedenen Fortbildungsveranstaltungen teilgenommen haben. Mit Blick auf den Fokus des vorliegenden Kapitels wird im Folgenden spezifisch die digitalisierungsbezogene Fortbildungsteilnahme betrachtet. Für eine vertiefende Betrachtung der Fortbildungsteilnahme von Lehrkräften in verschiedenen weiteren Bereichen sei auf Kapitel 10 in diesem Band verwiesen.

Die Abbildungen 5.3 (für Mathematik) und 5.4 (für den naturwissenschaftlichen Sachunterricht) zeigen im internationalen Vergleich, wie viele der Viertklässler:innen von Lehrkräften unterrichtet werden, die nach eigenen Angaben in den letzten zwei Jahren Fortbildungen zur *Integration von Informationstechnologien* besucht haben. Darüber hinaus ist ein Vergleich der Zyklen von TIMSS 2019 und 2023 abgebildet, der die Angaben der Lehrkräfte zu der Teilnahme an Fortbildungen darstellt.

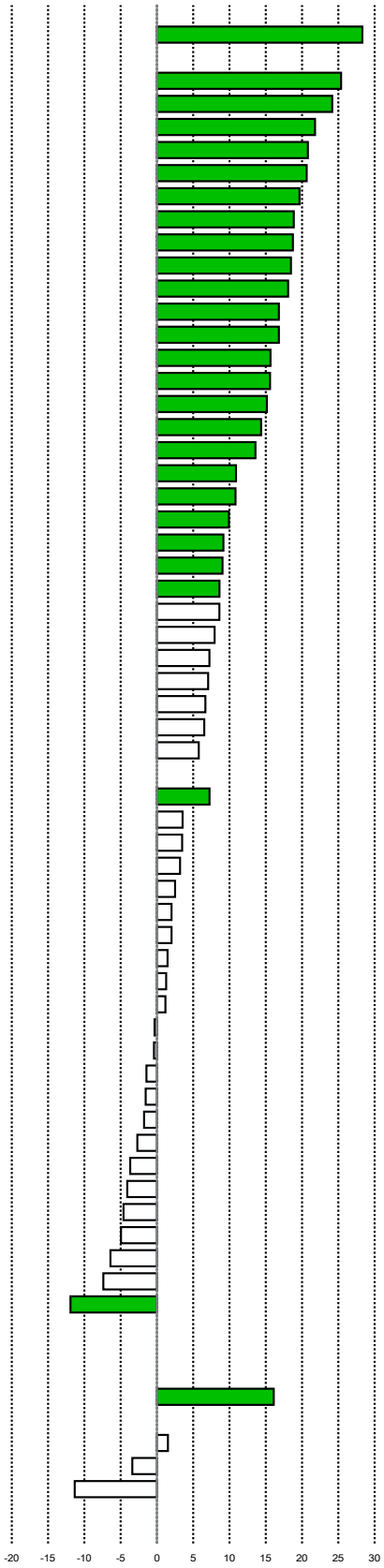
Für Deutschland zeigt sich, dass 2023 fast jedes vierte Grundschulkind (24.0 %) in der vierten Jahrgangsstufe in Deutschland von einer Mathematiklehrkraft unterrichtet wird, die nach eigenen Angaben in den letzten zwei Jahren an einer Fortbildung zur *Integration von Informationstechnologien* teilgenommen hat (siehe Abbildung 5.3). Dieser Anteil liegt signifikant unter dem internationalen Mittelwert (41.7 %) und auch unter den Anteilen in den Vergleichsgruppen EU (33.8 %) und OECD (33.4 %) und fällt signifikant höher aus als noch im vorherigen Studienzyklus von 2019 (8.4 %). Nur in einem Staat (Iran) ist die Differenz zwischen 2023 und 2019 signifikant größer als in Deutschland. Die Unterschiede zwischen 2019 und 2023 für die Vergleichsgruppen des internationalen Mittelwertes, der EU und der OECD sind nicht signifikant.

Abbildung 5.4 zeigt die Fortbildungsteilnahme der befragten Sachunterrichtslehrkräfte zur *Integration von Informationstechnologien* in Deutschland im internationalen Vergleich von 2019 und 2023. Der rechte Teil der Abbildung stellt die Differenz von 2019 auf 2023 grafisch dar. Grün eingefärbte Balken zeigen, dass es sich um eine signifikante Differenz handelt. Der Blick auf die Fortbildungsteilnahme von Sachunterrichtslehrkräften zeigt im Vergleich zu den Mathematiklehrkräften ein deutlich abweichendes Bild.

Für Deutschland zeigt sich in 2023, dass nur etwa jedes achte Grundschulkind (12.7 %) von einer Sachunterrichtslehrkraft unterrichtet wird, die nach eigenen Angaben in den letzten zwei Jahren an einer Fortbildung zur *Integration von Informationstechnologien* teilgenommen hat (siehe Abbildung 5.4). Damit liegt Deutschland im internationalen Vergleich im unteren Mittelfeld. Im Vergleich zu 2019 sind für Deutschland keine signifikanten Veränderungen in den Teilnahmequoten festzustellen. Signifikant höhere Angaben für die Teilnahme an Fortbildungen zu digitalen Medien für 2023 sind für den internationalen Mittelwert (33.8 %), die Vergleichsgruppen der OECD (22.9 %) und der EU (24.3 %) zu verzeichnen. Insgesamt zeigen sich in TIMSS 2023 in zehn Staaten signifikant höhere Anteile als in 2019 in Bezug auf die Fortbildungsteilnahme zur *Integration von Informationstechnologien.*

Abbildung 5.3: Fortbildungsteilnahme der Mathematiklehrkräfte zur *Integration von Informationstechnologien* in Deutschland im internationalen Vergleich von 2019 und 2023 (Anteile der Viertklässler:innen mit entsprechenden Fortbildungen der Mathematiklehrkräfte in Prozent)

Teilnehmer	2019 %	*(SE)*	2023 %	*(SE)*	Differenz A $\%_{23} - \%_{19}$	*(SE)*	
Iran	24.2	(3.5)	52.5	(3.0)	28.3	(4.6)	▲
Taiwan	40.5	(3.8)	65.9	(2.6)	25.4	(4.6)	■
2 *Litauen*	55.3	(3.5)	79.5	(2.3)	24.2	(4.2)	■
2 *Georgien*	57.2	(3.6)	79.0	(3.0)	21.8	(4.7)	■
Republik Korea (Südkorea)	14.7	(2.9)	35.6	(3.8)	20.8	(4.8)	■
Oman	41.8	(3.3)	62.4	(2.6)	20.6	(4.2)	■
Marokko	17.9	(3.2)	37.6	(2.6)	19.6	(4.1)	■
2 *Armenien*	48.4	(4.1)	67.2	(3.2)	18.9	(5.2)	■
23 *Albanien*	31.9	(4.0)	50.6	(6.1)	18.7	(7.2)	■
2 *Polen*	62.6	(3.6)	81.1	(2.9)	18.5	(4.6)	■
2 *Serbien*	27.3	(4.0)	45.4	(4.2)	18.1	(5.9)	■
123 *Türkei (5. Jgst.)*	32.4	(3.5)	49.2	(4.4)	16.8	(5.6)	■
23 *Belgien (Fläm. Gem.)*	7.8	(1.9)	24.6	(3.3)	16.8	(3.8)	■
Deutschland	8.4	(2.0)	24.0	(2.5)	15.7	(3.2)	■
Portugal	24.6	(3.1)	40.2	(3.0)	15.6	(4.3)	■
2 *Montenegro*	14.2	(1.9)	29.3	(2.3)	15.1	(3.0)	■
23 *Chile*	17.2	(3.4)	31.6	(3.7)	14.4	(5.0)	■
2 *Spanien*	28.2	(3.3)	41.8	(2.3)	13.6	(4.0)	■
Slowakei	36.6	(3.0)	47.5	(3.3)	10.9	(4.4)	■
Bulgarien	26.0	(3.5)	36.8	(3.6)	10.8	(5.0)	■
Japan	11.3	(2.6)	21.3	(3.4)	9.9	(4.2)	■
3 *Niederlande*	4.3	(1.5)	13.5	(3.0)	9.2	(3.3)	■
2 *Italien*	33.0	(3.0)	42.0	(3.3)	9.0	(4.5)	■
2 *Tschechien*	31.1	(3.4)	39.7	(2.4)	8.6	(4.2)	■
Aserbaidschan	33.5	(3.7)	42.1	(3.1)	8.6	(4.8)	■
Internationaler Mittelwert	**33.8**	**(3.2)**	**41.7**	**(3.2)**	**8.0**	**(4.5)**	■
VG EU	**26.6**	**(3.0)**	**33.8**	**(3.0)**	**7.2**	**(4.2)**	■
VG OECD	**26.4**	**(3.0)**	**33.4**	**(3.1)**	**7.1**	**(4.3)**	■
Nordmazedonien	21.6	(3.8)	28.3	(3.6)	6.7	(5.3)	■
123 *England*	17.5	(4.4)	24.0	(4.0)	6.5	(6.0)	■
2 *Kasachstan*	62.2	(3.7)	68.0	(3.0)	5.8	(4.8)	■
Vereinigte Arabische Emirate	69.7	(1.3)	76.9	(1.2)	7.2	(1.8)	▼
Irland	23.4	(3.2)	26.9	(3.4)	3.5	(4.6)	▼
2 *Singapur*	61.1	(2.3)	64.6	(2.2)	3.5	(3.2)	▼
2 *Bosnien und Herzegowina*	12.2	(2.1)	15.3	(3.3)	3.2	(3.9)	▼
2 *Zypern*	25.8	(3.0)	28.3	(2.8)	2.5	(4.1)	▼
Kuwait	70.1	(4.0)	72.1	(3.7)	2.0	(5.4)	▼
2 *Ungarn*	12.0	(2.0)	14.0	(2.8)	2.0	(3.4)	▼
23 *USA*	47.0	(2.4)	48.5	(2.0)	1.5	(3.1)	▼
2 *Saudi-Arabien*	52.5	(3.4)	53.8	(3.1)	1.3	(4.6)	▼
Katar	70.0	(3.5)	71.2	(2.2)	1.2	(4.2)	▼
123 *Neuseeland*	25.7	(2.9)	25.4	(3.7)	-0.3	(4.7)	▼
Lettland	48.3	(3.7)	47.9	(3.8)	-0.4	(5.3)	▼
2 *Kosovo*	16.3	(3.3)	14.9	(2.8)	-1.4	(4.3)	▼
2 *Frankreich*	12.5	(2.4)	11.0	(2.1)	-1.6	(3.2)	▼
1 *Südafrika (5. Jgst.)*	46.3	(3.9)	44.5	(2.6)	-1.8	(4.7)	▼
3 *Hongkong*	76.5	(3.9)	73.8	(4.7)	-2.7	(6.1)	▼
23 *Kanada*	37.2	(3.0)	33.5	(1.4)	-3.7	(3.3)	▼
12 *Norwegen (5. Jgst.)*	17.6	(3.8)	13.5	(2.9)	-4.1	(4.8)	▼
Finnland	15.1	(2.3)	10.5	(2.0)	-4.6	(3.0)	▼
23 *Dänemark*	20.0	(3.1)	15.1	(2.8)	-5.0	(4.2)	▼
Australien	29.8	(3.3)	23.4	(2.8)	-6.4	(4.3)	▼
Bahrain	72.0	(3.1)	64.6	(3.1)	-7.4	(4.4)	▼
2 *Schweden*	29.4	(4.1)	17.5	(3.9)	-12	(5.6)	▼
Benchmark-Teilnehmer							
Abu Dhabi, VAE	58.2	(2.7)	74.3	(2.4)	16.1	(3.6)	■
23 *Québec, Kanada*	22.3	(4.0)	23.8	(2.8)	1.5	(4.9)	▼
Dubai, VAE	77.1	(1.8)	73.7	(2.2)	-3.4	(2.8)	▼
2 *Ontario, Kanada*	48.6	(5.5)	37.3	(2.0)	-11	(5.9)	▼

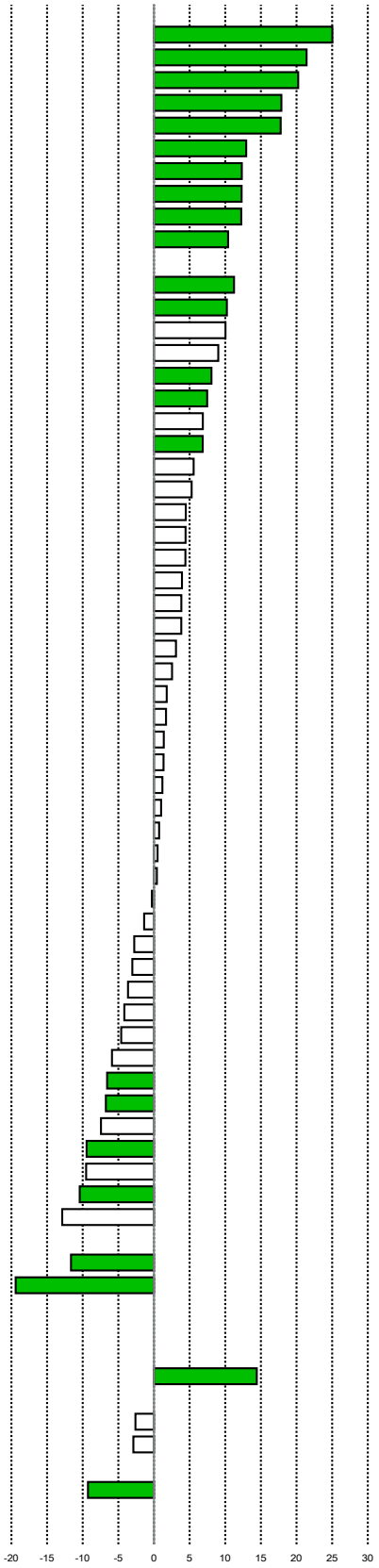


□ Keine statistisch signifikante Veränderung zwischen 2019 und 2023 ($p > .05$)
■ Statistisch signifikante Veränderung zwischen 2019 und 2023 ($p \leq .05$)
▲ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Abbildung 5.4: Fortbildungsteilnahme der Sachunterrichtslehrkräfte zur *Integration von Informationstechnologien* in Deutschland im internationalen Vergleich von 2019 und 2023 (Anteile der Schüler:innen mit entsprechenden Fortbildungen der Sachunterrichtslehrkräfte in Prozent)

Teilnehmer	2019 %	*(SE)*	2023 %	*(SE)*	Differenz[A] $\%_{23} - \%_{19}$	*(SE)*	
[23]*Albanien*	25.6	(3.9)	50.6	(5.2)	25.0	(6.5)	▲
Iran	26.5	(2.9)	47.9	(2.9)	21.4	(4.1)	▲
[2] *Polen*	55.7	(3.8)	75.9	(3.4)	20.2	(5.1)	▲
[2] *Litauen*	34.8	(3.5)	52.6	(3.0)	17.9	(4.6)	▲
[2] *Serbien*	25.9	(3.6)	43.7	(3.9)	17.8	(5.3)	▲
Aserbaidschan	29.8	(3.3)	42.8	(3.4)	12.9	(4.7)	▲
Bulgarien	18.6	(2.8)	30.9	(3.6)	12.3	(4.6)	▲
[2] *Spanien*	28.8	(3.3)	41.1	(2.4)	12.3	(4.1)	▲
Oman	37.2	(2.9)	49.4	(2.7)	12.2	(3.9)	▲
[2] *Tschechien*	13.9	(2.3)	24.3	(2.5)	10.4	(3.4)	▲
[123]*Türkei (5. Jgst.)*	31.5	(3.8)	42.7	(4.0)	11.2	(5.5)	■
Nordmazedonien	13.7	(2.8)	23.9	(3.2)	10.2	(4.2)	■
[2] *Zypern*	23.2	(4.2)	33.2	(3.2)	10.0	(5.3)	■
[2] *Armenien*	41.1	(4.4)	50.1	(4.8)	9.0	(6.6)	■
[23]*Belgien (Fläm. Gem.)*	10.0	(2.3)	18.0	(3.2)	8.1	(4.0)	■
[2] *Montenegro*	15.1	(1.9)	22.6	(2.2)	7.5	(2.9)	■
Taiwan	53.3	(4.0)	60.2	(2.9)	6.9	(5.0)	■
Portugal	15.1	(2.4)	22.0	(2.3)	6.8	(3.4)	■
Republik Korea (Südkorea)	27.4	(3.7)	33.0	(3.7)	5.6	(5.2)	■
Katar	64.9	(3.4)	70.2	(2.1)	5.3	(4.0)	■
[2] *Singapur*	57.7	(2.5)	62.2	(2.2)	4.5	(3.3)	■
Vereinigte Arabische Emirate	69.0	(2.2)	73.4	(0.9)	4.4	(2.4)	■
[2] *Kosovo*	9.3	(2.4)	13.8	(3.2)	4.4	(4.0)	■
Slowakei	35.3	(2.7)	39.2	(3.5)	3.9	(4.5)	■
[2] *Italien*	19.2	(3.1)	23.1	(2.8)	3.8	(4.2)	■
[123]*England*	17.3	(6.4)	21.1	(3.5)	3.8	(7.3)	■
Internationaler Mittelwert	**30.7**	**(3.3)**	**33.8**	**(3.0)**	**3.1**	**(4.5)**	■
VG EU	**21.8**	**(3.2)**	**24.3**	**(2.7)**	**2.5**	**(4.2)**	■
[3] *Bosnien und Herzegowina*	10.3	(1.7)	12.1	(2.9)	1.8	(3.4)	■
Japan	10.7	(2.4)	12.4	(2.9)	1.7	(3.8)	■
[2] *Georgien*	50.3	(4.4)	51.6	(3.3)	1.4	(5.5)	■
[2] *Saudi-Arabien*	50.6	(3.6)	52.0	(3.4)	1.3	(5.0)	■
[2] *Kasachstan*	56.8	(4.2)	58.0	(3.3)	1.2	(5.3)	■
Kuwait	75.0	(3.4)	76.0	(3.9)	1.0	(5.1)	■
VG OECD	**22.2**	**(3.3)**	**22.9**	**(2.7)**	**0.7**	**(4.3)**	■
Marokko	22.6	(3.1)	23.1	(2.6)	0.5	(4.0)	■
[23]*USA*	28.9	(2.2)	29.3	(2.2)	0.4	(3.1)	■
Deutschland	13.0	(2.7)	12.7	(2.2)	-0.3	(3.5)	■
Lettland	44.3	(3.9)	42.9	(3.9)	-1.4	(5.5)	■
[23]*Kanada*	18.0	(1.7)	15.2	(1.1)	-2.8	(2.0)	■
[23]*Chile*	16.3	(3.2)	13.3	(2.4)	-3.0	(4.0)	■
[2] *Ungarn*	8.5	(2.5)	4.8	(1.7)	-3.6	(3.0)	■
[12] *Norwegen (5. Jgst.)*	7.8	(2.6)	3.6	(1.5)	-4.2	(3.0)	■
[123]*Neuseeland*	19.8	(2.5)	15.2	(3.3)	-4.6	(4.2)	■
[1] *Südafrika (5. Jgst.)*	59.4	(3.5)	53.5	(3.0)	-5.9	(4.6)	■
[2] *Frankreich*	9.8	(2.5)	3.3	(1.4)	-6.6	(2.8)	■
Finnland	10.9	(2.1)	4.2	(1.5)	-6.8	(2.6)	■
[2] *Schweden*	18.9	(3.4)	11.5	(2.7)	-7.4	(4.4)	■
Bahrain	80.0	(2.3)	70.6	(2.9)	-9.5	(3.7)	■
[3]*Niederlande*	16.1	(4.6)	6.6	(2.3)	-9.5	(5.2)	■
[23]*Dänemark*	16.2	(3.9)	5.8	(2.2)	-10	(4.5)	■
[3]*Hongkong*	65.8	(4.9)	52.9	(5.4)	-13	(7.3)	■
Irland	21.7	(3.3)	10.1	(1.9)	-12	(3.8)	▼
Australien	34.5	(3.8)	15.1	(1.7)	-19	(4.2)	▼
Benchmark-Teilnehmer							
Abu Dhabi, VAE	55.0	(2.7)	69.5	(1.7)	14.4	(3.2)	▲
[23]*Québec, Kanada*	12.4	(3.0)	9.8	(2.0)	-2.6	(3.6)	■
[2] *Ontario, Kanada*	20.6	(3.3)	17.7	(1.8)	-2.9	(3.8)	■
Dubai, VAE	77.6	(2.0)	68.4	(1.9)	-9.3	(2.8)	▼



□ Keine statistisch signifikante Veränderung zwischen 2019 und 2023 ($p > .05$)
■ Statistisch signifikante Veränderung zwischen 2019 und 2023 ($p \leq .05$)
▲ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

 © TIMSS 2023

5.8 Zusammenfassung

Zielsetzung des vorliegenden Kapitels ist es, nicht nur die international eingesetzten Fragen zum Lehren und Lernen mit digitalen Medien in der Grundschule auszuwerten, sondern vor allem die Ergebnisse des umfassenderen Teilprojekts *TIMSS 2023 – Digitalisierung in der Grundschule* mit national ergänzten Fragestellungen vorzustellen und im Zusammenhang mit der Gesamtstudie TIMSS 2023 einzuordnen. So wurde untersucht, wie digitale Medien im Mathematik- und dem naturwissenschaftlichen Sachunterricht zum Zeitpunkt der TIMSS-2023-Erhebung an Grundschulen in Deutschland und im internationalen Vergleich genutzt wurden und welche Veränderungen in der Verfügbarkeit von häuslichen und schulischen digitalen Ressourcen im Trend seit 2007 bzw. im Vergleich zu TIMSS 2019 festzustellen sind. Die zentralen Erkenntnisse werden im Folgenden zusammengefasst und diskutiert.

Sowohl hinsichtlich der Ausstattung mit digitalen Medien im häuslichen als auch im schulischen Kontext konnte in TIMSS 2023 eine hohe Verfügbarkeit verzeichnet werden (siehe Abbildung 5.1 und Tabelle 5.1). Für die Ausstattung mit digitalen Medien im häuslichen Kontext zeigt sich, dass der Anteil der Viertklässler:innen in Deutschland, der in einem Haushalt lebt, für den die Erziehungsberechtigten angeben, über einen zuverlässigen Internetzugang zu verfügen, bei fast 100 Prozent liegt. Die Verfügbarkeit von digitalen Geräten wie Computer oder Laptop ist nur bei rund jedem 20. Kind nicht gegeben – ein Tablet oder einen E-Reader finden mehr als 85 Prozent der Grundschüler:innen zu Hause vor.

Übereinstimmend mit der KIM-2022-Studie (mpfs, 2022) dokumentieren die Befunde von TIMSS 2023 damit eine gute Verfügbarkeit digitaler Medien in den Familien in Deutschland. Dabei muss jedoch bedacht werden, dass die Angaben der Erziehungsberechtigten hinsichtlich der Verfügbarkeit digitaler Medien im häuslichen Kontext nicht explizit dahingehend ausgerichtet waren, ob *diesem* bzw. *jedem* Kind innerhalb der Familie digitale Geräte zur Verfügung stehen. Sowohl herkunftsbedingte Unterschiede im Zugang zu digitalen Medien als auch die gemeinsame Nutzung der digitalen Medien unter Geschwisterkindern ist somit zu berücksichtigen, kann mit den vorliegenden Daten allerdings nicht in den Blick genommen werden.

Auch wenn die Ausstattung im häuslichen Umfeld mit digitalen Medien im internationalen Vergleich in Teilen eher hoch ist, bleibt dennoch festzustellen, dass nach Angaben der in TIMSS 2023 befragten Kinder auch im Frühjahr 2023 noch rund 5 Prozent zu Hause über keinen Computer oder Laptop bzw. 13 Prozent nicht über ein Tablet oder einen E-Reader verfügen. Dies schränkt die Nutzung digitaler Medien für schulische Zwecke z. B. für die Bearbeitung von Hausaufgaben ein. Um allen Schüler:innen, auch im Sinne eines Abbaus von Bildungsbenachteiligungen, Zugang zu digitalen Medien zu ermöglichen, erscheint es bedeutsam, sicherzustellen, dass alle Schüler:innen entsprechende Medien auch zu Hause für schulische Belange nutzen können. Dies kann als notwendige Voraussetzung dafür betrachtet werden, fachliche Kompetenzen durch erweiterte Möglichkeiten des digitalgestützten Lehrens und Lernens im Fachunterricht und auch zu Hause zu fördern.

Die Befunde zeigen weiterhin, dass rund 40 Prozent der Viertklässler:innen in Deutschland ihre Kompetenzen im Umgang mit digitalen Medien als gering einschätzen. Auch wenn es sich um Selbsteinschätzungen und nicht um eine objektive Kompetenzmessung im Bereich der digitalen Kompetenzen handelt, weist dieses Ergebnis im internationalen Vergleich dennoch auf deutliche Handlungsbe-

darfe für Schulen hin, Kinder verstärkt auf einen selbstbestimmten und kritischen Umgang mit digitalen Medien vorzubereiten.

In der vertieften Betrachtung für Deutschland in diesem Kapitel wurde zudem deutlich, dass etwas mehr als ein Viertel der Grundschüler:innen im Unterricht kein Tablet bzw. Laptop nutzt, obwohl diese Geräte verfügbar wären (siehe Tabellen 5.2 und 5.3). Dieser Anteil erscheint auch vor dem Hintergrund der internationalen Befunde überraschend groß, wenn an die fortschreitenden Anforderungen an die kompetente Nutzung von digitalen Medien in der Gesellschaft gedacht und erwartet wird, auch Grundschüler:innen an digitale Medien heranzuführen (KMK, 2020). Diese Perspektive umfasst, dass auch schon Grundschulkinder auf eine kompetente Nutzung digitaler Medien zur Vorbereitung auf jetzige und zukünftige Herausforderungen in einer sich wandelnden Gesellschaft vorbereitet werden sollten. Besonders bedenklich ist, dass mehr als ein Drittel der Viertklässler:innen von Lehrkräften unterrichtet wird, denen für den konkreten Mathematikunterricht keine digitalen Medien zur Verfügung stehen (Tabelle 5.2). Dieser Anteil liegt für den naturwissenschaftlichen Sachunterricht bei 17.4 Prozent (Tabelle 5.3). Die hier vorgelegte Bestandsaufnahme macht deutlich, dass im Grundschulunterricht in Deutschland digitale Medien im weiteren Umfang zur Verfügung gestellt und dann auch stärker einbezogen werden könnten, um allen Grundschüler:innen entsprechende Lerngelegenheiten mit digitalen Medien zu bieten.

Diese Befunde überraschen gerade angesichts der Annahme, dass auch in Grundschulen in Deutschland die COVID-19-Pandemie zu einem ‚Digitalisierungsschub' (siehe auch Lorenz et al., 2023) geführt haben sollte. So, wie in diesem Zusammenhang in Bezug auf die PISA-Studie festgestellt wurde, dass die digitalisierungsbezogenen Entwicklungen, abgesehen von der Ausweitung der IT-Ausstattung, eher nur von kurzfristiger Dauer waren und mittlerweile nicht mehr weiterwachsen (Lewalter et al., 2023), scheint dieses Erklärungsmuster in Ansätzen auch für die Grundschulen in Deutschland zuzutreffen. Hier zeigt sich die Diskrepanz zwischen verfügbaren Geräten und deren Nutzung im Unterricht. Dennoch kann in Bezug auf die IT-Ausstattung im schulischen Kontext positiv hervorgehoben werden, dass es als Folge des DigitalPakt Schule angesehen werden kann, dass es zu einer erheblichen Ausweitung der schulischen IT-Ausstattung gekommen ist. Positiv zu verzeichnen ist zusätzlich, dass etwas mehr als 57 Prozent der Schüler:innen von Lehrkräften unterrichtet wurden, denen es gelang, in der Folge der COVID-19-Pandemie die Anpassung des Unterrichts an die herausfordernden Bedingungen in weniger als vier Monaten zu realisieren.

Hinsichtlich der Nutzungshäufigkeit digitaler Medien in Lehr- und Lernkontexten ist festzustellen, dass lediglich der Aspekt, dass digitale Medien im häuslichen und schulischen Kontext zur Verfügung stehen, nicht bedeutet, dass Kinder auch den Umgang mit digitalen Medien lernen oder Lehrkräfte wissen, wie diese didaktisch im Fachunterricht zur Lehrunterstützung einzusetzen sind. Dies spiegelt sich auch in den Ergebnissen von ICILS 2023 für die Sekundarstufe I sowie im Hinblick auf die tatsächlich gemessenen digitalen Kompetenzen wider (Eickelmann, Bos et al., 2024). Hinsichtlich des Einsatzes und der Nutzungshäufigkeit digitaler Medien in Deutschland im internationalen Vergleich kann trotz der Verfügbarkeit digitaler Geräte sowohl für den Mathematikunterricht als auch für den naturwissenschaftlichen Sachunterricht ein unterdurchschnittlicher Einsatz verzeichnet werden. Dabei stellt die unterrichtliche Nutzung digitaler Medien, um Lehrbücher zu lesen bzw. Lernvideos anzusehen, bei der Betrachtung der beiden Domänen eine Ausnahme dar, bei der die Angaben der Lehr-

kräfte vergleichsweise hohe Werte abbilden. Auch bei der national vertieften Betrachtung möglicher Einsatzoptionen digitaler Medien im Fachunterricht zeigt sich, dass der Einsatz digitaler Medien im Fachunterricht in Deutschland noch deutlich ausbaufähig ist. Bei der näheren Betrachtung möglicher Hindernisse, die Lehrkräfte davon abhalten, digitale Medien im Fachunterricht zu nutzen, wird sowohl von Mathematiklehrkräften als auch von Sachunterrichtslehrkräften angegeben, dass ein unzureichender Zugang zu digitalen Medien am ehesten ein Hindernis für den Einsatz digitaler Medien darstellt.

Mit Blick auf die von den Schüler:innen erreichten Kompetenzstände in Mathematik und den Naturwissenschaften (siehe Tabellen 5.6 und 5.7) ist festzustellen, dass – neben der generellen Ausstattung der Haushalte mit bildungsrelevanten Gütern – auch verschiedene in den Haushalten verfügbare digitale Medien mit den Kompetenzen in einem positiven Zusammenhang stehen. Dass die in den Schulen verfügbaren digitalen Medien, deren Nutzung sowie potenzielle Hindernisse für die Nutzung im Unterricht in keinem Zusammenhang mit den Kompetenzständen stehen, sollte nicht als generelle Unwirksamkeit des Einsatzes digitaler Medien für das fachliche Lernen im Unterricht fehlinterpretiert werden. Stattdessen kann angenommen werden, dass die fehlenden Zusammenhänge möglicherweise in einer nicht angemessenen didaktisch sinnvollen Einbettung in den Unterricht begründet sein könnten. Es erscheint zielführend, an diese explorativen Befunde vertiefende Analysen anzuschließen. Es liegt jedoch nahe, dass eine lernförderliche Nutzung digitaler Medien im Fachunterricht verstärkt durch Professionalisierungsmaßnahmen wie Lehrkräftekooperationen oder Fort- und Weiterbildungen unterstützt werden sollte. In diesem Kontext werden aktuell im vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Kompetenzverbund lernen:digital im Kompetenzzentrum MINT forschungsbasiert Angebote für die Gestaltung digital-gestützten Unterrichts u. a. im Fach Sachunterricht entwickelt und erprobt. Auch eine feste Verankerung in der ersten und zweiten Phase der Lehrkräfteausbildung sowie die Überprüfung von Unterrichtskonzepten auf ihre Wirksamkeit hin erscheint zielführend. Dabei sollte aber die Frage des *Wie* stärker betont werden als die einfache Betrachtung des *Wie oft*.

Hinsichtlich der Teilnahme an digitalisierungsbezogenen Fortbildungen von Lehrkräften ist für Mathematik ein Zuwachs zu verzeichnen. Im Vergleich zu 2019 ist in TIMSS 2023 der Anteil der Mathematiklehrkräfte, die an Fortbildungen zur *Integration von Informationstechnologien* teilgenommen haben, signifikant höher ausgefallen. Es könnte angenommen werden, dass die erhöhte Fortbildungsteilnahme perspektivisch dazu führt, dass mehr Lehrkräfte digitale Medien in ihrem Mathematikunterricht einsetzen. Für den naturwissenschaftlichen Sachunterricht zeigen die Befunde nur eine geringe Fortbildungsteilnahme der Lehrkräfte. Diesbezüglich könnten vertiefende Analysen Schlüsse darüber zulassen, inwiefern dies mit der geringen Nutzungshäufigkeit digitaler Medien im Sachunterricht im Zusammenhang steht, um zumindest mittelfristig eine höhere Fortbildungsteilnahme und so eine breitere Nutzung digitaler Medien im Sachunterricht zu erreichen.

Vor dem Hintergrund der hier vorgelegten Bestandsaufnahme für Grundschulen in Deutschland im internationalen Vergleich im Jahr 2023 ist in den Blick zu nehmen, wie es den Lehrkräften, aber auch Schulen und dem Bildungssystem in Deutschland insgesamt gelingen kann, digitale Medien gewinnbringend für den Unterricht und damit zur Kompetenzförderung für die Schüler:innen zu nutzen.

Vertiefte Analysen von Staaten, die digitale Medien bereits umfangreicher und vielfältiger einsetzen und insgesamt ein höheres fachliches Kompetenzniveau

ihrer Schüler:innen erreichen, könnten hier Ansatzpunkte für die Weiterentwicklung der Unterrichtsgestaltung in Deutschland mit und durch digitale Medien bieten.

Literatur

Blömeke, S., Bremerich-Vos, A., Haudeck, H., Kaiser, G., Nold, G., Schwippert, K., & Willenberg, H. (Hrsg.). (2011). *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT*. Waxmann.

BMBF (Bundesministerium für Bildung und Forschung). (2016). *Bildungsoffensive für die digitale Wissensgesellschaft. Strategie des Bundesministeriums für Bildung und Forschung*. https://www.kmk.org/fileadmin/pdf/Themen/Digitale-Welt/Bildungsoffensive_fuer_die_digitale_Wissensgesellschaft.pdf

Caruso, C., & Bruns, M. (2021). (Medienbezogene) Lerngelegenheiten und Kompetenzbedarfe im Vorbereitungsdienst. Explorative Perspektiven von Lehrkräften angesichts der veränderten Rahmenbedingungen durch die Corona-Pandemie. In C. Reintjes, R. Porsch & G. im Brahm (Hrsg.), *Das Bildungssystem in Zeiten der Krise. Empirische Befunde, Konsequenzen und Potenziale für das Lehren und Lernen* (S. 239–257). Waxmann. https://doi.org/10.31244/9783830993629

Drossel, K., Eickelmann, B., Schaumburg, H., & Labusch, A. (2019). Nutzung digitaler Medien und Prädiktoren aus der Perspektive der Lehrerinnen und Lehrer im internationalen Vergleich. In B. Eickelmann, W. Bos, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil & J. Vahrenhold (Hrsg.), *ICILS 2018 #Deutschland. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking* (S. 205–240). Waxmann. https://doi.org/10.25656/01:18166

Drossel, K., Wendt, H., Schmitz, S., & Eickelmann, B. (2012). Merkmale der Lehr- und Lernbedingungen im Primarbereich. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 171–202). Waxmann.

Eickelmann, B., Bos, W., Gerick, J., & Fröhlich, N. (2024). Computer- und informationsbezogene Kompetenzen von Schüler*innen der 8. Jahrgangsstufe in Deutschland im dritten internationalen Vergleich. In B. Eickelmann, N. Fröhlich, W. Bos, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil, & J. Vahrenhold (Hrsg.), *ICILS 2023 #Deutschland. Computer- und informationsbezogene Kompetenzen und Kompetenzen im Bereich Computational Thinking von Schüler*innen im internationalen Vergleich* (S. 47–72). Waxmann. https://doi.org/10.31244/9783830999492.02

Eickelmann, B., Fröhlich, N., Bos, W., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., Senkbeil, M., & Vahrenhold, J. (Hrsg.). (2024*). ICILS 2023 #Deutschland. Computer- und informationsbezogene Kompetenzen und Kompetenzen im Bereich Computational Thinking von Schüler*innen im internationalen Vergleich*. Waxmann. https://doi.org/10.31244/9783830999492

Eickelmann, B., & Drossel, K. (2020). *Schule auf Distanz. Perspektiven und Empfehlungen für den Schulalltag. Eine repräsentative Befragung von Lehrkräften in Deutschland*. Vodafone Stiftung Deutschland. https://www.vodafone-stiftung.de/wp-content/uploads/2020/05/Vodafone-Stiftung-Deutschland_Studie_Schule_auf_Distanz.pdf

Eickelmann, B., & Gerick, J. (2020). Lernen mit digitalen Medien. Zielsetzungen in Zeiten von Corona und unter besonderer Berücksichtigung von sozialen Ungleichheiten. In D. Fickermann & B. Edelstein (Hrsg.), *„Langsam vermisse ich die Schule …". Schule während und nach der Corona-Pandemie* (S. 153–162). Waxmann. https://doi.org/10.31244/9783830992318.09

Eickelmann, B., & Vennemann, M. (2014). Nutzung digitaler Medien im naturwissenschaftlichen Unterricht der Grundschule. In B. Eickelmann, R. Lorenz, M. Vennemann, J. Gerick & W. Bos (Hrsg.), *Grundschule in der digitalen Gesellschaft. Befunde aus den Schulleistungsstudien IGLU und TIMSS 2011* (S. 73–84). Waxmann.

Eickelmann, B., Bos, W., & Labusch, A. (2019). Die Studie ICILS 2018 im Überblick – Zentrale Ergebnisse und mögliche Entwicklungsperspektiven. In B. Eickelmann, W. Bos, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil & J. Vahrenhold (Hrsg.), *ICILS 2018 #Deutschland. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking* (S. 7–31). Waxmann. https://doi.org/10.25656/01:18166

Endberg, M., Engec, L.-I., & van Ackeren, I. (2021). ‚Optimierung' durch Fortbildung und Unterstützung für Schulen?! Modellvorschlag zu Unterstützungsleistungen für Schulentwicklung im Kontext der Digitalisierung und erste Ergebnisse des Projekts ForU-SE-digi für Nordrhein-Westfalen. *Zeitschrift für Theorie und Praxis der Medienbildung*, *42*, 108–133. https://doi.org/10.21240/mpaed/42/2021.04.07.X

Erikson, R., & Goldthorpe, J. H. (2002). Intergenerational inequality: A sociological perspective. *Journal of Economic Perspectives*, *16*(3), 31–44. https://doi.org/10.1257/089533002760278695

Erikson, R., Goldthorpe, J. H., & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology*, *30*(4), 415–441. https://doi.org/10.2307/589632

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. https://doi.org/10.1016/0049-089X(92)90017-B

Gerick, J., Annemann, C., Niemann, T., & Drossel, K. (2024). Digitalisierungsbezogene Lehrkräftefortbildungen – Analysen zu Zusammenhängen mit Lehrpersonen- und Schulmerkmalen sowie zum wahrgenommenen Fortbildungserfolg durch Lehrkräfte in Deutschland. *Zeitschrift für Erziehungswissenschaft*, *27*, 661–683. https://doi.org/10.1007/s11618-024-01225-8

Gerick, J., Massek, C., Eickelmann, B., & Labusch, A. (2019). Computer- und informationsbezogene Kompetenzen von Mädchen und Jungen im zweiten internationalen Vergleich. In B. Eickelmann, W. Bos, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil & J. Vahrenhold (Hrsg.), *ICILS 2018 #Deutschland. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking* (S. 271–300). Waxmann. https://doi.org/10.25656/01:18166

Helmke, A. (2010). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts* (3. Aufl.). Klett-Kallmeyer.

Herzig, B., & Grafe, S. (2010). Digitale Medien in Schule und Alltagswelt. Zur Verbindung von formellen und informellen Lernprozessen. In B. Bachmair (Hrsg.), *Medienbildung in neuen Kulturräumen. Die deutschprachige und britische Diskussion* (S. 183–195). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92133-4

Herzig, B. (2004). Medienpädagogische Kompetenz. In S. Blömeke, P. Reinhold., G. Tulodziecki & J. Wildt (Hrsg.), *Handbuch Lehrerbildung* (S. 578–594). Klinkhardt.

ILO (International Labour Office). (2012). *International Standard Classification of Occupations. Structure, group definitions and correspondence tables*. https://www.ilo.org/media/352556/download

Irion, T., & Peschel, M. (2016). Grundschule und neue Medien – Neue Entwicklungen. In M. Peschel & T. Irion (Hrsg.), *Neue Medien in der Grundschule 2.0. Grundlagen – Konzepte – Perspektiven*. (S. 11–15). Grundschulverband e. V.

Kasper, D., Wendt, H., Schwippert, K., & Köller, O. (2020). Trends in Schülerzusammensetzungen und mathematischen und naturwissenschaftlichen Kompetenzen. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 331–349). Waxmann. https://doi.org/10.31244/9783830993193

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2016). *Bildung in der digitalen Welt. Strategie der Kultusministerkonferenz*. https://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2017/Strategie_neu_2017_datum_1.pdf

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2020). *Ländervereinbarung über die gemeinsame Grundstruktur des Schul-*

wesens und die gesamtstaatliche Verantwortung der Länder in zentralen bildungspolitischen Fragen. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2020/2020_10_15-Laendervereinbarung.pdf

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2021). *Lehren und Lernen in der digitalen Welt. Die ergänzende Empfehlung zur Strategie zur „Bildung in der digitalen Welt"*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2021/2021_12_09-Lehren-und-Lernen-Digi.pdf

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2023). *Jahresbericht der Kultusministerkonferenz zur Bildung in der digitalen Welt. (Beschluss der KMK vom 07./08.12.2023)* https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2023/2023_12_07-Jahresbericht-Bildung-in-der-digitalen-Welt_2022-2023.pdf

Kunkel, C., & Peschel, M. (2020). Lernen mit und über digitale Medien im Sachunterricht. Entwicklung eines vielperspektivischen Konzepts zur Erschliessung digitaler Medien. In K. Rummler, I. Koppel, S. Aßmann, P. Bettinger & K. D. Wolf (Hrsg.), *Lernen mit und über Medien in einer digitalen Welt* (S. 455–476). https://doi.org/10.21240/mpaed/jb17/2020.05.18.X

Lachner, A., Scheiter, K., & Stürmer, K. (2020). Digitalisierung und Lernen mit digitalen Medien als Gegenstand der Lehrerinnen- und Lehrerbildung. In C. Cramer, J. König, M. Rothland & S. Blömeke (Hrsg.), *Handbuch Lehrerinnen- und Lehrerbildung* (S. 67–75). UTB. https://doi.org/10.35468/hblb2020-007

Lewalter, D., Kastorff, T., & Moser, S. (2023). Digitalisierungsbezogene Lerngelegenheiten und -aktivitäten in Schule und Freizeit. In D. Lewalter, J. Diedrich, F. Goldhammer, O. Köller & K. Reiss (Hrsg.), *PISA 2022. Analyse der Bildungsergebnisse in Deutschland* (S. 237–271). Waxmann. https://doi.org/10.31244/9783830998488

Lorenz, R. (2018). Ressourcen, Einstellungen und Lehrkraftbildung im Bereich Digitalisierung. In N. McElvany, F. Schwabe, W. Bos & H. G. Holtappels (Hrsg.), *Digitalisierung in der schulischen Bildung. Chancen und Herausforderungen* (S. 53–68). Waxmann.

Lorenz, R., Goldhammer, F., & Glondys, M. (2023). Digitalisierung in der Grundschule. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher & T. C. Stubbe (Hrsg.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (S. 197–214). Waxmann. https://doi.org/10.31244/9783830997009

McElvany, N., Lorenz, R., Frey, A., Goldhammer, F., Schilcher, A., & Stubbe, T. C. (Hrsg.). (2023). *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre*. Waxmann. https://doi.org/10.31244/9783830997009

Melzer, C., & Hillenbrand, C. (2015). Aufgabenprofile. Welche Aufgaben bewältigen sonderpädagogische Lehrkräfte in verschiedenen schulischen Tätigkeitsfeldern? *Zeitschrift für Heilpädagogik*, *66*(5), 230–242.

mpfs (Medienpädagogischer Forschungsverbund Südwest) (Hrsg.). (2022). *KIM-Studie 2022. Kindheit, Internet, Medien. Basisuntersuchung zum Medienumgang 6- bis 13-Jähriger*. https://www.mpfs.de/fileadmin/files/Studien/KIM/2022/KIM-Studie2022_website_final.pdf

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science*. TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2019/international-results/

Peschel, M. (2019). Milliarden für die Bildung. Grundlegende Forderungen an den Digitalpakt. *Die Grundschulzeitschrift*, *33*(318), 43.

Porsch, R., & Wendt, H. (2016). Aus- und Fortbildung von Mathematik- und Sachunterrichtslehrkräften. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert, D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 189–204). Waxmann. https://doi.org/10.25656/01:14022

Schaumburg, H. (2018). Empirische Befunde zur Wirksamkeit unterschiedlicher Konzepte des digital unterstützten Lernens. In N. McElvany, F. Schwabe, W. Bos & H. G. Holtappels (Hrsg.), *Digitalisierung in der schulischen Bildung. Chancen und Herausforderungen* (S. 27–40). Waxmann.

Scheiter, K. (2021). Lernen und Lehren mit digitalen Medien: Eine Standortbestimmung. *Zeitschrift für Erziehungswissenschaft*, *24*(5), 1039–1060. https://doi.org/10.1007/s11618-021-01047-y

Schwippert, K., Kasper, D., Köller, O., McElvany, N., Selter, C., Steffensky, M., & Wendt, H. (Hrsg.). (2020). *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Waxmann. https://doi.org/10.31244/9783830993193

Tamim, R. M., Bernard, R. M., Borokhovski, E., Abrami, P. C., & Schmid, R. F. (2011). What forty years of research says about the impact of technology on learning: A second-order meta-analysis and validation study. *Review of Educational Research*, *81*(1), 4–28. https://doi.org/10.3102/0034654310393361

UNESCO (United Nations Educational, Scientific and Cultural Organization). (2003). International Standard Classification of Education, ISCED 1997. In J. H. Hoffmeyer-Zlotnik & C. Wolf (Hrsg.), *Advances in cross-national comparison. A European working book for demographic and socio-economic variables* (S. 195–220). Plenum Press. https://doi.org/10.1007/978-1-4419-9186-7_10

Wendt, H., Kasper, D., Bos, W., Vennemann, M., & Goy, M. (2017). Wie viele Punkte auf der TIMSS-Metrik entsprechen einem Lernjahr? Leistungszuwächse in Mathematik und Naturwissenschaften am Ende der Grundschulzeit. In T. Eckert & B. Gniewosz (Hrsg.), *Bildungsgerechtigkeit* (S. 121–152). Springer. https://doi.org/10.1007/978-3-658-15003-7_8

Züchner, I., & Jäkel, H. R. (2021). Fernbeschulung während der COVID-19 bedingten Schulschließungen weiterführender Schulen: Analysen zum Gelingen aus Sicht von Schülerinnen und Schülern. *Zeitschrift für Erziehungswissenschaft*, *24*, 479–502. https://doi.org/10.1007/s11618-021-01006-7

Kapitel 6
Geschlechterunterschiede in mathematischen und naturwissenschaftlichen Kompetenzen

Sonja Nonte, Eva Grommé und Luise A. Scholz

6.1 Einleitung

Obwohl es viele Bemühungen gibt, Geschlechterdisparitäten in schulischen Leistungen abzubauen, sind in Deutschland weiterhin signifikante Geschlechterunterschiede in schulischen Leistungen zu verzeichnen (Stanat et al., 2023). Hinsichtlich mathematischer und naturwissenschaftlicher Kompetenzen konnte eine umfangreiche Metaanalyse zeigen, dass Mädchen jeweils nach wie vor seltener in der Gruppe der leistungsstärkeren Schüler:innen anzutreffen sind und auch ihr Interesse für Mathematik und Naturwissenschaften geringer ist als das von Jungen (Keller et al., 2022). In diesem Kapitel wird untersucht, ob und ggf. welche Geschlechterunterschiede in Bezug auf die mathematisch-naturwissenschaftlichen Leistungen bei Grundschulkindern vorliegen.

Als Novum im aktuellen Zyklus der *Trends in International Mathematics and Science Study* (TIMSS) lässt sich die Erfassung der Geschlechtsidentität benennen. Anstatt der Abfrage des männlichen oder weiblichen Geschlechts wie in den vorangegangenen Studienzyklen wurde in diesem Durchgang zum ersten Mal auch eine dritte Antwortoption gegeben: *anderes*. Ausgehend von einer vermeintlich natürlichen Dichotomie der Geschlechter gibt es theoretische Positionen, die eine Dekonstruktion der Binarität von ‚männlich' und ‚weiblich' ermöglichen (Kuster, 2019). Dabei gilt als eine zentrale Erkenntnis die Differenzierung zwischen dem biologischen Geschlecht (*sex*) und der gesellschaftlich-kulturell geprägten sozialen Geschlechtsidentität (*gender*). Selbst wenn es ausschließlich zwei biologische Geschlechter gäbe, die binär zueinanderstehen, muss das nicht heißen, dass sich daraus ausschließlich zwei Geschlechtsidentitäten ergeben, sodass sich die Geschlechtsidentität als Spektrum auffassen ließe (Nielsen et al., 2021). Auch bei Grundschulkindern kann es vorkommen, dass sich die Wahrnehmung des eigenen Geschlechts nicht auf eine der binären geschlechtsspezifischen Identitäten begrenzt (Pauletti et al., 2017). In manchen Fällen identifizieren

sich Kinder außerhalb oder innerhalb der binären Geschlechter und können dies benennen (Olson et al., 2015). Um den befragten Kindern bei TIMSS eine Angabe außerhalb der binären Geschlechterkategorien zu ermöglichen, wurde im Kontext der Erhebung in Deutschland die Option *anderes* zur Auswahl gestellt. Diese Alternative wurde von ca. 2 Prozent der Schüler:innen gewählt. Aufgrund dieser geringen Fallzahlen und der nicht vergleichbaren sowie national bezogenen Antwortoption dieser Kategorie können jedoch keine repräsentativen Aussagen für diese Schülergruppe getroffen werden. In den Analysen werden deshalb ausschließlich Befunde zu Schüler:innen präsentiert, die sich der Kategorie *Mädchen* oder *Junge* zugeordnet haben.

Bezogen auf die binären Geschlechter werden in diesem Kapitel zunächst theoretische Erklärungsansätze geschlechterbezogener Disparitäten in mathematischen und naturwissenschaftlichen Kompetenzen skizziert. Anschließend werden aktuelle Befunde aus Schulleistungsstudien differenziert nach den Domänen Mathematik und Naturwissenschaften vorgestellt, um im Anschluss daran die Befunde aus TIMSS 2023 zu berichten. Hier werden Geschlechterunterschiede nicht nur bei Viertklässler:innen in Deutschland, sondern insbesondere auch im internationalen Vergleich sowie im Hinblick auf unterschiedliche Inhalts- und Anforderungsbereiche betrachtet. Auch die Unterschiede in der Verteilung auf die Kompetenzstufen werden tiefergehend analysiert, um abschließend die Befunde unter Berücksichtigung des Theorie- und Forschungstandes zu diskutieren und Implikationen für das deutsche Bildungssystem zu formulieren.

6.2 Theoretische Erklärungsansätze zu Geschlechterunterschieden in mathematischen und naturwissenschaftlichen Leistungen

Mit Bezug auf die immer wieder berichteten geringeren Leistungen von Mädchen im mathematischen, aber auch naturwissenschaftlichen Bereich, wie sie u. a. in den Schulleistungsstudien TIMSS (Schwippert et al., 2020), dem Bildungstrend des *Instituts für Qualitätsentwicklung im Bildungswesen* (IQB; Stanat et al., 2022) und dem *Programme for International Student Assessment* (PISA; OECD, 2023) gefunden wurden, werden verschiedene theoretische Modelle herangezogen, um ein vertieftes Verständnis für die Gründe dieser Geschlechterunterschiede zu entwickeln. Ausführlich werden diese theoretischen Erklärungsansätze im Berichtsband zu TIMSS 2019 dargestellt (Nonte et al., 2020). Im vorliegenden Kapitel erfolgen Ergänzungen zu den dort skizzierten theoretischen Modellen (*Stereotype Threat* und Erwartungs-Wert-Modell).

Modelle, die sich heranziehen lassen, um signifikante Leistungsunterschiede zwischen Schüler:innen zu erklären, setzen auf der individuellen, aber auch gesellschaftlichen Ebene an. Nur wenige Ansätze beziehen sich dabei ausschließlich auf das Individuum, z. B. durch die Untersuchung biologischer Determinanten in Form von neurologischen Aktivitäten (Young et al., 2012). Verbreiteter sind Modelle, die gerade in der Interaktion von Individuum und Umwelt Erklärungsansätze für die letztlich auftretenden geschlechterbezogenen Leistungsunterschiede sehen, wie beispielsweise der *Stereotype Threat* (Fiske & Neuberg, 1990; Holder & Kessels, 2017; Steele & Aronson, 1995). Dieser besagt, dass internalisierte und pauschalisierende Vorannahmen im Kontext der Zugehörigkeit zu einer Gruppe, wie z. B. Mädchen, bei den Betroffenen zu Sorge darüber führen, ein Stereotyp zu erfüllen und/oder danach beurteilt zu werden. Somit hätten Mädchen, beispielsweise in Bezug auf Mathematik, die Sorge, in Relation zu Jun-

gen schlechter abzuschneiden oder bewertet zu werden (van Mier et al., 2018). Die Wahrnehmung des Stereotyps des mathematisch leistungsschwachen Mädchens wird durch äußere Vorannahmen geformt und kann somit das individuelle Erleben (Flore & Wicherts, 2015), darunter auch das Fähigkeitsselbstkonzept, beeinflussen, welches als wichtiger Prädiktor für die Erklärung von Geschlechterunterschieden in Bezug auf Leistung gilt (Heyder et al., 2021; Möller et al., 2020). McCoy et al. (2022) stellten fest, dass Mütter und Lehrkräfte die mathematischen Leistungen von Mädchen im Vergleich zu denen der Jungen unterschätzten, besonders, wenn die Mädchen leistungsstark waren, und auch, wenn die tatsächlichen Mathematikleistungen berücksichtigt wurden.

Mit den motivationalen Merkmalen von Schüler:innen und ihrem Einfluss auf schulische Leistungen befasst sich auch das Erwartungs-Wert-Modell (Eccles & Wigfield, 2002). Laut diesem kann das Fähigkeitsselbstkonzept, aber auch die intrinsische Motivation beeinflussen, inwiefern sich mit Mathematik oder den Naturwissenschaften auseinandergesetzt wird. Aktuelle Ergebnisse einer umfassenden Metaanalyse von mathematisch leistungsstarken Schüler:innen aus sechs PISA-Zyklen bestätigten, dass ein geringeres mathematisch-naturwissenschaftliches Interesse auch zu schlechteren schulischen Leistungen führen kann (Keller et al., 2022). Es zeigten sich außerdem im Bereich der Naturwissenschaften Geschlechterunterschiede, wobei trotz eines allgemein höheren Interesses der Jungen an naturwissenschaftlichen Themen unterschiedliche Interessensprofile identifiziert werden konnten. So berichteten Mädchen ein höheres Interesse an medizinisch-biologischen Inhalten und Jungen an Themen der Physik. Passend zu diesen Befunden konnte in aktuellen längsschnittlichen Untersuchungen im Panel-Design festgestellt werden, dass Mädchen nicht nur ein geringeres Interesse an Mathematik zeigen, sondern auch die Berufsorientierung hierdurch beeinflusst wird (Costache et al., 2022; Lazarides et al., 2020).

6.3 Befunde zu Geschlechterunterschieden in mathematischen und naturwissenschaftlichen Leistungen in Schulleistungsstudien

Erste Befunde zu Geschlechterunterschieden in mathematischen und naturwissenschaftlichen Kompetenzen im internationalen Vergleich unter Teilnahme Deutschlands wurden im Kontext der *First International Mathematics Study* (FIMS) im Jahr 1964 mit einem Fokus auf 13-jährige Schüler:innen berichtet (IEA, o. J.). Damals dokumentierte Keeves (1973), dass sich Mädchen und Jungen in allen Teilnehmerstaaten signifikant in ihren mathematischen Leistungen unterschieden. In Deutschland lag der Geschlechterunterschied mit 20 Differenzpunkten zugunsten der Jungen oberhalb des internationalen Durchschnitts von 12.5 Punkten. Wie die Kompetenzunterschiede zwischen Mädchen und Jungen derzeit in aktuellen Schulleistungsstudien ausfallen, wird in diesem Abschnitt erläutert. Dabei wird der Fokus auf den Primarbereich gelegt und nur dort, wo keine Daten für den Primarbereich vorliegen, werden aktuelle Befunde für die Sekundarstufe vorgestellt.

6.3.1 Geschlechterunterschiede in mathematischen Kompetenzen

Bereits zum dritten Mal wurden 2021 im Kontext des IQB-Bildungstrends (bis 2012: Ländervergleich) Kompetenzen in den Fächern Deutsch und Mathematik von Viertklässler:innen erfasst. Die Ergebnisse im Bereich Mathematik zeigen einen geringen Effekt für Geschlechterunterschiede von Viertklässler:innen zugunsten der Jungen von d = .23 (Cohens d; zur Einordnung siehe Nonte et al., 2020). Insgesamt nahmen die Geschlechterunterschiede über die Erhebungszyklen zu (d = .16 im Jahr 2011; d = .18 im Jahr 2016; Gentrup et al., 2022). Auch wird ersichtlich, dass Mädchen die Mindeststandards signifikant häufiger verfehlen als Jungen und auch seltener den Optimalstandard erreichen. Ein Vergleich der Testergebnisse der Jahre 2011 und 2021 zeigt, dass Mädchen bei den Erhebungen im Jahr 2021 durchschnittlich 43 Punkte weniger erreichten als noch im Jahr 2011. Dies entspricht einem mittleren Effekt (d = -.41). Jungen verloren ebenfalls, jedoch vergleichsweise weniger Punkte im zeitlichen Verlauf. Hier liegt die Differenz der Jahre 2011 und 2021 bei 33 Punkten (d = -.31).

Auch in TIMSS 2011, 2015 und 2019 konnte beobachtet werden, dass Jungen in Deutschland in den oberen mathematischen Kompetenzstufen stärker repräsentiert waren, Mädchen hingegen in den mittleren oder einer der beiden unteren Kompetenzstufen (Nonte et al., 2020). Auch auf der Gesamtskala Mathematik war der Unterschied signifikant: Die Kompetenzen der Mädchen lagen 10 Punkte unterhalb der Kompetenzwerte der Jungen, dies entsprach in etwa dem Durchschnittswert der EU. Eine differenzierte Betrachtung der Inhaltsbereiche in TIMSS 2019 zeigt für Deutschland Geschlechterunterschiede insbesondere im Bereich *Messen und Geometrie* mit einer Differenz von 14 Punkten zugunsten der Jungen. Im Bereich *Arithmetik* lag der Unterschied noch bei 11 Punkten und im Bereich *Daten* bei 7 Punkten. In allen Bereichen erreichten Jungen signifikant höhere Punktwerte als Mädchen. In den kognitiven Anforderungsbereichen in Mathematik zeigte sich ein ähnliches Bild. Auch hier erreichten Jungen durchschnittlich höhere Werte in den Bereichen *Reproduzieren*, *Anwenden* und *Problemlösen*. Am stärksten war der Geschlechterunterschied hier für den Bereich *Reproduzieren* (16 Punkte Differenz) ausgeprägt. Der über die Zyklen zu verzeichnende Leistungsabfall ist im Vergleich zu vorherigen Zyklen jedoch als gering zu bewerten und war weder bei den Mädchen noch bei den Jungen signifikant.

Für die Sekundarstufe konnte im Rahmen von PISA 2022 (Diedrich, Reinhold et al., 2023) festgestellt werden, dass sich die im Grundschulbereich festgestellten Geschlechterdisparitäten fortsetzen. Während Jungen in fast allen Teilnehmerstaaten höhere Mathematikkompetenzen erreichten als Mädchen, zeigen sich für Mädchen ungünstigere motivationale Voraussetzungen (z. B. Angst). Besonders interessant erscheint der Befund, dass Mädchen weniger Vertrauen äußern, sich durch Anstrengung in Mathematik verbessern zu können, obgleich sie gegenüber Jungen häufiger angeben, sich im Unterricht anzustrengen („statisches Selbstbild“; Diedrich, Patzl et al., 2023, S. 101f.).

6.3.2 Geschlechterunterschiede in naturwissenschaftlichen Kompetenzen

Für die naturwissenschaftlichen Kompetenzen konnten in TIMSS 2019 keine signifikanten Geschlechterunterschiede für Deutschland beobachtet werden. Eine differenzierte Betrachtung der Inhaltsbereiche zeigte jedoch im Inhaltsbereich *Biologie* einen signifikanten Vorsprung der Mädchen von insgesamt 7 Punkten gegenüber den Jungen. Im Inhaltsbereich *Geografie* wiederum lagen die Jungen mit 13 Differenzpunkten vor den Mädchen. Dabei zeigte sich in TIMSS 2019 im Trend keine signifikante Veränderung bei den Mädchen, jedoch bei den Jungen: Im Vergleich zu TIMSS 2011 verloren Jungen in TIMSS 2019 durchschnittlich 13 Punkte. Damit kann konstatiert werden, dass sich durch den signifikanten Abfall der Leistungen der Jungen in den Naturwissenschaften die Werte der Mädchen denen der Jungen in TIMSS 2019 anglichen. Insgesamt konnte zudem eine signifikante Abnahme der naturwissenschaftlichen Testleistungen in TIMSS 2019 im Vergleich zu TIMSS 2007 und TIMSS 2011 in der Gesamtstichprobe beobachtet werden (Nonte et al., 2020).

Naturwissenschaftliche Kompetenzen werden im IQB-Bildungstrend 2021 lediglich für die Sekundarstufe I berichtet (Stanat et al., 2022). In fast allen untersuchten Bereichen (mit Ausnahme von Physik) erreichten Mädchen hier signifikant höhere Testleistungen als Jungen (Gentrup et al., 2022). Im Kontext von PISA 2022 konnten wiederum keine signifikanten Geschlechterunterschiede in den naturwissenschaftlichen Kompetenzen sowohl für Deutschland als auch im OECD-Durchschnitt beobachtet werden (Kastorff et al., 2023).

Zusammenfassend kann für die mathematischen Kompetenzen von Mädchen im Kontext der Schulleistungsstudien der letzten Jahre beobachtet werden, dass diese signifikant hinter denen der Jungen zurückliegen. Zudem sind sie in den unteren Kompetenzstufen über- und in den höheren Kompetenzstufen unterrepräsentiert. Uneinheitlicher stellen sich die Befunde zu den naturwissenschaftlichen Kompetenzen von Viertklässler:innen dar. Hier erreichten Mädchen insbesondere im Inhaltsbereich *Biologie,* Jungen hingegen im Inhaltsbereich *Geografie* signifikant höhere Testwerte. In *Physik/Chemie* zeigte sich hingegen kein einheitliches Muster.

6.4 Ergebnisse zu geschlechterspezifischen Leistungsunterschieden in Mathematik

Im Folgenden werden die aktuellen Befunde aus TIMSS 2023 zu Unterschieden in mathematischen und naturwissenschaftlichen Leistungen von Mädchen und Jungen[1] vorgestellt. Die Ergebnisse werden zunächst im internationalen Vergleich präsentiert, anschließend folgen vertiefende nationale Befunde zur Verteilung auf den jeweiligen Kompetenzstufen, zu den fachspezifischen Einstellungen und Selbstkonzepten sowie Trendanalysen unter Berücksichtigung vergangener TIMSS-Zyklen (u. a. Nonte et al., 2020; Wendt et al., 2016).

1 Wie bereits berichtet basieren die Analysen lediglich auf den binären Geschlechterangaben Mädchen/Junge. Eine dritte Kategorie bleibt aufgrund der geringen Fallzahlen bei den Analysen unberücksichtigt.

6.4.1 Geschlechterspezifische Leistungsunterschiede auf der Gesamtskala Mathematik

Eine Einordnung der Testleistungen von Mädchen und Jungen im internationalen Vergleich ermöglicht die Darstellung in Abbildung 6.1. Dort werden in der linken Spalte zunächst alle Teilnehmer der Vergleichsgruppe von TIMSS 2023 aufgelistet. Daneben finden sich die prozentualen Anteile (%) von Mädchen und Jungen in der zugrundeliegenden Stichprobe mit dem zugehörigen Standardfehler (*SE*). Zuletzt wird in der fünften Spalte die Differenz in den erreichten Leistungspunkten (M_J-M_M) sowie der zugehörige Standardfehler berichtet. Diese deskriptiven Angaben werden durch eine Visualisierung in Form von waagerecht dargestellten Histogrammen in der rechten Spalte ergänzt. Signifikante Unterschiede werden in Form von grün eingefärbten Balken gekennzeichnet, wobei die Staaten entsprechend dieser Unterschiede absteigend in der Abbildung sortiert sind.

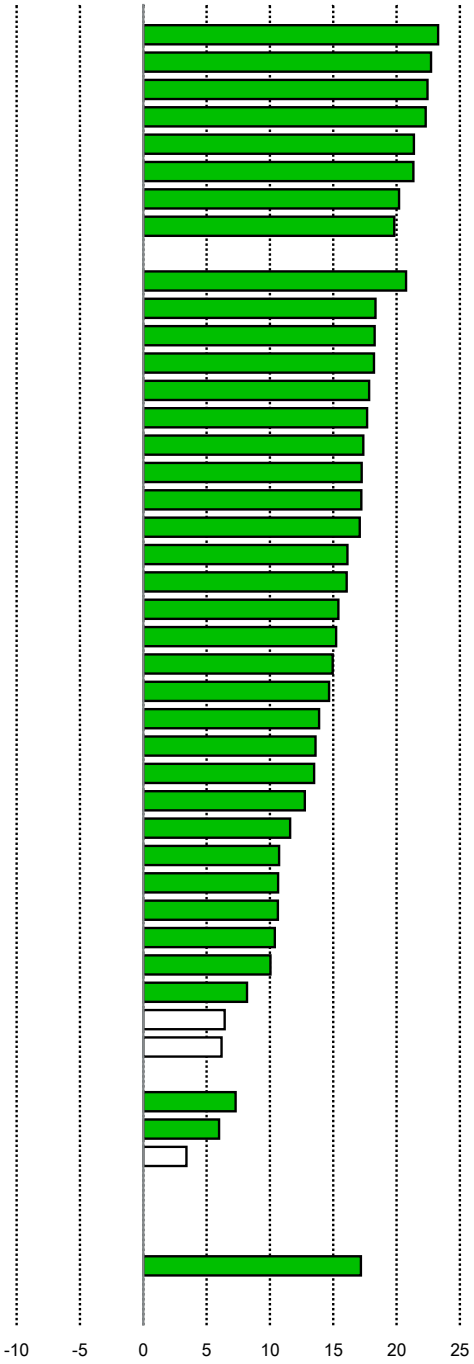
Anhand von Abbildung 6.1 wird deutlich, dass der Unterschied in den Testleistungen in Mathematik von Mädchen und Jungen in Deutschland bei 13 Punkten liegt. Dieser Unterschied ist signifikant zugunsten der Jungen und liegt nominell oberhalb der internationalen mittleren Differenz von 11 Punkten und unterhalb der Vergleichswerte für die EU (15 Punkte) und OECD (16 Punkte), unterscheidet sich von diesen aber nicht signifikant. Für fast alle Teilnehmerstaaten (inklusive Benchmark-Teilnehmer) kann beobachtet werden, dass Jungen auf der Gesamtskala Mathematik höhere Testleistungen erreichen als Mädchen (in Irland, Rumänien und Bulgarien sind diese Unterschiede jedoch nicht signifikant). Am stärksten ausgeprägt ist die Leistungsdifferenz auf der Gesamtskala Mathematik in Frankreich und Australien mit einem Unterschied von 23 Leistungspunkten zugunsten der Jungen. Daran schließen sich die Staaten Italien und Portugal mit 22 Differenzpunkten an. Tiefere Einblicke in die Geschlechterdisparitäten in Deutschland ermöglicht die differenzierte Betrachtung der Leistungsunterschiede nach Kompetenzstufen, wie sie in Abbildung 6.2 beschrieben sind.

Abgebildet sind hier die prozentualen Anteile (%) von Mädchen und Jungen auf den jeweiligen Kompetenzstufen sowie die im Mittel erreichten Leistungspunkte für Mathematik. Auf der rechten Seite findet sich die Differenz des prozentualen Anteils von Mädchen und Jungen auf den jeweiligen Stufen abgebildet. Anhand von Abbildung 6.2 wird ersichtlich, dass Jungen in den oberen zwei Kompetenzstufen (IV und V) mit 3.5 Prozent bzw. 3.7 Prozent häufiger, hingegen auf der zweiten Kompetenzstufe mit 4.6 Prozentpunkten signifikant seltener vertreten sind als Mädchen. Die Differenzen für die erste und dritte Kompetenzstufe sind nicht signifikant, deuten in der Tendenz jedoch ebenfalls auf einen geringeren Anteil von Jungen auf diesen Stufen hin. Im Vergleich zu TIMSS 2019 haben die Unterschiede hier noch einmal zugenommen und auch die Verteilung auf den oberen zwei Kompetenzstufen weist deutlich auf einen Geschlechterunterschied beim Erreichen höherer mathematischer Kompetenzen zu Lasten von Mädchen hin.

Nachdem Geschlechterunterschiede in den mathematischen Leistungen im internationalen Vergleich und im Hinblick auf die Kompetenzstufen für Deutschland beschrieben wurden, werden im Folgenden Einblicke in Unterschiede nach mathematischen Inhaltsbereichen im internationalen Vergleich dargestellt.

Abbildung 6.1: Testleistungen nach Geschlecht – Gesamtskala Mathematik

	Mädchen			Jungen			Differenz[A]		
Teilnehmer	%	M_M	*(SE)*	%	M_J	*(SE)*	$M_J - M_M$	*(SE)*	
2 *Frankreich*	50.5	473	(3.1)	49.5	496	(3.2)	23	(2.8)	▲
Australien	52.0	514	(2.1)	48.0	537	(2.5)	23	(2.6)	▲
2 *Italien*	48.3	501	(2.9)	51.7	524	(3.2)	22	(2.6)	▲
Portugal	50.0	506	(3.0)	50.0	528	(3.2)	22	(2.9)	▲
123 *Neuseeland*	49.2	479	(2.9)	50.8	501	(3.2)	21	(3.3)	▲
2 *Belgien (Frz. Gem.)*	50.5	479	(2.4)	49.5	500	(2.5)	21	(2.3)	▲
23 *Kanada*	50.7	494	(1.6)	49.3	514	(2.0)	20	(1.8)	▲
Macau	48.3	572	(1.0)	51.7	592	(1.3)	20	(1.7)	▲
2 *Zypern*	49.0	506	(2.5)	51.0	526	(3.1)	21	(3.0)	■
23 *Belgien (Fläm. Gem.)*	48.7	511	(3.2)	51.3	530	(2.8)	18	(3.5)	■
123 *England*	49.9	543	(3.5)	50.1	561	(3.1)	18	(3.5)	■
2 *Ungarn*	49.8	511	(3.3)	50.2	529	(4.2)	18	(2.8)	■
23 *USA*	49.4	508	(2.5)	50.6	526	(3.0)	18	(1.7)	■
2 *Spanien*	49.5	489	(1.9)	50.5	507	(2.1)	18	(1.6)	■
3 *Niederlande*	50.1	528	(2.5)	49.9	546	(2.4)	17	(2.8)	■
Republik Korea (Südkorea)	49.7	586	(3.0)	50.3	603	(2.9)	17	(2.8)	■
Slowakei	49.5	506	(3.7)	50.5	523	(3.0)	17	(3.2)	■
23 *Chile*	47.2	435	(2.8)	52.8	452	(3.2)	17	(2.6)	■
VG OECD	**49.3**	**517**	**(0.5)**	**50.7**	**533**	**(0.6)**	**16**	**(0.5)**	■
2 *Schweden*	50.9	522	(2.9)	49.1	538	(3.1)	16	(2.7)	■
23 *Dänemark*	50.9	516	(2.3)	49.1	532	(2.5)	15	(2.5)	■
2 *Tschechien*	49.3	523	(2.0)	50.7	538	(2.6)	15	(2.3)	■
12 *Norwegen (5. Jgst.)*	50.5	523	(2.4)	49.5	538	(2.3)	15	(2.6)	■
VG EU	**49.4**	**516**	**(0.6)**	**50.6**	**531**	**(0.7)**	**15**	**(0.6)**	■
123 *Türkei (5. Jgst.)*	48.1	546	(4.5)	51.9	560	(5.0)	14	(4.7)	■
3 *Hongkong*	48.8	587	(4.3)	51.2	601	(4.4)	14	(3.3)	■
Deutschland	48.9	517	(2.2)	51.1	530	(2.1)	13	(2.2)	■
2 *Litauen*	49.2	554	(2.7)	50.8	567	(2.6)	13	(2.3)	■
2 *Singapur*	49.3	609	(2.4)	50.7	621	(2.5)	12	(1.9)	■
2 *Serbien*	50.7	518	(3.5)	49.3	528	(3.9)	11	(3.6)	■
2 *Polen*	49.8	541	(2.4)	50.2	551	(2.7)	11	(3.2)	■
Internationaler Mittelwert	**49.3**	**498**	**(0.4)**	**50.7**	**508**	**(0.4)**	**11**	**(0.4)**	■
Japan	50.7	586	(2.5)	49.3	596	(2.7)	10	(2.5)	■
2 *Slowenien*	49.2	509	(2.2)	50.8	519	(2.2)	10	(2.5)	■
Lettland	48.6	530	(3.4)	51.4	538	(2.9)	8	(3.4)	■
Irland	48.7	542	(3.8)	51.3	549	(3.4)	6	(4.0)	■
23 *Rumänien*	49.2	539	(5.2)	50.8	545	(5.3)	6	(4.2)	■
Taiwan	47.7	603	(1.7)	52.3	611	(1.7)	7	(1.8)	▼
Finnland	49.3	526	(2.6)	50.7	532	(2.7)	6	(2.6)	▼
Bulgarien	47.1	528	(3.8)	52.9	532	(4.1)	3	(3.5)	▼
Benchmark-Teilnehmer									
Dubai, VAE	48.3	548	(2.5)	51.7	565	(1.4)	17	(2.7)	■

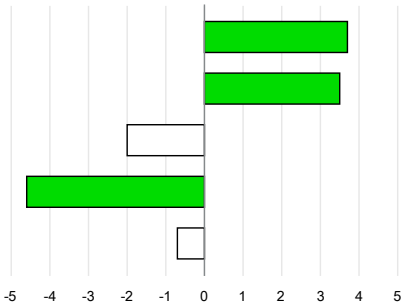


□ Keine statistisch signifikanten Unterschiede zwischen Mädchen und Jungen ($p > .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen
■ Statistisch signifikante Unterschiede zwischen Mädchen und Jungen ($p \leq .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen
▲ = Differenzwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Abbildung 6.2: Mittlere Testleistungen von Mädchen und Jungen und prozentuale Verteilung auf die Kompetenzstufen in Deutschland – Mathematik

KS	Mädchen				Jungen				Differenz A	
	%	(SE)	M	(SE)	%	(SE)	M	(SE)	$\%_J$-$\%_M$	(SE)
V	6.4	(0.6)	638	(21.7) ■	10.1	(0.6)	642	(20.3)	3.7	(0.8)
IV	27.8	(1.2)	577	(7.7) ■	31.3	(1.2)	580	(5.8)	3.5	(1.6)
III	38.0	(1.6)	514	(3.4) ■	36.1	(1.2)	517	(4.0)	-2.0	(1.7)
II	21.6	(1.7)	449	(9.5) ■	17.0	(0.7)	452	(10.2)	-4.6	(1.8)
I	6.1	(0.8)	383	(25.4) ■	5.5	(0.6)	376	(24.9)	-0.7	(0.9)



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

☐ Keine statistisch signifikanten Unterschiede zwischen Mädchen und Jungen ($p > .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

■ Statistisch signifikante Unterschiede zwischen Mädchen und Jungen ($p \leq .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

■ = Unterschied in den mittleren Leistungswerten zu Jungen nicht statistisch signifikant ($p > .05$)

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

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

6.4.2 Geschlechterspezifische Leistungsunterschiede nach mathematischen Inhaltsbereichen

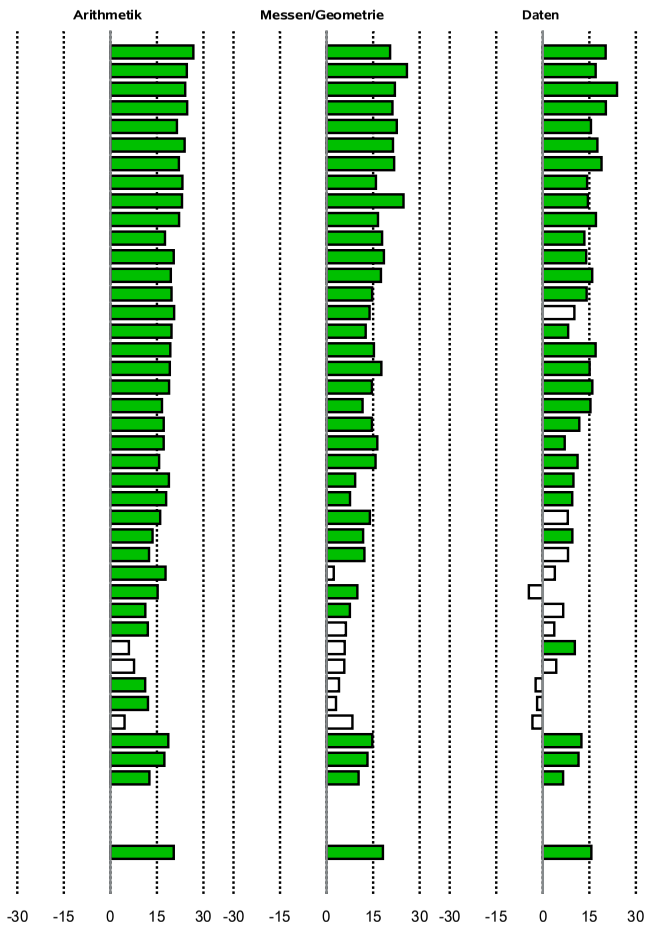
Die Konzeption von TIMSS sieht für die Domäne Mathematik drei Inhaltsbereiche vor. Dies sind *Arithmetik*, *Messen und Geometrie* und *Daten* (siehe Kapitel 3 in diesem Band). Die Darstellung der Inhaltsbereiche erfolgt im Vergleich zu anderen ausgewählten Teilnehmerstaaten und Benchmark-Teilnehmern (für die Vergleichsgruppen in TIMSS siehe Kapitel 2 in diesem Band) und ermöglicht so differenzierte Einblicke in Ausprägungen von Geschlechterunterschieden nach einzelnen Inhaltsbereichen. Die Darstellung in Abbildung 6.3 dokumentiert analog zu Abbildung 6.1 Geschlechterunterschiede nach Inhaltsbereichen in Mathematik im internationalen Vergleich.

Im Bereich *Arithmetik* fallen bei allen Teilnehmerstaaten in der Vergleichsgruppe die Differenzen der Leistungswerte zugunsten der Jungen aus. Lediglich in drei Staaten (Irland, Rumänien und Bulgarien; ohne Benchmark-Teilnehmer) sind sie nicht signifikant. Die größten Geschlechterunterschiede liegen in Frankreich mit 27 Punkten sowie in Australien und Portugal mit 25 Punkten zugunsten der Jungen vor. Der internationale Mittelwert liegt bei 13 Punkten, die Differenz in der Vergleichsgruppe OECD liegt im Durchschnitt bei 19 Punkten, die der EU bei 17 Punkten. Diese Geschlechterunterschiede sind jeweils signifikant und fallen wiederum zugunsten der Jungen aus. Deutschland befindet sich mit einem signifikanten Unterschied von 18 Differenzpunkten leicht über dem mittleren Wert der EU und leicht unter dem mittleren Wert der OECD, wobei Mädchen im Inhaltsbereich *Arithmetik* 515 Punkte und Jungen 533 Punkte erreichen. Insgesamt zeigen sich im Bereich *Arithmetik* sowohl die größten als auch konsistentesten Geschlechterunterschiede im Vergleich zu den beiden anderen mathematischen Inhaltsbereichen.

Auch im Inhaltsbereich *Messen und Geometrie* ist in allen Teilnehmerstaaten eine Differenz zugunsten der Jungen zu verzeichnen, die mit Ausnahme von sieben Teilnehmern signifikant ist. Die größten Unterschiede liegen in Australien mit 26 Punkten sowie in Zypern mit 25 Punkten zugunsten der Jungen vor.

Abbildung 6.3: Testleistungen in Inhaltsbereichen nach Geschlecht im internationalen Vergleich – Mathematik

	Arithmetik						Messen/Geometrie						Daten					
	Mädchen		Jungen		Differenz ^		Mädchen		Jungen		Differenz ^		Mädchen		Jungen		Differenz ^	
Teilnehmer	M_M	*(SE)*	M_J	*(SE)*	$M_J - M_M$	*(SE)*	M_M	*(SE)*	M_J	*(SE)*	$M_J - M_M$	*(SE)*	M_M	*(SE)*	M_J	*(SE)*	$M_J - M_M$	*(SE)*
2 *Frankreich*	466	(3.3)	492	(3.0)	27	(2.7)	484	(3.4)	505	(3.6)	20	(3.4)	470	(3.3)	490	(3.4)	20	(3.3)
Australien	508	(2.3)	533	(3.0)	25	(3.2)	509	(3.1)	535	(2.8)	26	(3.4)	532	(2.4)	549	(2.8)	17	(3.1)
2 *Italien*	503	(2.9)	527	(2.8)	24	(2.5)	500	(4.3)	522	(3.5)	22	(4.2)	492	(4.1)	516	(3.3)	24	(3.1)
Portugal	503	(3.3)	528	(3.1)	25	(3.3)	499	(4.1)	520	(3.2)	21	(4.4)	518	(3.5)	538	(3.4)	20	(3.7)
123 *Neuseeland*	476	(3.0)	497	(3.2)	21	(4.0)	472	(3.4)	495	(2.9)	23	(3.9)	495	(3.3)	511	(3.2)	16	(4.2)
2 *Belgien (Frz. Gem.)*	468	(2.9)	492	(2.9)	24	(3.6)	496	(3.6)	517	(3.2)	21	(5.3)	482	(3.5)	499	(3.0)	18	(2.8)
23 *Kanada*	490	(1.7)	512	(2.0)	22	(1.8)	488	(1.8)	510	(2.1)	22	(1.9)	505	(2.3)	524	(2.2)	19	(2.8)
Macau	566	(1.2)	589	(1.6)	23	(2.3)	582	(2.0)	598	(1.4)	16	(2.9)	576	(1.9)	590	(1.3)	14	(2.6)
2 *Zypern*	515	(2.6)	538	(2.9)	23	(3.3)	496	(3.1)	521	(3.1)	25	(3.7)	494	(3.3)	509	(3.9)	15	(3.5)
23 *Belgien (Fläm. Gem.)*	502	(2.9)	524	(2.7)	22	(3.0)	527	(3.5)	544	(3.6)	17	(3.8)	515	(3.3)	532	(3.1)	17	(3.4)
123 *England*	547	(3.6)	565	(3.5)	18	(3.9)	530	(4.0)	548	(3.9)	18	(4.5)	555	(4.2)	568	(4.1)	13	(4.2)
2 *Ungarn*	517	(3.2)	538	(4.0)	20	(2.7)	507	(3.6)	525	(4.4)	18	(3.4)	497	(4.1)	511	(5.0)	14	(4.9)
23 *USA*	513	(2.6)	533	(2.9)	19	(2.0)	490	(3.0)	508	(3.4)	18	(2.9)	511	(2.8)	527	(2.9)	16	(2.3)
2 *Spanien*	487	(2.2)	506	(2.2)	20	(1.5)	490	(2.3)	504	(2.7)	15	(2.5)	495	(1.9)	509	(2.9)	14	(3.1)
3 *Niederlande*	526	(2.5)	547	(3.3)	21	(3.9)	527	(2.9)	541	(3.7)	14	(3.7)	539	(4.2)	549	(3.0)	10	(5.4)
Republik Korea (Südkorea)	576	(3.0)	596	(3.0)	20	(2.9)	599	(3.2)	612	(3.3)	13	(3.6)	602	(4.0)	610	(3.0)	8	(3.8)
Slowakei	508	(3.5)	527	(2.9)	19	(3.1)	505	(3.5)	520	(3.3)	15	(3.3)	499	(4.2)	516	(3.9)	17	(3.9)
23 *Chile*	422	(2.6)	441	(2.9)	19	(2.8)	433	(3.0)	451	(2.9)	18	(3.5)	456	(3.4)	471	(3.5)	15	(3.7)
2 *Schweden*	518	(2.7)	537	(3.6)	19	(3.7)	525	(3.1)	540	(3.1)	15	(3.4)	528	(3.3)	544	(3.4)	16	(3.3)
23 *Dänemark*	508	(2.4)	525	(3.0)	17	(3.0)	524	(2.4)	536	(2.7)	12	(3.0)	524	(2.7)	540	(2.7)	15	(2.9)
2 *Tschechien*	525	(2.1)	542	(2.4)	17	(2.2)	529	(1.9)	544	(2.5)	15	(2.3)	506	(3.0)	518	(3.2)	12	(2.9)
12 *Norwegen (5. Jgst.)*	522	(3.1)	539	(2.4)	17	(3.6)	517	(2.7)	534	(3.5)	16	(3.7)	533	(2.4)	540	(2.5)	7	(2.6)
123 *Türkei (5. Jgst.)*	542	(5.0)	558	(5.3)	16	(4.9)	548	(5.8)	564	(5.4)	16	(7.0)	551	(4.4)	562	(4.6)	11	(4.7)
3 *Hongkong*	588	(4.6)	607	(4.4)	19	(3.4)	593	(5.2)	603	(4.8)	9	(4.5)	583	(4.8)	593	(4.0)	10	(3.5)
Deutschland	515	(2.5)	533	(2.4)	18	(2.1)	524	(2.4)	531	(2.5)	8	(2.4)	516	(2.6)	525	(2.8)	9	(2.9)
2 *Litauen*	554	(2.6)	570	(2.9)	16	(3.0)	549	(2.6)	563	(3.1)	14	(2.7)	563	(3.7)	571	(3.3)	8	(4.6)
2 *Singapur*	606	(2.5)	620	(2.7)	14	(2.3)	613	(2.7)	625	(2.8)	12	(2.5)	611	(2.9)	620	(3.4)	10	(2.2)
2 *Serbien*	522	(3.4)	535	(4.0)	12	(3.5)	518	(3.5)	530	(4.8)	12	(4.0)	501	(4.7)	509	(4.4)	8	(4.3)
2 *Polen*	532	(2.7)	550	(2.7)	18	(3.1)	556	(3.0)	558	(2.9)	2	(3.3)	544	(3.4)	548	(3.0)	4	(4.4)
Japan	574	(2.8)	589	(2.7)	15	(2.8)	600	(3.3)	610	(3.8)	10	(4.8)	600	(3.5)	595	(3.0)	-4	(4.2)
2 *Slowenien*	508	(2.3)	519	(2.5)	11	(2.9)	511	(2.4)	519	(2.9)	8	(3.5)	511	(2.7)	518	(3.7)	7	(4.8)
Lettland	527	(3.3)	539	(3.2)	12	(3.2)	537	(3.7)	543	(3.1)	6	(3.2)	530	(3.8)	534	(3.0)	4	(3.6)
Irland	545	(3.8)	551	(3.6)	6	(4.2)	537	(4.0)	543	(3.6)	6	(3.9)	541	(4.2)	551	(3.6)	10	(4.4)
23 *Rumänien*	548	(5.2)	555	(5.7)	8	(4.1)	535	(6.0)	541	(5.7)	6	(5.2)	517	(5.6)	521	(6.4)	4	(5.5)
Taiwan	596	(1.4)	607	(1.7)	11	(1.7)	620	(1.9)	624	(1.9)	4	(2.5)	602	(1.5)	600	(2.3)	-2	(1.9)
Finnland	516	(2.6)	528	(2.9)	12	(2.5)	537	(2.6)	540	(3.5)	3	(3.3)	537	(3.1)	535	(3.3)	-2	(3.2)
Bulgarien	542	(3.7)	547	(3.7)	5	(3.4)	522	(4.6)	531	(4.3)	8	(4.3)	508	(5.2)	504	(5.6)	-3	(5.0)
VG OECD	**513**	**(0.6)**	**532**	**(0.6)**	**19**	**(0.6)**	**519**	**(0.6)**	**534**	**(0.6)**	**15**	**(0.7)**	**522**	**(0.7)**	**534**	**(0.6)**	**12**	**(0.7)**
VG EU	**515**	**(0.6)**	**532**	**(0.7)**	**17**	**(0.7)**	**519**	**(0.7)**	**532**	**(0.7)**	**13**	**(0.8)**	**515**	**(0.8)**	**526**	**(0.8)**	**12**	**(0.8)**
Internationaler Mittelwert	**500**	**(0.4)**	**513**	**(0.5)**	**13**	**(0.5)**	**499**	**(0.5)**	**510**	**(0.5)**	**10**	**(0.5)**	**497**	**(0.5)**	**503**	**(0.5)**	**7**	**(0.6)**
Benchmark-Teilnehmer																		
Dubai, VAE	548	(2.2)	569	(1.6)	20	(2.6)	539	(2.4)	558	(1.7)	18	(2.9)	555	(2.5)	571	(1.8)	16	(3.1)



□ Keine statistisch signifikanten Unterschiede zwischen Mädchen und Jungen ($p > .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

■ Statistisch signifikante Unterschiede zwischen Mädchen und Jungen ($p \leq .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

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

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

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

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

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

© TIMSS 2023

Polen und Finnland sind mit 2 bzw. 3 Punkten die Staaten mit der kleinsten Differenz, wobei diese Unterschiede nicht signifikant ausfallen. Der internationale Mittelwert liegt bei 10 Punkten Differenz, der EU-Durchschnitt bei 13 Punkten und der der OECD bei 15 Punkten. Diese Differenzen fallen zugunsten der Jungen aus und sind signifikant. Mit 8 Punkten Differenz liegt Deutschland somit unter dem EU- und OECD-Durchschnitt der Geschlechterunterschiede im Inhaltsbereich *Messen und Geometrie*. Mit 524 Punkten, die Mädchen erreichen, und 531 Punkten, die Jungen in diesem Inhaltsbereich erzielen, fällt dieser Unterschied ebenfalls signifikant aus.

Im Inhaltsbereich *Daten* erreichen Jungen in der Mehrzahl der Teilnehmerstaaten höhere Punktwerte als Mädchen. In vier Staaten erreichen Mädchen leicht höhere Werte, die Differenz ist jedoch nicht signifikant. Am größten sind die Differenzen zugunsten der Jungen in Italien (24 Punkte) sowie in Frankreich und Portugal (jeweils 20 Punkte). Insgesamt liegt der internationale Mittelwert bei 7 Differenzpunkten, was dem niedrigsten Wert im Vergleich der drei Inhaltsbereiche entspricht. Die Vergleichsgruppen OECD und EU weisen beide eine Differenz von 12 Punkten auf, jeweils zugunsten der Jungen. Deutschland liegt mit 9 Punkten unter diesen Vergleichswerten im Inhaltsbereich *Daten*. Mädchen erreichen 516 Punkte und Jungen 525 Punkte. Der Unterschied ist signifikant. Insgesamt sind die Unterschiede in diesem Inhaltsbereich über die Teilnehmerstaaten hinweg nicht so deutlich ausgeprägt und konsistent in der Richtung wie in den zwei anderen Inhaltsbereichen.

6.4.3 Geschlechterspezifische Leistungsunterschiede nach kognitiven Anforderungsbereichen in Mathematik

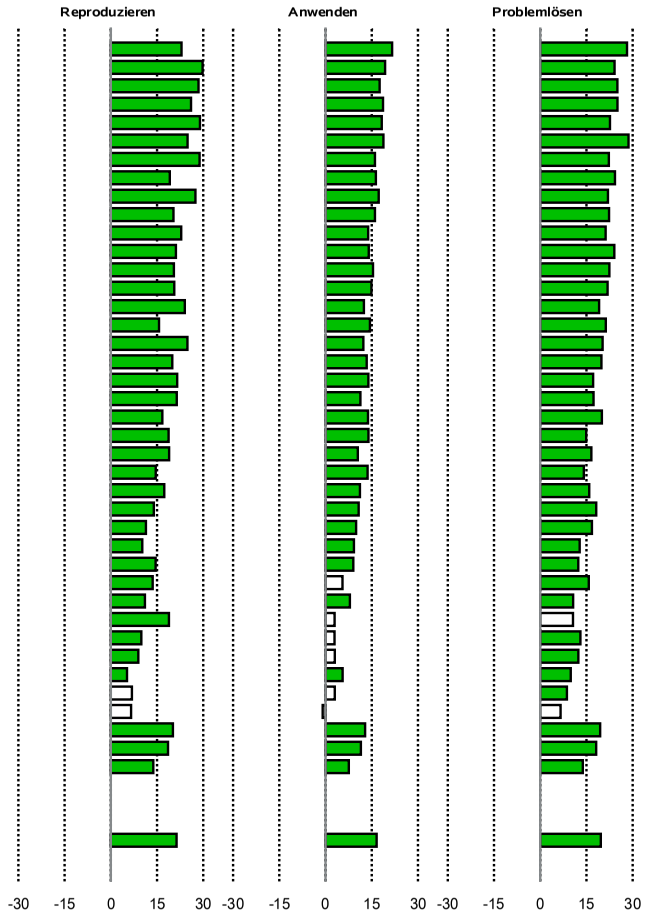
Neben einer Differenzierung nach Inhaltsbereichen ermöglicht das Testdesign von TIMSS auch eine Differenzierung nach kognitiven Anforderungsbereichen. Hierzu gehören die drei Bereiche *Reproduzieren*, *Anwenden* und *Problemlösen* (siehe Kapitel 3 in diesem Band). Erneut können Unterschiede zwischen Mädchen und Jungen in den jeweiligen kognitiven Anforderungsbereichen sowohl für Deutschland als auch im internationalen Vergleich beobachtet werden. Die Ergebnisse sind in Abbildung 6.4 analog zu den vorherigen Darstellungsformen dokumentiert.

Wie bereits für die mathematischen Inhaltsbereiche sind auch hier die kognitiven Anforderungsbereiche nebeneinander abgebildet. In allen kognitiven Anforderungsbereichen können Differenzen zugunsten der Jungen in beinahe allen abgebildeten Teilnehmerstaaten festgestellt werden.

Wie in Abbildung 6.4 ersichtlich, kann im kognitiven Anforderungsbereich *Reproduzieren* bei 35 Teilnehmerstaaten ein signifikanter Unterschied zugunsten der Jungen verzeichnet werden. Der größte Geschlechterunterschied lässt sich für Australien berichten, wo Jungen durchschnittlich 30 Punkte mehr erreichen als Mädchen. Aber auch in Neuseeland und Kanada mit jeweils 29 Punkten und in Italien mit 28 Punkten zeigen sich deutliche Unterschiede zwischen den Geschlechtern. Die Differenz im internationalen Mittel beträgt 14 Punkte. In den Vergleichsgruppen fallen die Differenzen in der OECD mit 20 Punkten und der EU mit 19 Punkten jeweils signifikant zugunsten der Jungen aus. In Deutschland liegt ebenfalls eine signifikante Differenz vor. Diese beträgt 17 Punkte und liegt damit unter den Vergleichswerten der OECD und der EU. Jungen erreichen 541 Punkte und Mädchen 524 Punkte.

Abbildung 6.4: Testleistungen in kognitiven Anforderungsbereichen nach Geschlecht im internationalen Vergleich – Mathematik

	Reproduzieren						Anwenden						Problemlösen					
	Mädchen		Jungen		Differenz^		Mädchen		Jungen		Differenz^		Mädchen		Jungen		Differenz^	
Teilnehmer	M_M	(SE)	M_J	(SE)	$M_J - M_M$	(SE)	M_M	(SE)	M_J	(SE)	$M_J - M_M$	(SE)	M_M	(SE)	M_J	(SE)	$M_J - M_M$	(SE)
2 *Frankreich*	473	(3.5)	496	(4.0)	23	(3.4)	473	(3.4)	495	(3.5)	22	(3.5)	468	(3.2)	496	(3.2)	28	(3.0)
Australien	515	(2.7)	544	(3.3)	30	(3.6)	513	(2.2)	533	(2.6)	19	(2.8)	515	(3.0)	539	(2.9)	24	(3.5)
2 *Italien*	496	(3.3)	524	(3.3)	28	(3.3)	504	(3.0)	522	(3.4)	18	(2.9)	500	(3.1)	526	(2.9)	25	(2.6)
Portugal	504	(3.0)	530	(3.3)	26	(3.2)	507	(3.2)	525	(3.3)	19	(3.0)	506	(3.4)	531	(3.0)	25	(3.3)
123 *Neuseeland*	473	(3.6)	502	(3.1)	29	(4.0)	479	(3.0)	498	(3.0)	18	(3.7)	483	(2.9)	506	(3.2)	23	(3.8)
2 *Belgien (Frz. Gem.)*	483	(3.1)	508	(2.8)	25	(3.3)	479	(2.8)	498	(3.1)	19	(2.7)	469	(2.6)	498	(2.9)	29	(2.7)
23 *Kanada*	488	(1.7)	517	(2.1)	29	(2.0)	496	(1.5)	512	(2.2)	16	(1.9)	493	(2.1)	515	(2.4)	22	(2.5)
Macau	572	(1.3)	591	(2.1)	19	(2.2)	575	(1.9)	591	(1.5)	16	(2.6)	570	(1.1)	594	(2.1)	24	(2.5)
2 *Zypern*	505	(2.8)	532	(3.0)	27	(3.3)	505	(3.0)	523	(3.2)	17	(3.4)	504	(2.7)	526	(3.3)	22	(3.6)
23 *Belgien (Fläm. Gem.)*	523	(3.4)	543	(3.3)	20	(3.0)	509	(3.0)	525	(2.8)	16	(3.2)	505	(3.7)	527	(3.1)	22	(3.7)
123 *England*	547	(4.3)	569	(3.6)	23	(5.2)	543	(3.5)	557	(3.4)	14	(4.0)	539	(4.0)	560	(3.7)	21	(3.8)
2 *Ungarn*	517	(3.7)	538	(4.6)	21	(4.6)	508	(3.4)	522	(4.0)	14	(3.0)	508	(3.4)	532	(4.1)	24	(3.4)
23 *USA*	510	(3.0)	531	(3.4)	20	(2.8)	509	(2.7)	525	(3.0)	15	(1.7)	501	(3.2)	524	(3.0)	22	(2.6)
2 *Spanien*	489	(2.7)	510	(2.8)	21	(2.1)	490	(1.8)	505	(2.2)	15	(2.3)	489	(2.2)	511	(2.3)	22	(2.4)
3 *Niederlande*	528	(3.9)	552	(3.0)	24	(4.4)	530	(2.7)	543	(2.7)	12	(3.1)	527	(2.8)	547	(2.9)	19	(3.4)
Republik Korea (Südkorea)	592	(3.2)	608	(4.2)	16	(4.4)	586	(3.1)	600	(2.9)	14	(3.0)	582	(3.5)	603	(3.4)	21	(4.0)
Slowakei	500	(4.2)	525	(3.1)	25	(3.6)	508	(3.9)	520	(3.0)	12	(3.1)	507	(3.6)	528	(3.1)	20	(2.8)
23 *Chile*	423	(2.9)	443	(3.8)	20	(4.0)	437	(3.0)	451	(3.2)	13	(4.1)	442	(3.9)	462	(4.4)	20	(5.9)
2 *Schweden*	514	(2.8)	536	(3.2)	22	(3.4)	524	(2.9)	537	(3.0)	14	(3.2)	525	(2.6)	542	(3.5)	17	(3.1)
23 *Dänemark*	510	(2.5)	531	(2.6)	21	(2.6)	517	(2.6)	529	(2.9)	11	(2.6)	517	(3.6)	534	(2.6)	17	(3.8)
2 *Tschechien*	526	(2.7)	542	(2.9)	17	(2.6)	521	(1.9)	535	(2.8)	14	(2.6)	518	(2.4)	538	(3.3)	20	(3.2)
12 *Norwegen (5. Jgst.)*	516	(3.1)	535	(3.1)	19	(3.9)	524	(2.5)	538	(2.3)	14	(2.7)	527	(2.7)	542	(2.5)	15	(2.9)
123 *Türkei (5. Jgst.)*	533	(6.0)	552	(5.0)	19	(5.1)	554	(4.8)	564	(5.2)	10	(4.9)	542	(5.2)	559	(5.3)	17	(5.4)
3 *Hongkong*	591	(4.3)	606	(4.6)	15	(3.8)	585	(4.6)	599	(4.4)	14	(3.3)	588	(4.6)	602	(5.2)	14	(4.2)
Deutschland	524	(3.0)	541	(2.5)	17	(2.4)	513	(2.6)	525	(2.3)	11	(2.1)	516	(2.4)	531	(3.2)	16	(3.5)
2 *Litauen*	549	(2.5)	563	(2.6)	14	(2.3)	561	(2.6)	571	(2.7)	11	(2.7)	545	(2.9)	563	(3.1)	18	(3.1)
2 *Singapur*	619	(2.5)	630	(2.6)	11	(2.3)	610	(2.4)	620	(2.8)	10	(2.5)	600	(2.7)	617	(3.0)	17	(3.1)
2 *Serbien*	517	(3.7)	528	(3.8)	10	(3.7)	517	(3.7)	527	(4.1)	9	(4.0)	517	(3.8)	530	(3.9)	13	(3.9)
2 *Polen*	531	(2.7)	546	(3.6)	15	(4.5)	542	(2.9)	551	(2.9)	9	(3.4)	543	(3.3)	556	(2.9)	12	(3.7)
Japan	585	(4.0)	598	(4.1)	14	(6.5)	594	(2.9)	600	(3.1)	6	(3.3)	568	(3.1)	584	(3.2)	16	(4.0)
2 *Slowenien*	511	(2.3)	522	(2.6)	11	(2.8)	510	(2.3)	518	(2.4)	8	(3.0)	504	(2.5)	515	(2.4)	11	(3.0)
Lettland	525	(3.3)	543	(2.9)	19	(3.5)	533	(3.4)	536	(2.9)	3	(3.3)	528	(4.0)	539	(4.0)	11	(5.6)
Irland	546	(3.9)	555	(4.1)	10	(4.6)	545	(3.9)	547	(3.4)	3	(4.0)	534	(4.0)	547	(3.6)	13	(4.6)
23 *Rumänien*	533	(4.9)	542	(5.6)	9	(4.5)	541	(5.0)	544	(5.6)	3	(3.6)	536	(5.3)	549	(5.7)	12	(3.6)
Taiwan	616	(1.4)	621	(2.0)	5	(2.1)	609	(1.8)	614	(2.0)	6	(2.2)	584	(1.5)	594	(2.0)	10	(2.2)
Finnland	534	(3.2)	541	(3.1)	7	(3.7)	524	(2.8)	527	(2.8)	3	(2.9)	524	(3.0)	532	(3.3)	9	(2.6)
Bulgarien	525	(3.6)	531	(4.7)	7	(4.2)	533	(3.9)	532	(4.1)	-1	(3.4)	518	(5.0)	525	(5.4)	7	(4.9)
VG OECD	**516**	**(0.6)**	**536**	**(0.6)**	**20**	**(0.7)**	**518**	**(0.6)**	**531**	**(0.6)**	**13**	**(0.6)**	**514**	**(0.6)**	**534**	**(0.6)**	**19**	**(0.7)**
VG EU	**516**	**(0.7)**	**534**	**(0.7)**	**19**	**(0.7)**	**517**	**(0.7)**	**529**	**(0.7)**	**11**	**(0.7)**	**513**	**(0.7)**	**531**	**(0.7)**	**18**	**(0.8)**
Internationaler Mittelwert	**498**	**(0.5)**	**511**	**(0.5)**	**14**	**(0.5)**	**502**	**(0.5)**	**509**	**(0.5)**	**8**	**(0.5)**	**497**	**(0.5)**	**510**	**(0.5)**	**14**	**(0.6)**
Benchmark-Teilnehmer																		
Dubai, VAE	546	(2.1)	567	(1.5)	21	(2.5)	548	(2.2)	564	(1.6)	17	(2.4)	548	(2.8)	567	(1.6)	20	(2.8)



Keine statistisch signifikanten Unterschiede zwischen Mädchen und Jungen ($p > .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

Statistisch signifikante Unterschiede zwischen Mädchen und Jungen ($p \leq .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

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

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

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

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

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

© TIMSS 2023

Im Bereich *Anwenden* sind die Geschlechterunterschiede in 31 Teilnehmerstaaten zugunsten der Jungen signifikant. Frankreich weist mit 22 Punkten die größten Geschlechterunterschiede auf, Bulgarien die geringste (nicht signifikante) Geschlechterdifferenz (1 Punkt Differenz zugunsten der Mädchen). Die Vergleichsgruppe der OECD verzeichnet eine Differenz von 13 Punkten, die der EU eine von 11 Punkten (jeweils signifikant). Der internationale Mittelwert liegt bei 8 Punkten und ist ebenfalls signifikant. Deutschland ordnet sich mit 11 Punkten zugunsten der Jungen im EU- und unter dem OECD-Durchschnitt ein. Diese Differenz ist auch in Deutschland signifikant.

Die Ergebnisse für Geschlechterunterschiede im kognitiven Anwendungsbereich *Problemlösen* ähneln stark den Befunden zu *Reproduzieren* und *Anwenden*. Auch hier sind die Geschlechterunterschiede in 35 Teilnehmerstaaten signifikant. Diesmal weist die französischsprachige Gemeinschaft Belgiens die größte Geschlechterdifferenz mit 29 Punkten auf, danach folgt Frankreich mit 28 Punkten. Die geringste Differenz findet sich in Finnland (9 Punkte) und Bulgarien (7 Punkte). Der internationale Mittelwert beträgt 14 Punkte, der Mittelwert in der Vergleichsgruppe EU 18 Punkte und in der OECD 19 Punkte. Alle drei Geschlechterunterschiede sind signifikant. Deutschland liegt mit einer signifikanten Differenz von 16 Punkten zugunsten der Jungen unter diesen Vergleichswerten, aber über dem internationalen Durchschnitt.

6.4.4 Trends in geschlechterspezifischen Leistungsunterschieden in Mathematik

Nachdem bisher die Befunde basierend auf den Daten aus dem aktuellen TIMSS-Zyklus 2023 beschrieben wurden, wird im Folgenden nun der Fokus auf etwaige Veränderungen im Vergleich zu vergangenen TIMSS-Zyklen gelegt. Diese Trendanalysen ermöglichen eine Einordnung der berichteten Befunde vor dem Hintergrund möglicher Veränderungen, auch im internationalen Vergleich. Bei der Interpretation ist zu beachten, dass die Zusammensetzung der Schülerschaft, aber auch gesellschaftliche Ereignisse, wie etwa die COVID-19-Pandemie, eine enorme Bedeutung für die berichteten Unterschiede haben können. Aus diesem Grund finden sich komplexere Analysen unter Berücksichtigung von weiteren Prädiktorvariablen wie etwa dem sozioökonomischen Status oder der Herkunft in Kapitel 9 in diesem Band.

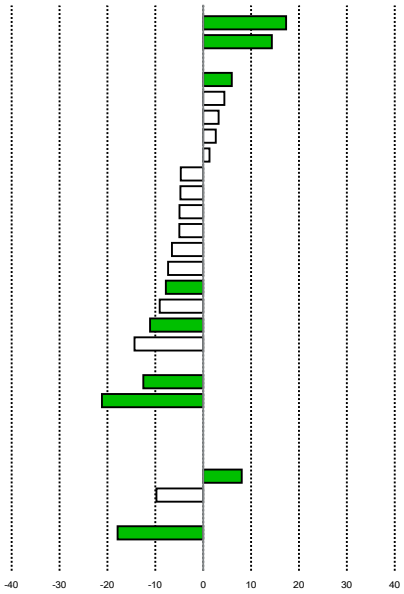
In Abbildung 6.5 lassen sich die Entwicklungen aus den Jahren 2007, 2019 und 2023 von 19 Teilnehmerstaaten und drei Benchmark-Teilnehmern getrennt nach Mädchen und Jungen ablesen. Bei den Mädchen können 13 Teilnehmerstaaten einen nominellen Zuwachs der Leistungspunkte seit 2007 verzeichnen, wobei dieser bei acht Staaten signifikant ist. Die nominell größte positive Entwicklung im Trend verzeichnet Georgien mit 55 Punkten von 440 auf 495 Leistungspunkte. Die Entwicklung von Tschechien ist insofern bemerkenswert, als dass zwischen 2007 und 2023 insgesamt eine Zunahme von 40 Punkten verzeichnet werden kann. Bei genauerer Betrachtung ist dies jedoch insbesondere auf einen enormen Anstieg von 2007 bis 2019 zurückzuführen, im Zeitraum von 2019 bis 2023 hingegen liegt kein signifikanter Anstieg mehr vor.

Bei den Benchmark-Teilnehmern zeigt sich ein ähnliches Bild für Dubai. Mit 96 Punkten in TIMSS 2023 konnte im Vergleich zu 2007 ein deutlicher Zuwachs verzeichnet werden, der Unterschied zu 2019 beträgt allerdings lediglich 8 Leistungspunkte. Im aktuellen Zyklus verzeichnen drei Teilnehmerstaaten signifikante

Abbildung 6.5: Vergleich der Leistungsentwicklungen von Mädchen und Jungen in Mathematik in TIMSS 2007, 2019 und 2023 (internationaler Vergleich)

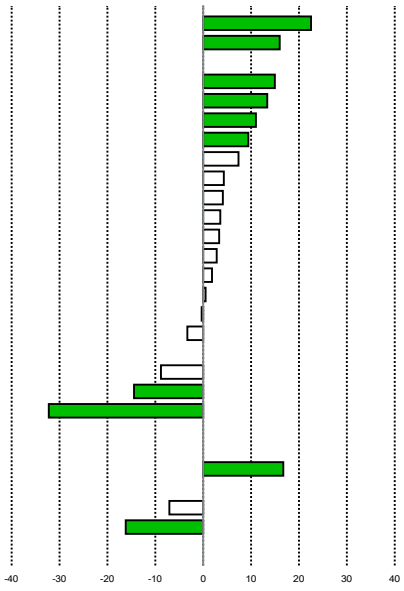
Mädchen

Fußnoten 2007	2019	2023	T	Teilnehmer	2007 M_{07}	(SE)	2019 M_{19}	(SE)	2023 M_{23}	(SE)	Veränderungen[A] M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)	
2	2	2		Georgien	440	(4.2)	478	(3.9)	495	(3.3)	55	(5.3)	+	17	(5.1)	+ ▲
2	2	2	8	Litauen	530	(2.8)	540	(2.9)	554	(2.7)	24	(3.9)	+	14	(4.0)	+ ▲
				Taiwan	575	(2.0)	597	(2.4)	603	(1.7)	28	(2.6)	+	6	(2.9)	+ ■
	2	2		Schweden	499	(2.4)	518	(3.2)	522	(2.9)	23	(3.7)	+	4	(4.3)	■
				Australien	513	(4.2)	511	(2.9)	514	(2.1)	1	(4.7)		3	(3.6)	■
	2			Slowakei	493	(4.6)	503	(3.5)	506	(3.7)	13	(5.9)	+	3	(5.1)	■
				Deutschland	519	(2.5)	516	(2.8)	517	(2.2)	-2	(3.3)		1	(3.6)	■
		2		Tschechien	483	(3.3)	527	(2.7)	523	(2.0)	40	(3.8)	+	-5	(3.4)	■
3	3	3	8	Niederlande	530	(2.7)	533	(2.2)	528	(2.5)	-2	(3.7)		-5	(3.4)	■
1 2	1 2	1 2 3	8	Neuseeland	492	(2.4)	484	(3.7)	479	(2.9)	-13	(3.8)	–	-5	(4.7)	■
3	3	2 3	8	Dänemark	520	(2.9)	521	(2.2)	516	(2.3)	-4	(3.7)		-5	(3.2)	■
		2		Ungarn	508	(4.6)	518	(3.0)	511	(3.3)	3	(5.7)		-7	(4.5)	■
2		2	8	Italien	499	(3.2)	509	(2.7)	501	(2.9)	2	(4.3)		-7	(4.0)	■
	3		8	Japan	568	(2.5)	593	(2.2)	586	(2.5)	18	(3.5)	+	-8	(3.3)	– ■
1 3	1 2	1 2 3	8	England	541	(3.2)	552	(4.0)	543	(3.5)	2	(4.7)		-9	(5.3)	■
2 3	3	3	8	Hongkong	605	(3.2)	599	(3.5)	587	(4.3)	-18	(5.3)	–	-11	(5.5)	– ■
			4 8	Iran	409	(5.2)	439	(6.4)	425	(4.4)	16	(6.8)	+	-14	(7.7)	■
	2	2	6	Singapur	603	(3.8)	621	(4.0)	609	(2.4)	6	(4.5)		-13	(4.7)	– ▼
2 3	2 3	2 3		USA	526	(2.7)	529	(3.0)	508	(2.5)	-18	(3.7)	–	-21	(3.9)	– ▼
				Benchmark-Teilnehmer												
2 3 6	2		8	Dubai, VAE	452	(4.0)	540	(3.2)	548	(2.5)	96	(4.7)	+	8	(4.0)	+ ■
2		2	8	Ontario, Kanada	509	(3.2)	502	(4.9)	493	(2.5)	-16	(4.0)	–	-10	(5.5)	■
2		3 2 3	8	Québec, Kanada	515	(3.5)	523	(3.0)	505	(2.6)	-10	(4.4)	–	-18	(4.0)	– ▼



Jungen

Fußnoten 2007	2019	2023	T	Teilnehmer	2007 M_{07}	(SE)	2019 M_{19}	(SE)	2023 M_{23}	(SE)	Veränderungen[A] M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)	
2	2	2	8	Litauen	530	(3.2)	544	(3.7)	567	(2.6)	37	(4.1)	+	23	(4.5)	+ ▲
				Australien	519	(3.6)	521	(3.3)	537	(2.5)	18	(4.4)	+	16	(4.1)	+ ▲
2	2	2		Georgien	437	(4.9)	486	(4.1)	500	(3.7)	63	(6.1)	+	15	(5.5)	+ ■
	2	2		Schweden	506	(3.1)	525	(3.1)	538	(3.1)	32	(4.4)	+	13	(4.4)	+ ■
1 2	1 2	1 2 3	8	Neuseeland	493	(3.1)	490	(3.3)	501	(3.2)	8	(4.4)		11	(4.6)	+ ■
				Taiwan	577	(2.0)	601	(2.3)	611	(1.7)	34	(2.6)	+	9	(2.9)	+ ■
	2			Slowakei	499	(4.7)	516	(4.2)	523	(3.0)	24	(5.6)	+	7	(5.2)	■
				Deutschland	531	(2.5)	526	(2.4)	530	(2.1)	-1	(3.3)		4	(3.2)	■
3	3	3	8	Niederlande	540	(2.4)	542	(3.0)	546	(2.4)	6	(3.4)		4	(3.8)	■
3	3	2 3	8	Dänemark	526	(3.2)	528	(2.6)	532	(2.5)	6	(4.1)		4	(3.7)	■
	3		8	Japan	568	(2.7)	593	(1.9)	596	(2.7)	28	(3.9)	+	3	(3.4)	■
2		2	8	Italien	514	(3.6)	521	(3.2)	524	(3.2)	10	(4.8)	+	3	(4.5)	■
1 3	1 2	1 2 3	8	England	542	(3.6)	560	(3.0)	561	(3.1)	19	(4.8)	+	2	(4.4)	■
		2		Ungarn	511	(3.8)	529	(3.1)	529	(4.2)	18	(5.7)	+	0	(5.2)	■
		2		Tschechien	489	(3.0)	538	(3.1)	538	(2.6)	49	(4.0)	+	0	(4.1)	■
2 3	3	3	8	Hongkong	609	(4.4)	604	(3.9)	601	(4.4)	-8	(6.2)		-3	(5.9)	■
	2	2	8	Singapur	596	(4.1)	629	(4.2)	621	(2.5)	25	(4.8)	+	-9	(4.9)	▼
2 3	2 3	2 3		USA	532	(2.7)	540	(2.9)	526	(3.0)	-6	(4.0)		-14	(4.1)	– ▼
		4 8		Iran	396	(5.5)	447	(5.3)	414	(5.1)	18	(7.5)	+	-32	(7.4)	– ▼
				Benchmark-Teilnehmer												
2 3 6	2		8	Dubai, VAE	438	(4.9)	548	(2.3)	565	(1.4)	127	(5.1)	+	17	(2.7)	+ ▲
2		2	8	Ontario, Kanada	514	(3.7)	521	(3.2)	514	(3.2)	-1	(4.9)		-7	(4.6)	▼
2		3 2 3	8	Québec, Kanada	524	(3.3)	541	(2.4)	525	(2.5)	1	(4.1)		-16	(3.5)	– ▼



Keine statistisch signifikanten Unterschiede zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Unterschiede zwischen 2019 und 2023 ($p \leq .05$)
▲ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
\+ = Mittelwert in 2023 signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

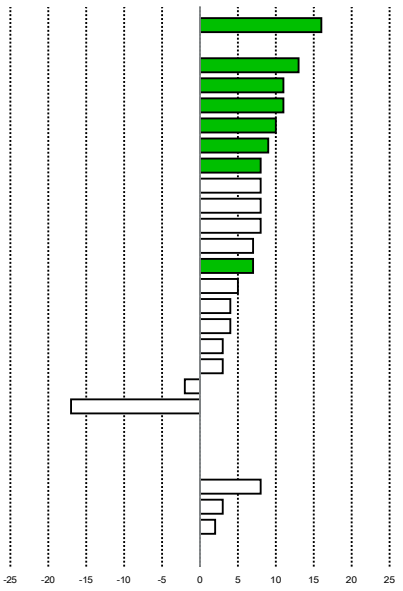
Leistungsrückgänge bei Mädchen im Bereich Mathematik im Vergleich zu 2007: Neuseeland mit 13 Punkten sowie Hongkong und die USA mit jeweils 18 Punkten. Für die USA und Hongkong (wie auch für Singapur und Japan) zeigt sich darüber hinaus auch signifikante Abnahme im Vergleich von 2019 und 2023. Mädchen in Deutschland verzeichnen im Jahr 2023 1 Punkt mehr als noch im Jahr 2019, jedoch 2 Punkte weniger als 2007. Dieser Trend ist nicht signifikant. Somit scheinen die mathematischen Leistungen der Mädchen in Deutschland im zeitlichen Verlauf relativ stabil zu sein.

Im unteren Abschnitt von Abbildung 6.5 sind die Leistungstrends der Jungen in den Jahren 2007, 2019 und 2023 abgebildet. Bei 13 Teilnehmerstaaten (ohne Benchmark-Teilnehmer) lässt sich ein signifikanter Zuwachs der Mathematikleistungen seit 2007 verzeichnen, bei sechs lassen sich zwar leichte, aber nicht signifikante Auf- oder Abwärtstrends erkennen. Den größten Zuwachs bei der Entwicklung der Mathematikleistungen bei Jungen kann erneut Georgien verzeichnen. Mit einem Anstieg von 437 Punkten (2007) auf 486 Punkte (2019) sowie auf 500 Leistungspunkte im Jahr 2023 zeigt sich ein positiver und statistisch signifikanter Trend. Der größte Zuwachs zwischen 2019 und 2023 ist für Litauen mit 23 Punkten sowie für Australien mit 16 Punkten zu verzeichnen. Analog zur Entwicklung von Mädchen kann wiederum für Tschechien beobachtet werden, dass sich der positive Trend für Jungen zwischen 2007 und 2019 in TIMSS 2023 nicht fortsetzen konnte. Auffällige Entwicklungen zeigen sich zudem für die Testleistungen von Jungen im Iran: In 2023 erreichen Jungen durchschnittlich 32 Punkte weniger als noch im Jahr 2019. Bei den Benchmark-Teilnehmern zeigt wiederum Dubai einen starken Leistungszuwachs der Jungen von 127 Punkten im Vergleich zum Jahr 2007. Jungen in Deutschland erreichen in TIMSS 2023 1 Leistungspunkt weniger im Vergleich zu 2007, seit 2019 verzeichnen sie jedoch wieder 4 Punkte mehr, wobei diese Differenzen über die Zeit nicht signifikant sind. Damit scheinen die Mathematikleistungen der Jungen, analog zu denen der Mädchen, in Deutschland relativ stabil zu sein. Ein signifikanter Negativtrend seit 2007 kann bei keinem der Teilnehmerstaaten (inklusive Benchmark-Teilnehmern) beobachtet werden.

Abbildung 6.6 ermöglicht eine Einschätzung, wie sich die Geschlechterunterschiede in den Mathematikleistungen in den Jahren 2007, 2019 und 2023 in Deutschland und in ausgewählten Teilnehmerstaaten entwickelt haben. In der Logik vorheriger Abbildungen entsprechen negative Differenzen einem Vorsprung der Mädchen und positive Differenzen einem Vorsprung der Jungen, wobei sich die rechts in der Tabelle verzeichneten Veränderungen auf die Entwicklungen zwischen 2007 bzw. 2019 und 2023 sowie ganz rechts auf die Entwicklung der Geschlechterunterschiede zwischen 2019 und 2023 (inklusive farbigen Balken) beziehen. Der Abbildung ist zu entnehmen, dass zwölf der abgebildeten Trend-Teilnehmerstaaten einen signifikanten Zuwachs von Geschlechterdifferenzen in den Mathematikleistungen von 2007 zu 2023 aufweisen. Am deutlichsten ausgeprägt ist die Zunahme von Geschlechterunterschieden in Neuseeland (von 1 Punkt im Jahr 2007 auf 21 Punkte im Jahr 2023). Im Vergleich von 2019 und 2023 zeigt sich für insgesamt acht Staaten ein signifikanter Anstieg der Geschlechterdifferenzen. In Deutschland kann hingegen keine signifikante Veränderung der Geschlechterunterschiede in den Mathematikleistungen über die Zeit beobachtet werden. Mit einer Differenz von 12 Punkten im Jahr 2007 sowie 10 Punkten im Jahr 2019 und 13 Punkten in 2023 bleibt diese relativ stabil. Ein Rückgang der nach wie vor signifikanten Geschlechterunterschiede auf der Gesamtskala Mathematik kann damit nicht beobachtet werden.

Abbildung 6.6: Vergleich der Leistungsunterschiede in Mathematik von Mädchen und Jungen in TIMSS 2007, 2019 und 2023 (internationaler Vergleich)

Fußnoten 2007	2019	2023	T	Teilnehmer	2007 M_{07}	(SE)	2019 M_{19}	(SE)	2023 M_{23}	(SE)	Veränderungen^ M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)	
1 2	1 2	1 2 3	8	Neuseeland	1	(3.0)	5	(4.6)	21	(3.3)	20	(4.5)	+	16	(5.7)	+ ▲
				Australien	6	(3.4)	10	(2.9)	23	(2.6)	17	(4.3)	+	13	(3.9)	+ ■
	3		8	Japan	0	(3.1)	-1	(2.2)	10	(2.5)	10	(4.0)	+	11	(3.3)	+ ■
1 3	1 2	1 2 3	8	England	0	(3.7)	7	(3.8)	18	(3.5)	18	(5.1)	+	11	(5.2)	+ ■
2		2	8	Italien	15	(2.5)	12	(3.4)	22	(2.6)	7	(3.6)		10	(4.3)	+ ■
	2	2		Schweden	6	(2.4)	7	(2.8)	16	(2.7)	10	(3.6)	+	9	(3.9)	+ ■
3	3	2 3	8	Dänemark	7	(3.7)	7	(2.9)	15	(2.5)	8	(4.5)		8	(3.8)	+ ■
2 3	3	3	8	Hongkong	4	(2.9)	6	(3.3)	14	(3.3)	10	(4.4)	+	8	(4.7)	■
2	2	2	8	Litauen	0	(3.6)	5	(3.8)	13	(2.3)	13	(4.3)	+	8	(4.4)	■
3	3	3	8	Niederlande	10	(2.7)	9	(3.0)	17	(2.8)	7	(3.9)		8	(4.1)	■
		2		Ungarn	3	(4.7)	11	(3.0)	18	(2.8)	15	(5.5)	+	7	(4.1)	■
2 3	2 3	2 3		USA	6	(2.4)	11	(2.9)	18	(1.7)	12	(2.9)	+	7	(3.4)	+ ■
	2			Slowakei	6	(2.7)	12	(3.6)	17	(3.2)	11	(4.2)	+	5	(4.8)	■
		2		Tschechien	6	(2.8)	11	(2.9)	15	(2.3)	9	(3.6)	+	4	(3.7)	■
	2	2	8	Singapur	-6	(2.7)	8	(2.8)	12	(1.9)	18	(3.3)	+	4	(3.4)	■
				Taiwan	2	(2.1)	4	(2.7)	7	(1.8)	5	(2.8)		3	(3.2)	■
				Deutschland	12	(2.1)	10	(2.5)	13	(2.2)	1	(3.0)		3	(3.3)	■
2	2	2		Georgien	-3	(3.7)	7	(3.3)	5	(3.7)	8	(5.2)		-2	(5.0)	■
		4	8	Iran	-14	(7.0)	7	(8.8)	-10	(6.4)	4	(9.5)		-17	(10.9)	■
				Benchmark-Teilnehmer												
2 3 6	2		8	Dubai, VAE	-14	(8.1)	9	(4.4)	17	(2.7)	31	(8.5)	+	8	(5.2)	■
2		2	8	Ontario, Kanada	6	(3.0)	18	(4.4)	21	(3.1)	15	(4.3)	+	3	(5.4)	■
2	3	2 3	8	Québec, Kanada	9	(3.1)	18	(3.0)	20	(2.1)	11	(3.7)	+	2	(3.7)	■



□ Keine statistisch signifikante Veränderung zwischen 2019 und 2023 ($p > .05$)
■ Statistisch signifikante Veränderung zwischen 2019 und 2023 ($p \leq .05$)
▲ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
\+ = Mittelwert in 2023 signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

6.5 Ergebnisse zu geschlechterspezifischen Leistungsunterschieden in den Naturwissenschaften

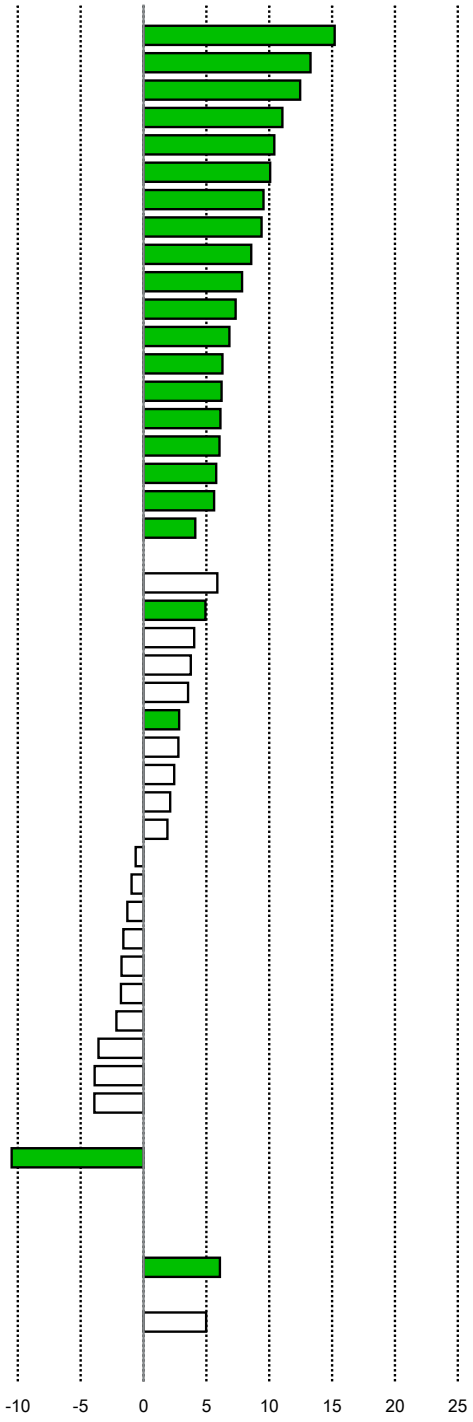
Im Folgenden werden Unterschiede zwischen den naturwissenschaftlichen Leistungen von Mädchen und Jungen im internationalen Vergleich unter Berücksichtigung der Verteilung auf den Kompetenzstufen sowie nach Inhalts- und Anforderungsbereichen getrennt berichtet.

6.5.1 Geschlechterspezifische Leistungsunterschiede auf der Gesamtskala Naturwissenschaften

Die geschlechterbezogene Differenz in den naturwissenschaftlichen Testleistungen im internationalen Vergleich wird in Abbildung 6.7 dargestellt. Der Aufbau der Abbildung erfolgt analog zur Domäne Mathematik (siehe Abschnitt 6.4.1). Die Differenz zwischen den Testleistungen für die Naturwissenschaften zwischen Mädchen und Jungen beträgt in Deutschland 2 Punkte zugunsten von Mädchen, wobei der Unterschied nicht signifikant ist. Für viele andere Teilnehmerstaaten zeigt sich ein ähnlicher Befund (z. B. Norwegen, Rumänien und Bulgarien). In 19 Teilnehmerstaaten (ohne Benchmark-Teilnehmer) hingegen ist die Leistungsdifferenz zugunsten der Jungen signifikant. Weitere Teilnehmerstaaten

Abbildung 6.7: Testleistung nach Geschlecht – Gesamtskala Naturwissenschaften

Teilnehmer	Mädchen %	Mädchen M_M	Mädchen (SE)	Jungen %	Jungen M_J	Jungen (SE)	Differenz[A] $M_J - M_M$	Differenz[A] (SE)	
Republik Korea (Südkorea)	49.7	576	(3.0)	50.3	591	(2.9)	15	(3.0)	▲
Portugal	50.0	504	(2.7)	50.0	517	(2.8)	13	(3.1)	▲
23 *Belgien (Fläm. Gem.)*	48.7	482	(3.2)	51.3	494	(2.8)	12	(2.8)	▲
Macau	48.3	530	(1.3)	51.7	541	(1.9)	11	(2.0)	▲
3 *Hongkong*	48.8	540	(4.6)	51.2	550	(3.9)	10	(3.5)	▲
Australien	52.0	545	(2.1)	48.0	555	(2.4)	10	(2.9)	▲
2 *Singapur*	49.3	603	(2.4)	50.7	612	(2.6)	10	(1.9)	▲
2 *Italien*	48.3	506	(2.6)	51.7	515	(3.0)	9	(2.7)	▲
2 *Frankreich*	50.5	484	(3.3)	49.5	492	(3.1)	9	(2.5)	▲
2 *Zypern*	49.0	483	(3.5)	51.0	491	(3.5)	8	(3.6)	▲
23 *USA*	49.4	529	(2.6)	50.6	536	(2.6)	7	(2.0)	▲
2 *Belgien (Frz. Gem.)*	50.5	477	(3.0)	49.5	484	(2.6)	7	(2.4)	▲
2 *Ungarn*	49.8	521	(3.1)	50.2	527	(3.6)	6	(2.5)	▲
2 *Tschechien*	49.3	523	(2.2)	50.7	529	(2.9)	6	(2.7)	▲
Japan	50.7	552	(2.5)	49.3	558	(2.8)	6	(2.3)	▲
3 *Niederlande*	50.1	514	(3.1)	49.9	520	(3.2)	6	(2.4)	▲
Taiwan	47.7	570	(1.8)	52.3	575	(2.0)	6	(2.5)	▲
23 *Kanada*	50.7	518	(1.8)	49.3	524	(2.0)	6	(1.6)	▲
VG OECD	**49.3**	**524**	**(0.6)**	**50.7**	**528**	**(0.6)**	**4**	**(0.5)**	▲
Slowakei	49.5	518	(4.1)	50.5	523	(3.0)	6	(3.0)	■
2 *Slowenien*	49.2	523	(2.6)	50.8	528	(2.7)	5	(2.5)	■
123 *England*	49.9	555	(2.9)	50.1	559	(3.4)	4	(3.4)	■
2 *Serbien*	50.7	508	(3.3)	49.3	512	(4.0)	4	(3.5)	■
123 *Türkei (5. Jgst.)*	48.1	568	(3.8)	51.9	572	(4.0)	4	(3.8)	■
VG EU	**49.4**	**517**	**(0.7)**	**50.6**	**519**	**(0.7)**	**3**	**(0.6)**	■
123 *Neuseeland*	49.2	516	(3.5)	50.8	518	(3.3)	3	(3.8)	■
2 *Spanien*	49.5	503	(1.9)	50.5	505	(2.3)	2	(1.9)	■
23 *Chile*	47.2	478	(2.9)	52.8	480	(3.1)	2	(3.2)	■
2 *Schweden*	50.9	532	(3.2)	49.1	534	(3.5)	2	(2.7)	■
Internationaler Mittelwert	**49.3**	**495**	**(0.4)**	**50.7**	**494**	**(0.5)**	**-1**	**(0.5)**	■
2 *Polen*	49.8	550	(2.6)	50.2	549	(2.6)	-1	(2.9)	■
23 *Dänemark*	50.9	523	(2.9)	49.1	521	(2.8)	-1	(2.6)	■
Deutschland	48.9	516	(2.8)	51.1	515	(2.9)	-2	(2.5)	■
12 *Norwegen (5. Jgst.)*	50.5	531	(3.0)	49.5	530	(2.7)	-2	(2.7)	■
23 *Rumänien*	49.2	527	(4.9)	50.8	525	(5.2)	-2	(3.4)	■
Bulgarien	47.1	531	(4.8)	52.9	529	(5.4)	-2	(3.8)	■
Lettland	48.6	527	(3.1)	51.4	524	(3.3)	-4	(2.9)	■
2 *Litauen*	49.2	539	(2.6)	50.8	535	(2.8)	-4	(2.2)	■
Irland	48.7	534	(3.9)	51.3	530	(3.5)	-4	(3.8)	■
Finnland	49.3	547	(2.7)	50.7	537	(3.4)	-10	(2.4)	▼
Benchmark-Teilnehmer									
2 *Ontario, Kanada*	51.6	522	(2.8)	48.4	528	(3.0)	6	(2.6)	▲
Dubai, VAE	48.3	560	(2.5)	51.7	565	(1.8)	5	(2.8)	■



Keine statistisch signifikanten Unterschiede zwischen Mädchen und Jungen ($p > .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

Statistisch signifikante Unterschiede zwischen Mädchen und Jungen ($p \leq .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

▲ = Differenzwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)

▼ = Differenzwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)

■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)

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

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

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

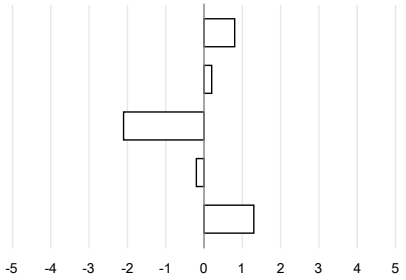
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.

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

© TIMSS 2023

Abbildung 6.8: Mittlere Testleistungen von Mädchen und Jungen und prozentuale Verteilungen auf die Kompetenzstufen in Deutschland – Naturwissenschaften

KS	Mädchen				Jungen				Differenz [A]	
	%	*(SE)*	*M*	*(SE)*	%	*(SE)*	*M*	*(SE)*	$\%_J$-$\%_M$	*(SE)*
V	8.3	(0.7)	636	(28.1) ■	9.1	(0.8)	636	(27.4)	0.8	(1.0)
IV	27.1	(1.2)	577	(8.6) ■	27.2	(1.1)	579	(8.2)	0.2	(1.7)
III	35.5	(1.4)	515	(4.5) ■	33.4	(1.5)	515	(6.4)	-2.1	(2.1)
II	20.6	(1.3)	450	(11.1) ■	20.5	(1.2)	448	(11.5)	-0.2	(1.7)
I	8.5	(0.8)	374	(26.1) ■	9.8	(0.8)	366	(26.1)	1.3	(1.0)



□ Keine statistisch signifikanten Unterschiede zwischen Mädchen und Jungen ($p > .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen
■ Statistisch signifikante Unterschiede zwischen Mädchen und Jungen ($p \leq .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen
■ = Unterschied in den mittleren Leistungswerten zu Jungen nicht statistisch signifikant ($p > .05$)
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

verzeichnen einen nominellen Vorsprung der Mädchen gegenüber den Jungen (u. a. Polen, Dänemark und Irland), diese Geschlechterunterschiede sind jedoch nicht signifikant mit Ausnahme von Finnland. Hier erreichen Jungen im Durchschnitt 10 Leistungspunkte weniger als Mädchen. Die deutlichsten Unterschiede finden sich in der Republik Korea (Südkorea) mit einer Differenz von 15 Punkten und in Portugal mit einer Differenz von 13 Punkten zugunsten der Jungen. Der internationale Mittelwert liegt bei 1 Punkt Unterschied zugunsten der Mädchen, ist jedoch nicht signifikant; der Vergleichswert für die EU liegt bei 3 Punkten Differenz zugunsten der Jungen, der Wert für die OECD bei 4 Punkten, ebenfalls zugunsten der Jungen. Beide Werte sind signifikant. Der Geschlechterunterschied auf der Gesamtskala Naturwissenschaften in Deutschland hingegen ist nicht signifikant und beträgt 2 Punkte.

In Abbildung 6.8 sind die Verteilung der prozentualen Anteile von Mädchen und Jungen auf den fünf Kompetenzstufen sowie die im Mittel erreichten Leistungswerte für Naturwissenschaften abgebildet. Anhand der grafischen Darstellung auf der rechten Seite, aber auch unter Betrachtung der deskriptiven Ergebnisse auf der linken Seite wird ersichtlich, dass keine signifikanten Unterschiede zwischen Mädchen und Jungen hinsichtlich der Anteile auf den Kompetenzstufen vorliegen. In Relation zu den Ergebnissen von TIMSS 2019 scheinen die Unterschiede zwischen Mädchen und Jungen in den Naturwissenschaften auf den oberen Kompetenzstufen IV und V tendenziell weiter abgenommen zu haben. Jedoch konnten auch in TIMSS 2019 keine signifikanten Geschlechterunterschiede für die Verteilung auf den Kompetenzstufen beobachtet werden.

6.5.2 Geschlechterspezifische Leistungsunterschiede nach naturwissenschaftlichen Inhaltsbereichen

Während die Betrachtung der Geschlechterunterschiede auf der Gesamtskala der Naturwissenschaften einen groben Einblick gibt, ermöglicht Abbildung 6.9 durch die Aufteilung in die drei Inhaltsbereiche *Biologie*, *Physik/Chemie* und *Geografie* eine differenziertere Betrachtung (siehe ausführlich in Kapitel 4 in diesem Band). Der Aufbau der Abbildungen 6.9 und 6.10 ist analog zu den Abbildungen zu den mathematischen Inhaltsbereichen (Abbildung 6.3) und kognitiven Anforderungsbereichen (Abbildung 6.4) konzipiert und zeigt die Geschlechterunterschiede von 37 Teilnehmerstaaten sowie zwei Benchmark-Teilnehmern.

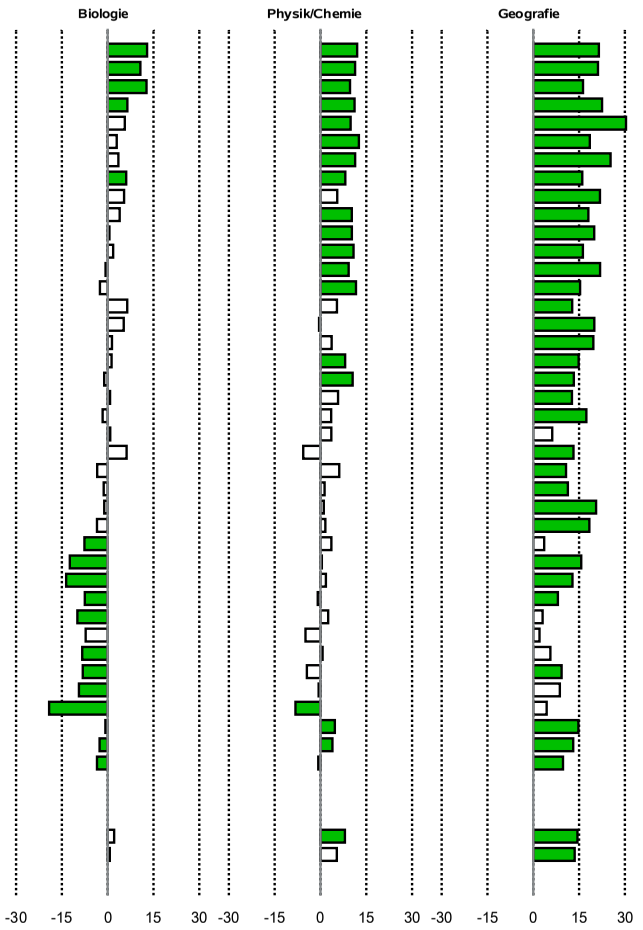
Signifikante Geschlechterunterschiede im Inhaltsbereich *Biologie* zeigen sich bei 14 Teilnehmerstaaten. In fünf Teilnehmerstaaten (u. a. Republik Korea [Südkorea], Portugal und in der flämischen Gemeinschaft in Belgien) fallen die Unterschiede zugunsten der Jungen aus und in neun Teilnehmerstaaten (u. a. Finnland, Irland und Deutschland) zugunsten der Mädchen. Die größte Differenz liegt in Finnland vor, mit 19 Punkten zugunsten der Mädchen. Im internationalen Mittel kann eine Differenz von 4 Punkten zugunsten der Mädchen beobachtet werden, in der Vergleichsgruppe OECD beträgt die Differenz lediglich 1 Punkt und im EU-Durchschnitt 3 Punkte, wobei alle Unterschiede signifikant sind und zugunsten von Mädchen ausfallen. Signifikante Unterschiede können auch in Deutschland beobachtet werden. Hier erreichen Mädchen durchschnittlich 14 Leistungspunkte mehr als Jungen.

Die Betrachtung des Inhaltsbereiches *Physik/Chemie* zeigt insgesamt eine stärkere Tendenz zugunsten der Jungen. Der einzige Teilnehmerstaat mit einer signifikanten Differenz zugunsten von Mädchen ist Finnland. Hier erreichen Mädchen im Mittel 8 Punkte mehr als Jungen. In 15 Teilnehmerstaaten sowie beim Benchmark-Teilnehmer Ontario verzeichnen Jungen signifikant bessere Leistungen. Die größten Differenzen zeigen dabei Australien (13 Punkte) sowie die Republik Korea (Südkorea) und Tschechien (jeweils 12 Punkte), in allen drei Staaten sind die Unterschiede signifikant. Der internationale Mittelwert liegt bei lediglich 1 Punkt Differenz zugunsten der Mädchen. In den Vergleichsgruppen OECD und EU beträgt die Differenz 5 bzw. 4 Punkte zugunsten von Jungen. Diese Differenzen sind signifikant, mit Ausnahme des internationalen Mittelwertes. In Deutschland liegt für den Inhaltsbereich *Physik//Chemie* kein signifikanter Unterschied vor. Die Leistungsdifferenz beträgt 2 Punkte zugunsten der Jungen und liegt so im Bereich des internationalen Mittelwertes.

Die Ergebnisse des Bereiches *Geografie* fallen in allen Teilnehmerstaaten sowie den Benchmark-Teilnehmern zugunsten der Jungen aus. Der Großteil der Geschlechterdifferenzen ist signifikant. Ein besonders großer Geschlechterunterschied liegt in Hongkong vor (30 Punkte), danach folgt Singapur mit 25 Punkten Differenz. Geringe und nicht signifikante Unterschiede finden sich nur vereinzelt (z. B. in Finnland, Irland, Lettland). Der internationale Mittelwert liegt bei 10 Punkten. In der Vergleichsgruppe OECD beträgt die Differenz 15 Punkte, in der EU 13 Punkte. Für Deutschland lässt sich eine signifikante Differenz von 13 Punkten zugunsten der Jungen im Inhaltsbereich *Geografie* feststellen.

Abbildung 6.9: Testleistungen in Inhaltsbereichen nach Geschlecht im internationalen Vergleich – Naturwissenschaften

	Biologie						Physik/Chemie						Geografie					
	Mädchen		Jungen		Differenz^		Mädchen		Jungen		Differenz^		Mädchen		Jungen		Differenz^	
Teilnehmer	M_M	*(SE)*	M_J	*(SE)*	$M_J - M_M$	*(SE)*	M_M	*(SE)*	M_J	*(SE)*	$M_J - M_M$	*(SE)*	M_M	*(SE)*	M_J	*(SE)*	$M_J - M_M$	*(SE)*
Republik Korea (Südkorea)	573	(3.5)	586	(3.0)	13	(3.4)	589	(3.1)	601	(2.9)	12	(3.0)	562	(3.8)	584	(3.9)	21	(3.2)
Portugal	505	(3.1)	516	(3.2)	11	(2.9)	501	(3.3)	512	(3.1)	11	(3.7)	505	(3.4)	526	(3.5)	21	(3.7)
23 *Belgien (Fläm. Gem.)*	485	(3.2)	498	(3.2)	13	(3.7)	478	(3.4)	487	(3.0)	10	(3.2)	482	(4.6)	498	(3.2)	16	(4.2)
Macau	526	(1.8)	532	(1.6)	6	(2.2)	540	(1.5)	551	(1.7)	11	(2.3)	523	(4.0)	545	(1.9)	23	(4.5)
3 *Hongkong*	539	(5.0)	544	(4.3)	6	(4.8)	548	(4.5)	558	(4.6)	10	(4.8)	524	(5.1)	554	(7.1)	30	(7.7)
Australien	553	(2.1)	556	(3.0)	3	(3.1)	539	(2.1)	551	(2.4)	13	(2.6)	540	(3.4)	558	(2.6)	19	(4.2)
2 *Singapur*	612	(2.6)	615	(2.8)	4	(2.0)	616	(2.8)	627	(2.7)	11	(2.8)	565	(3.2)	590	(2.8)	25	(3.8)
2 *Italien*	507	(3.3)	513	(3.1)	6	(2.8)	505	(2.6)	513	(3.7)	8	(3.8)	500	(3.3)	516	(3.6)	16	(3.0)
2 *Frankreich*	485	(4.5)	490	(3.4)	5	(4.0)	483	(4.1)	488	(3.2)	5	(4.5)	479	(5.8)	500	(4.2)	22	(5.7)
2 *Zypern*	490	(3.3)	494	(3.9)	4	(3.6)	479	(3.9)	490	(3.9)	10	(3.6)	460	(3.8)	478	(4.5)	18	(3.6)
23 *USA*	540	(2.6)	541	(2.8)	1	(2.4)	520	(3.3)	530	(2.8)	10	(2.0)	518	(3.3)	538	(4.1)	20	(4.1)
2 *Belgien (Frz. Gem.)*	477	(2.9)	479	(3.3)	2	(2.8)	479	(2.3)	490	(2.6)	11	(2.8)	466	(4.3)	482	(4.1)	16	(5.4)
2 *Ungarn*	528	(3.2)	527	(3.9)	-1	(2.9)	510	(3.3)	519	(3.5)	9	(2.8)	523	(4.1)	545	(4.8)	22	(4.8)
2 *Tschechien*	530	(2.4)	528	(2.4)	-3	(2.9)	519	(2.6)	531	(2.7)	12	(2.6)	512	(3.3)	527	(3.4)	15	(3.3)
Japan	541	(4.2)	547	(3.0)	6	(3.9)	570	(3.0)	576	(4.0)	5	(3.9)	535	(3.0)	548	(3.0)	13	(3.6)
3 *Niederlande*	515	(3.8)	520	(3.5)	5	(3.2)	509	(3.2)	509	(3.4)	0	(4.4)	517	(3.9)	537	(3.0)	20	(3.9)
Taiwan	561	(2.1)	563	(2.3)	1	(2.1)	580	(1.9)	584	(1.6)	4	(1.9)	563	(2.6)	583	(2.9)	20	(3.2)
23 *Kanada*	525	(2.1)	527	(2.6)	1	(2.5)	512	(1.8)	520	(2.1)	8	(2.6)	513	(2.2)	528	(3.0)	15	(2.6)
Slowakei	519	(4.3)	518	(4.0)	-1	(3.7)	518	(4.6)	528	(4.0)	11	(4.8)	512	(4.5)	525	(4.2)	13	(4.4)
2 *Slowenien*	519	(2.9)	520	(2.8)	1	(3.2)	530	(2.6)	536	(3.7)	6	(4.3)	517	(3.4)	530	(3.7)	13	(4.3)
123 *England*	556	(3.3)	555	(3.8)	-2	(3.7)	556	(4.1)	560	(3.7)	3	(4.5)	546	(4.9)	563	(4.1)	17	(5.5)
2 *Serbien*	505	(4.1)	505	(4.7)	1	(3.8)	518	(3.6)	522	(4.2)	4	(4.0)	495	(4.3)	501	(3.7)	6	(5.0)
123 *Türkei (5. Jgst.)*	551	(3.8)	557	(4.4)	6	(4.1)	592	(4.5)	586	(4.6)	-6	(4.8)	569	(4.8)	582	(5.0)	13	(4.0)
123 *Neuseeland*	521	(3.7)	518	(3.2)	-3	(3.7)	509	(4.1)	515	(3.0)	6	(3.9)	514	(3.5)	525	(3.0)	11	(4.4)
2 *Spanien*	503	(2.2)	502	(2.5)	-1	(1.8)	504	(2.2)	506	(2.9)	1	(2.9)	499	(2.0)	510	(3.4)	11	(2.9)
23 *Chile*	487	(2.8)	486	(2.8)	-1	(3.1)	471	(2.9)	472	(3.1)	1	(3.2)	461	(3.2)	482	(4.5)	21	(4.8)
2 *Schweden*	534	(3.7)	530	(3.5)	-4	(2.7)	531	(3.4)	532	(3.9)	2	(3.3)	528	(4.8)	547	(5.3)	18	(5.9)
2 *Polen*	553	(3.3)	546	(3.3)	-8	(3.8)	547	(3.3)	551	(2.9)	4	(4.1)	550	(3.4)	553	(3.4)	4	(4.2)
23 *Dänemark*	537	(2.9)	524	(3.3)	-12	(3.2)	510	(3.3)	510	(3.3)	0	(3.2)	517	(3.6)	533	(3.2)	16	(3.4)
Deutschland	522	(3.1)	509	(2.5)	-14	(2.7)	515	(3.0)	516	(2.7)	2	(2.4)	505	(4.3)	518	(3.8)	13	(5.2)
12 *Norwegen (5. Jgst.)*	538	(3.2)	530	(3.5)	-7	(3.1)	521	(3.1)	520	(3.5)	-1	(3.5)	539	(2.7)	548	(3.9)	8	(3.2)
23 *Rumänien*	529	(5.3)	519	(5.8)	-10	(4.3)	529	(4.8)	531	(6.1)	3	(4.5)	525	(5.6)	528	(6.4)	3	(6.1)
Bulgarien	533	(5.6)	526	(5.5)	-7	(4.7)	529	(5.0)	524	(5.8)	-5	(4.5)	534	(6.1)	536	(6.5)	2	(4.3)
Lettland	523	(4.3)	514	(3.1)	-8	(3.5)	533	(3.6)	533	(4.7)	1	(4.4)	519	(3.9)	525	(4.9)	6	(4.8)
2 *Litauen*	535	(2.5)	527	(3.0)	-8	(3.4)	546	(2.4)	542	(3.5)	-4	(3.6)	530	(2.9)	539	(4.3)	9	(4.6)
Irland	540	(4.4)	530	(3.9)	-9	(4.4)	528	(4.0)	527	(4.0)	-1	(4.3)	530	(4.7)	538	(4.9)	9	(5.2)
Finnland	553	(3.3)	534	(3.7)	-19	(4.2)	542	(2.7)	533	(3.2)	-8	(2.6)	548	(3.0)	553	(3.6)	4	(3.3)
VG OECD	**526**	**(0.6)**	**525**	**(0.6)**	**-1**	**(0.6)**	**523**	**(0.6)**	**528**	**(0.6)**	**5**	**(0.7)**	**518**	**(0.7)**	**533**	**(0.7)**	**15**	**(0.8)**
VG EU	**519**	**(0.8)**	**517**	**(0.8)**	**-3**	**(0.7)**	**515**	**(0.7)**	**519**	**(0.8)**	**4**	**(0.8)**	**512**	**(0.9)**	**525**	**(0.9)**	**13**	**(1.0)**
Internationaler Mittelwert	**497**	**(0.5)**	**494**	**(0.5)**	**-4**	**(0.5)**	**500**	**(0.5)**	**499**	**(0.5)**	**-1**	**(0.6)**	**491**	**(0.6)**	**501**	**(0.6)**	**10**	**(0.7)**
Benchmark-Teilnehmer																		
2 *Ontario, Kanada*	532	(3.0)	534	(3.4)	2	(3.4)	516	(2.9)	524	(2.9)	8	(3.5)	515	(3.3)	529	(4.2)	14	(3.4)
Dubai, VAE	554	(2.3)	555	(1.7)	1	(2.9)	568	(2.8)	573	(2.4)	5	(3.4)	562	(3.4)	576	(2.5)	14	(3.4)



Keine statistisch signifikanten Unterschiede zwischen Mädchen und Jungen ($p > .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

Statistisch signifikante Unterschiede zwischen Mädchen und Jungen ($p \leq .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

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

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

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

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

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

© TIMSS 2023

6.5.3 Geschlechterspezifische Unterschiede nach kognitiven Anforderungsbereichen in den Naturwissenschaften

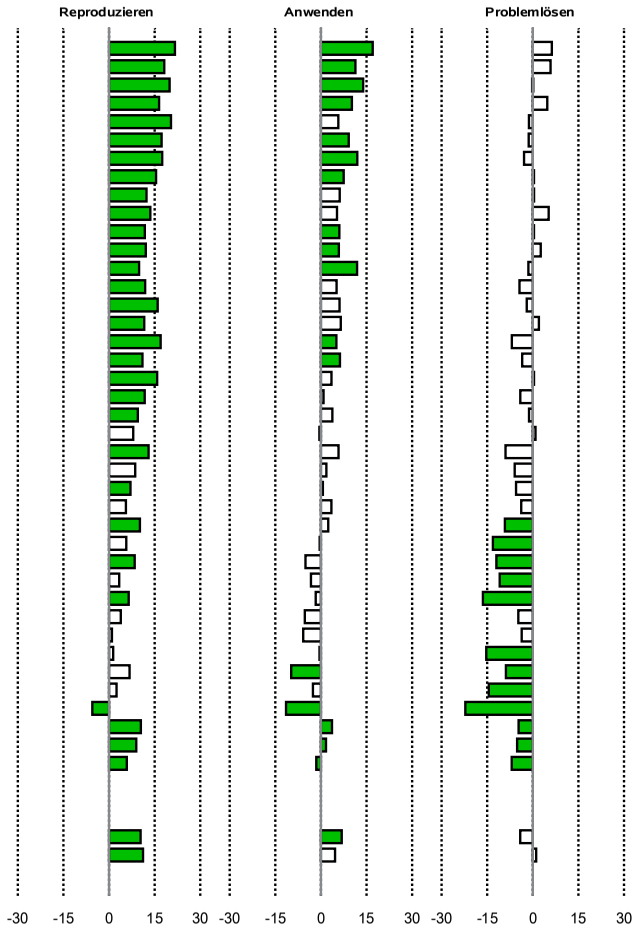
In den Naturwissenschaften wird, wie bereits bei der Rahmenkonzeption Mathematik, zwischen den drei kognitiven Anforderungsbereichen *Reproduzieren*, *Anwenden* und *Problemlösen* differenziert. Die Ergebnisse der Geschlechterunterschiede im internationalen Vergleich sind in Abbildung 6.10 dargestellt. Im Anforderungsbereich *Reproduzieren* liegen in beinahe allen Teilnehmerstaaten (inklusive der Benchmark-Teilnehmer) signifikante Differenzen zugunsten der Jungen vor. Lediglich Finnland weist mit 6 Leistungspunkten Differenz einen signifikanten Unterschied zugunsten der Mädchen auf. Größere Differenzen zugunsten der Jungen lassen sich für die Republik Korea (Südkorea) mit 22 Punkten, aber auch in der flämischen Gemeinschaft Belgiens und Hongkong (jeweils 20 Punkte) feststellen. Die Differenz im internationalen Durchschnitt ist signifikant und entspricht 6 Punkten. Die Vergleichsgruppen weisen ähnliche, statistisch relevante Differenzwerte auf (OECD: 10 Punkte, EU: 9 Punkte). Deutschland liegt mit 3 Punkten Differenz zugunsten der Jungen unterhalb der Vergleichswerte der OECD und der EU. Die Differenz der Testleistungen von Mädchen und Jungen in Deutschland ist zudem nicht signifikant.

Für den Anforderungsbereich *Anwenden* liegt bei der überwiegenden Mehrzahl der Teilnehmerstaaten (inklusive Benchmark-Teilnehmern) ein Vorsprung der Jungen vor (u. a. in Portugal, Macau, Singapur), wobei nur ein Teil dieser Unterschiede signifikant ist. Die größte Differenz zugunsten der Jungen weist die Republik Korea (Südkorea) mit 17 Punkten auf. In Finnland wiederum liegen die Mädchen mit 11 Leistungspunkten vor den Jungen, aber auch Litauen weist mit 10 Punkten eine statistisch signifikante Differenz zugunsten von Mädchen auf. Auch im internationalen Durchschnitt erreichen die Mädchen 2 Punkte mehr. Dieser Unterschied ist signifikant. Die Vergleichswerte der OECD (4 Punkte) und der EU (2 Punkte) fallen jedoch zugunsten der Jungen aus. In Deutschland erreichen die Mädchen zwar 3 Punkte mehr als Jungen, dieser Unterschied ist jedoch nicht signifikant.

Die Ergebnisse für den Bereich *Problemlösen* zeigen ein uneinheitliches Bild. Lediglich neun Teilnehmerstaaten weisen einen signifikanten Geschlechterunterschied auf. Diese statistisch relevanten Unterschiede fallen allesamt zugunsten der Mädchen aus. Die größte Differenz ist erneut für Finnland mit 22 Punkten Vorsprung auf Seiten der Mädchen feststellbar. Aber auch Norwegen (16 Punkte), Lettland (15 Punkte) und Irland (14 Punkte) verzeichnen Differenzwerte zugunsten der Mädchen. Obgleich in vielen Teilnehmerstaaten nur sehr geringe sowie nicht signifikante Geschlechterunterschiede im Anforderungsbereich *Problemlösen* festgestellt werden können, liegen im internationalen Durchschnitt Mädchen mit 7 Punkten Differenz signifikant vor den Jungen. Die Vergleichswerte der OECD und der EU weisen jeweils 5 Punkte Differenz zugunsten der Mädchen auf. Alle Differenzen der Vergleichsgruppen sind statistisch signifikant. In Deutschland erreichen Mädchen 11 Leistungspunkte mehr als Jungen. Hier liegt ein statistisch relevanter Geschlechterunterschied vor, der über den Vergleichswerten der OECD und der EU einzuordnen ist.

Abbildung 6.10: Testleistungen in kognitiven Anforderungsbereichen nach Geschlecht im internationalen Vergleich – Naturwissenschaften

	Reproduzieren						Anwenden						Problemlösen					
	Mädchen		Jungen		Differenz^		Mädchen		Jungen		Differenz^		Mädchen		Jungen		Differenz^	
Teilnehmer	M_M	*(SE)*	M_J	*(SE)*	$M_J - M_M$	*(SE)*	M_M	*(SE)*	M_J	*(SE)*	$M_J - M_M$	*(SE)*	M_M	*(SE)*	M_J	*(SE)*	$M_J - M_M$	*(SE)*
Republik Korea (Südkorea)	573	(3.2)	595	(3.7)	22	(4.0)	575	(2.9)	592	(3.5)	17	(4.0)	580	(2.7)	587	(3.3)	6	(3.6)
Portugal	497	(3.8)	515	(2.4)	18	(3.6)	505	(4.1)	516	(3.4)	11	(2.8)	509	(3.5)	515	(2.9)	6	(3.3)
2,3 *Belgien (Fläm. Gem.)*	478	(4.1)	498	(4.0)	20	(4.5)	477	(3.3)	491	(2.8)	14	(3.4)	498	(4.1)	497	(3.4)	0	(3.5)
Macau	530	(1.4)	546	(2.6)	16	(2.7)	529	(1.3)	539	(2.7)	10	(3.0)	531	(2.2)	536	(2.0)	5	(3.3)
3 *Hongkong*	537	(5.3)	558	(5.1)	20	(3.7)	540	(4.3)	545	(4.3)	6	(3.8)	543	(5.3)	542	(4.8)	-1	(4.3)
Australien	544	(2.5)	561	(2.8)	17	(2.5)	544	(2.0)	553	(2.6)	9	(2.6)	548	(3.2)	547	(3.9)	-1	(3.0)
2 *Singapur*	602	(3.0)	619	(2.9)	17	(3.4)	589	(2.6)	601	(2.6)	12	(2.9)	623	(2.2)	620	(2.6)	-3	(1.8)
2 *Italien*	503	(3.4)	519	(3.1)	15	(2.9)	505	(3.0)	512	(3.1)	7	(2.4)	506	(4.3)	506	(2.7)	0	(4.7)
2 *Frankreich*	485	(3.7)	497	(3.0)	12	(3.7)	484	(4.2)	490	(3.7)	6	(3.8)	479	(4.0)	479	(3.7)	0	(3.6)
2 *Zypern*	473	(3.7)	487	(4.7)	14	(4.8)	490	(3.1)	495	(3.7)	5	(3.2)	484	(3.2)	489	(3.5)	5	(3.5)
2,3 *USA*	529	(3.2)	541	(3.1)	12	(2.5)	527	(2.6)	533	(3.2)	6	(2.4)	530	(3.1)	530	(2.9)	0	(3.0)
2 *Belgien (Frz. Gem.)*	472	(3.9)	485	(2.2)	12	(3.5)	477	(3.8)	483	(2.5)	6	(2.8)	481	(3.2)	484	(2.4)	3	(3.4)
2 *Ungarn*	517	(4.1)	527	(4.1)	10	(3.8)	518	(3.2)	530	(4.6)	12	(3.7)	527	(5.2)	525	(4.5)	-1	(6.8)
2 *Tschechien*	518	(3.1)	530	(2.8)	12	(3.2)	520	(2.6)	526	(2.6)	5	(2.8)	531	(4.2)	527	(2.9)	-4	(4.5)
Japan	530	(3.9)	546	(3.4)	16	(3.1)	557	(3.3)	563	(3.2)	6	(3.1)	569	(2.6)	567	(3.1)	-2	(2.7)
3 *Niederlande*	512	(4.0)	524	(2.9)	12	(3.1)	510	(3.8)	517	(2.7)	7	(4.2)	519	(4.1)	521	(3.9)	2	(5.0)
Taiwan	571	(2.1)	588	(1.8)	17	(2.1)	566	(2.8)	571	(2.2)	5	(2.2)	570	(2.2)	564	(2.8)	-7	(3.9)
2,3 *Kanada*	522	(2.1)	532	(2.3)	11	(1.8)	517	(2.2)	523	(2.7)	6	(2.2)	517	(2.8)	513	(3.3)	-3	(3.2)
Slowakei	516	(4.1)	532	(3.3)	16	(3.2)	518	(4.5)	522	(3.3)	3	(3.5)	516	(5.7)	517	(4.4)	0	(4.7)
2 *Slowenien*	518	(2.6)	530	(2.6)	12	(2.9)	533	(2.6)	534	(3.5)	1	(3.3)	514	(2.9)	510	(3.1)	-4	(2.9)
1,2,3 *England*	550	(3.3)	560	(4.2)	9	(4.0)	556	(3.4)	560	(3.7)	4	(3.4)	557	(3.4)	555	(3.5)	-1	(4.1)
2 *Serbien*	498	(3.7)	505	(3.6)	8	(4.2)	514	(4.2)	514	(4.1)	-1	(4.6)	513	(4.8)	514	(4.5)	1	(5.3)
1,2,3 *Türkei (5. Jgst.)*	565	(4.6)	578	(4.5)	13	(4.2)	564	(4.7)	570	(4.2)	6	(4.0)	579	(4.8)	570	(4.4)	-9	(4.7)
1,2,3 *Neuseeland*	516	(3.4)	525	(3.1)	9	(4.8)	513	(3.9)	515	(3.3)	2	(4.9)	517	(3.5)	511	(3.8)	-6	(4.0)
2 *Spanien*	505	(1.8)	512	(2.7)	7	(2.3)	501	(2.1)	502	(2.1)	1	(2.2)	501	(2.8)	495	(2.5)	-6	(3.0)
2,3 *Chile*	476	(3.4)	482	(4.0)	6	(4.2)	477	(3.1)	480	(3.0)	3	(3.9)	479	(2.8)	475	(3.8)	-4	(3.8)
2 *Schweden*	518	(3.5)	528	(4.5)	10	(3.3)	531	(3.3)	533	(3.5)	2	(2.7)	551	(3.5)	542	(4.1)	-9	(3.5)
2 *Polen*	545	(2.6)	551	(2.6)	6	(3.5)	553	(2.9)	553	(2.7)	0	(3.3)	553	(2.8)	540	(3.3)	-13	(3.9)
2,3 *Dänemark*	516	(2.8)	525	(3.6)	8	(3.4)	523	(3.0)	518	(4.3)	-5	(3.4)	534	(2.8)	522	(3.7)	-12	(3.9)
Deutschland	513	(2.4)	517	(3.2)	3	(3.6)	514	(2.9)	510	(2.4)	-3	(2.4)	521	(3.8)	511	(2.5)	-11	(3.6)
1,2 *Norwegen (5. Jgst.)*	525	(3.0)	532	(2.8)	6	(3.1)	531	(2.7)	530	(2.9)	-2	(3.0)	542	(3.5)	526	(3.7)	-16	(3.7)
2,3 *Rumänien*	522	(4.7)	525	(5.8)	4	(4.8)	524	(5.0)	518	(5.7)	-5	(4.4)	542	(4.7)	538	(6.2)	-5	(5.3)
Bulgarien	535	(5.1)	535	(5.8)	1	(4.8)	531	(4.9)	525	(6.0)	-6	(4.7)	525	(5.6)	522	(6.0)	-4	(5.1)
Lettland	521	(3.0)	522	(3.9)	1	(4.0)	526	(3.5)	525	(4.4)	0	(4.6)	537	(4.0)	522	(4.2)	-15	(4.4)
2 *Litauen*	533	(2.4)	540	(3.6)	7	(4.2)	538	(2.3)	528	(3.0)	-10	(3.2)	548	(2.4)	539	(2.5)	-9	(2.8)
Irland	532	(4.0)	535	(3.5)	2	(4.2)	531	(4.8)	529	(3.7)	-3	(4.3)	538	(4.0)	524	(4.0)	-14	(4.8)
Finnland	544	(2.8)	539	(3.3)	-6	(2.8)	551	(3.0)	540	(3.4)	-11	(3.0)	551	(3.0)	529	(3.7)	-22	(3.6)
VG OECD	**520**	**(0.6)**	**530**	**(0.6)**	**10**	**(0.7)**	**523**	**(0.6)**	**527**	**(0.6)**	**4**	**(0.6)**	**528**	**(0.7)**	**523**	**(0.7)**	**-5**	**(0.7)**
VG EU	**512**	**(0.8)**	**521**	**(0.8)**	**9**	**(0.8)**	**516**	**(0.8)**	**518**	**(0.8)**	**2**	**(0.7)**	**521**	**(0.8)**	**516**	**(0.8)**	**-5**	**(0.9)**
Internationaler Mittelwert	**494**	**(0.5)**	**500**	**(0.5)**	**6**	**(0.6)**	**497**	**(0.5)**	**495**	**(0.5)**	**-2**	**(0.5)**	**501**	**(0.5)**	**494**	**(0.5)**	**-7**	**(0.6)**
Benchmark-Teilnehmer																		
2 *Ontario, Kanada*	527	(2.7)	537	(3.4)	10	(2.7)	524	(2.9)	530	(3.6)	7	(3.1)	520	(3.5)	516	(4.3)	-4	(3.7)
Dubai, VAE	572	(2.9)	583	(1.7)	11	(3.5)	558	(2.6)	563	(1.7)	5	(3.1)	546	(2.5)	547	(2.2)	1	(3.3)



□ Keine statistisch signifikanten Unterschiede zwischen Mädchen und Jungen ($p > .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

■ Statistisch signifikante Unterschiede zwischen Mädchen und Jungen ($p \leq .05$); Linksseitige Ausprägung = Vorsprung der Mädchen; Rechtsseitige Ausprägung = Vorsprung der Jungen

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

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

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

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

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

© TIMSS 2023

6.5.4 Trends in geschlechterspezifischen Leistungsunterschieden in den Naturwissenschaften

Die Entwicklung der Leistungen von Mädchen und Jungen in den Naturwissenschaften im internationalen Vergleich über die Zyklen 2007, 2019 und 2023 hinweg ist in Abbildung 6.11 abgebildet. Diese ist analog zur Abbildung 6.5 in Abschnitt 6.4.4 aufgebaut.

Die Betrachtung der Leistungsentwicklung der Mädchen im internationalen Vergleich zeigt, dass in dreizehn der hier betrachteten Teilnehmerstaaten im Trend seit 2007 ein Zuwachs an Leistungspunkten zu verzeichnen ist (Abbildung 6.11). Für neun Staaten handelt es sich dabei um einen statistisch relevanten Zuwachs, wobei Georgien mit einem deutlichen Plus von 40 Leistungspunkten seit 2007 das Land mit dem größten Zuwachs darstellt. Die größte signifikante Zunahme an Leistungspunkten von Mädchen im Vergleich von 2019 zu 2023 können England und Australien verzeichnen (jeweils 22 Punkte). Als Benchmark-Teilnehmer weist auch Dubai, ähnlich wie im mathematischen Bereich, eine enorme Leistungssteigerung seit 2007 mit 92 Punkten in den Naturwissenschaften auf. Von 2019 bis 2023 haben Mädchen hier 19 Leistungspunkte dazu gewonnen. Bei insgesamt fünf Teilnehmerstaaten und zwei Benchmark-Teilnehmern zeigen sich im Vergleich zu 2007 rückläufige Entwicklungen für die Naturwissenschaften. Signifikant werden die Rückläufe jedoch nur in Italien mit einem Punkteverlust bei den Mädchen von 14 Punkten im Zeitraum von 2007 bis 2023 (wobei seit 2019 wieder 9 Punkte mehr erreicht wurden) und dem Iran, wo ein Verlust von 17 Leistungspunkten seit 2019 verzeichnet werden kann. Mädchen in Deutschland erreichten 2007 noch 520 Punkte, 2019 waren es 516 Punkte und 2023 515 Punkte. Dieser Trend ist somit zwar leicht rückläufig, jedoch nicht statistisch signifikant.

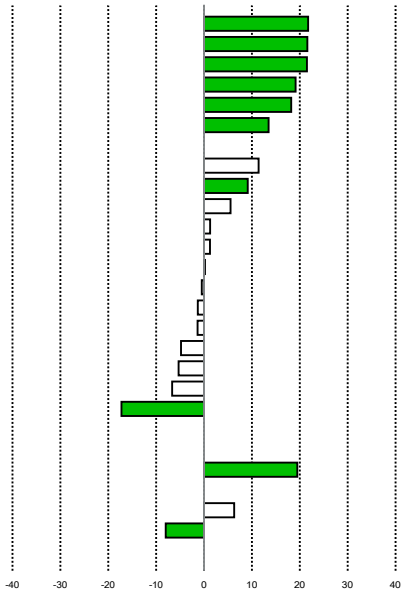
Das Gesamtbild der Leistungsentwicklung von Jungen in den Naturwissenschaften in den Jahren 2007, 2019 und 2023 zeigt, dass, stärker als in Mathematik, neben signifikanten positiven Entwicklungen, für manche Teilnehmerstaaten (auch Benchmark-Teilnehmer) ein rückläufiger Trend zu verzeichnen ist. Insgesamt weisen neun Teilnehmerstaaten und ein Benchmark-Teilnehmer im Vergleich zu 2007 positive Leistungstrends auf, wobei für Dubai mit 117 Leistungspunkten Zugewinn seit 2007 der größte Zuwachs vorliegt. Auch Jungen in Georgien haben seit 2007 insgesamt 50 Leistungspunkte mehr erreicht, wobei der Zugewinn seit 2019 nur bei 6 Punkten liegt. Ein deutlicher Rückgang im Vergleich zu 2007 ist für Jungen in Italien festzustellen, wo diese 26 Punkte weniger aufweisen. Im Vergleich zu 2019 zeigt sich der größte Rückgang für den Iran mit einer Abnahme von 19 Punkten. Auch in Deutschland haben Jungen seit 2007 durchschnittlich weniger Leistungspunkte erreicht. Im Vergleich zu 2019 sind die Leistungen 2023 um 6 Punkte abgesunken, im Verhältnis zu 2007 beträgt der Rückgang sogar 20 Punkte. Dieser negative Trend ist signifikant.

Eine Betrachtung der Geschlechterdifferenzen in den Naturwissenschaften in den Jahren 2007, 2019 und 2023 (Abbildung 6.12) verweist lediglich für Singapur und Japan auf eine signifikante Zunahme der Geschlechterunterschiede seit dem Jahr 2007. Im Vergleich zu 2019 zeigt sich eine signifikante Zunahme für Japan, Australien und Hongkong. Eine signifikante Abnahme von Unterschieden seit 2007 findet sich hingegen nur in Deutschland. Hier haben sich die Unterschiede in den Testleistungen von Jungen und Mädchen von 15 Punkten im Jahr 2007 zugunsten der Jungen auf 2 Punkte zugunsten der Mädchen verringert.

Abbildung 6.11: Vergleich der Leistungsentwicklungen von Mädchen und Jungen in Naturwissenschaften in TIMSS 2007, 2019 und 2023 (internationaler Vergleich)

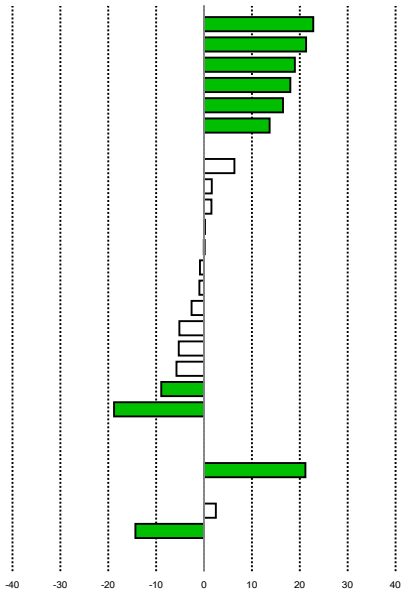
Mädchen

Fußnoten 2007	2019	2023	T	Teilnehmer	2007 M_{07}	(SE)	2019 M_{19}	(SE)	2023 M_{23}	(SE)	M_{23}-M_{07}	Veränderungen[A] (SE)		M_{23}-M_{19}	(SE)	
1 3	1 2	1 2 3	8	England	543	(3.1)	537	(3.6)	559	(3.4)	16	(4.6)	+	22	(5.0)	+ ▲
				Australien	525	(4.0)	533	(2.9)	555	(2.4)	30	(4.7)	+	22	(3.8)	+ ▲
	2	2	8	Singapur	587	(4.4)	591	(3.6)	612	(2.6)	25	(5.0)	+	21	(4.4)	+ ▲
2 3		3	3 8	Hongkong	553	(3.6)	531	(3.1)	550	(3.9)	-2	(5.3)		19	(5.0)	+ ▲
				Taiwan	556	(2.3)	557	(2.0)	575	(2.0)	20	(3.0)	+	18	(2.8)	+ ▲
1 2	1 2	1 2 3	8	Neuseeland	506	(2.8)	505	(3.2)	518	(3.3)	12	(4.4)	+	13	(4.7)	+ ▲
2	2	2		Georgien	423	(4.7)	452	(4.7)	463	(3.5)	40	(5.9)	+	11	(5.9)	■
2		2	8	Italien	529	(3.2)	506	(3.3)	515	(3.0)	-14	(4.4)	–	9	(4.4)	+ ■
	2			Slowakei	521	(5.2)	518	(3.8)	523	(3.0)	2	(6.0)		6	(4.9)	■
3	3	3	8	Niederlande	518	(3.0)	519	(3.1)	520	(3.2)	2	(4.3)		1	(4.4)	■
		2		Ungarn	535	(4.4)	526	(3.2)	527	(3.6)	-8	(5.7)		1	(4.8)	■
2 3	2 3	2 3		USA	536	(3.0)	536	(3.0)	536	(2.6)	0	(4.0)		0	(3.9)	■
		2		Tschechien	511	(3.7)	529	(3.0)	529	(2.9)	18	(4.7)	+	0	(4.2)	■
				Deutschland	520	(2.6)	516	(2.8)	515	(2.9)	-5	(3.9)		-1	(4.0)	■
3	3	2 3	8	Dänemark	514	(3.2)	523	(2.7)	521	(2.8)	7	(4.3)		-1	(3.9)	■
	2	2		Schweden	526	(2.7)	538	(3.6)	534	(3.5)	8	(4.4)		-5	(5.0)	■
2	2	2	8	Litauen	516	(2.7)	540	(2.8)	535	(2.8)	19	(3.9)	+	-5	(3.9)	■
	3		8	Japan	548	(2.5)	565	(2.0)	558	(2.8)	10	(3.8)	+	-7	(3.5)	■
			8	Iran	443	(5.6)	440	(6.6)	423	(5.5)	-20	(7.9)	–	-17	(8.6)	– ■
				Benchmark-Teilnehmer												
2 3 6	2		8	Dubai, VAE	473	(4.5)	545	(3.6)	565	(1.8)	92	(4.9)	+	19	(4.0)	+ ▲
2		2	8	Ontario, Kanada	532	(4.1)	522	(4.0)	528	(3.0)	-4	(5.1)		6	(5.0)	■
2		3	2 3 8	Québec, Kanada	516	(3.1)	519	(2.7)	511	(2.5)	-6	(3.9)		-8	(3.6)	– ■



Jungen

Fußnoten 2007	2019	2023	T	Teilnehmer	2007 M_{07}	(SE)	2019 M_{19}	(SE)	2023 M_{23}	(SE)	M_{23}-M_{07}	Veränderungen[A] (SE)		M_{23}-M_{19}	(SE)	
				Australien	530	(3.5)	532	(2.7)	555	(2.4)	25	(4.3)	+	23	(3.7)	+ ▲
1 3	1 2	1 2 3	8	England	540	(3.4)	537	(2.7)	559	(3.4)	18	(4.8)	+	21	(4.4)	+ ▲
2 3		3	3 8	Hongkong	556	(4.3)	531	(4.3)	550	(3.9)	-5	(5.8)		19	(5.8)	+ ▲
1 2	1 2	1 2 3	8	Neuseeland	502	(3.5)	500	(2.8)	518	(3.3)	17	(4.8)	+	18	(4.4)	+ ▲
				Taiwan	558	(2.4)	559	(2.2)	575	(2.0)	18	(3.1)	+	16	(3.0)	+ ▲
	2	2	8	Singapur	587	(4.4)	598	(3.8)	612	(2.6)	25	(5.1)	+	14	(4.5)	+ ▲
2	2	2		Georgien	413	(5.1)	457	(4.2)	463	(3.5)	50	(6.2)	+	6	(5.5)	■
3	3	3	8	Niederlande	528	(2.8)	518	(3.3)	520	(3.2)	-8	(4.2)	–	2	(4.6)	■
2		2	8	Italien	541	(3.7)	514	(3.3)	515	(3.0)	-26	(4.7)	–	2	(4.5)	■
	2			Slowakei	530	(4.8)	523	(4.4)	523	(3.0)	-6	(5.7)		0	(5.4)	■
3	3	2 3	8	Dänemark	520	(3.6)	522	(2.8)	521	(2.8)	2	(4.6)		0	(3.9)	■
2	2	2	8	Litauen	512	(2.9)	536	(3.3)	535	(2.8)	23	(4.0)	+	-1	(4.3)	■
	3		8	Japan	547	(2.4)	559	(2.1)	558	(2.8)	11	(3.7)	+	-1	(3.5)	■
	2	2		Schweden	524	(3.7)	536	(3.8)	534	(3.5)	10	(5.1)		-3	(5.2)	■
		2		Ungarn	538	(3.6)	533	(3.1)	527	(3.6)	-10	(5.1)	–	-5	(4.7)	■
2 3	2 3	2 3		USA	541	(3.1)	541	(3.2)	536	(2.6)	-5	(4.0)		-5	(4.1)	■
				Deutschland	535	(2.9)	520	(2.4)	515	(2.9)	-20	(4.1)	–	-6	(3.8)	■
		2		Tschechien	518	(3.4)	538	(3.0)	529	(2.9)	11	(4.5)	+	-9	(4.2)	– ■
			8	Iran	429	(6.0)	442	(5.4)	423	(5.5)	-6	(8.2)		-19	(7.7)	– ■
				Benchmark-Teilnehmer												
2 3 6	2		8	Dubai, VAE	448	(4.9)	544	(2.4)	565	(1.8)	117	(5.3)	+	21	(3.1)	+ ▲
2		2	8	Ontario, Kanada	539	(4.3)	526	(3.3)	528	(3.0)	-11	(5.3)	–	2	(4.4)	■
2		3	2 3 8	Québec, Kanada	518	(3.5)	525	(3.0)	511	(2.5)	-7	(4.3)		-14	(3.9)	– ■

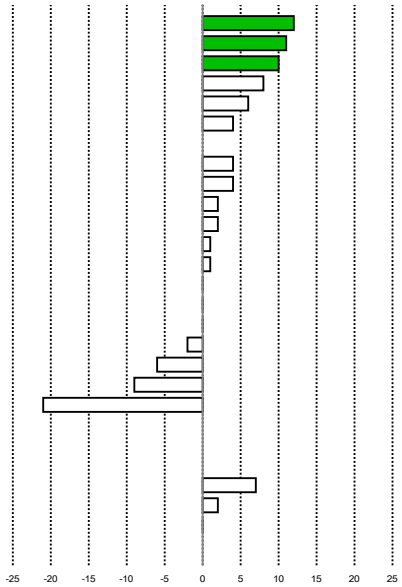


Keine statistisch signifikanten Unterschiede zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Unterschiede zwischen 2019 und 2023 ($p \leq .05$)
▲ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
\+ = Mittelwert in 2023 signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Abbildung 6.12: Vergleich der Leistungsunterschiede in den Naturwissenschaften von Mädchen und Jungen in TIMSS 2007, 2019 und 2023 (internationaler Vergleich)

Fußnoten 2007	2019	2023	T	Teilnehmer	2007 M_{07}	(SE)	2019 M_{19}	(SE)	2023 M_{23}	(SE)	Veränderungen^ M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)		
	3		8	Japan	-1	(2.6)	-6	(2.0)	6	(2.3)	7	(3.5)	+	12	(3.0)	+	▲
				Australien	5	(3.5)	-1	(2.9)	10	(2.9)	5	(4.5)		11	(4.1)	+	▲
2 3	3		3 8	Hongkong	3	(3.7)	0	(3.6)	10	(3.5)	7	(5.1)		10	(5.0)	+	▲
1 2	1 2	1 2	3 8	Neuseeland	-4	(3.5)	-5	(3.9)	3	(3.8)	7	(5.2)		8	(5.4)		▲
3	3		3 8	Niederlande	11	(2.5)	0	(2.8)	6	(2.4)	-5	(3.5)		6	(3.7)		▲
				Taiwan	2	(2.5)	2	(2.3)	6	(2.5)	4	(3.5)		4	(3.4)		▲
	2	2		Schweden	-2	(2.9)	-2	(3.3)	2	(2.7)	4	(4.0)		4	(4.3)		■
1 3	1 2	1 2	3 8	England	-3	(3.0)	0	(3.5)	4	(3.4)	7	(4.5)		4	(4.9)		■
	2	2	8	Singapur	0	(3.0)	8	(2.8)	10	(1.9)	10	(3.6)	+	2	(3.4)		■
2 3	2 3	2 3		USA	5	(2.7)	5	(2.7)	7	(2.0)	2	(3.4)		2	(3.4)		■
2		2	8	Italien	13	(2.6)	8	(2.8)	9	(2.7)	-4	(3.7)		1	(3.9)		■
	2			Slowakei	8	(2.9)	5	(3.8)	6	(3.0)	-2	(4.2)		1	(4.8)		■
3	3	2 3	8	Dänemark	6	(3.9)	-1	(2.8)	-1	(2.6)	-7	(4.7)		0	(3.8)		■
		2		Ungarn	3	(4.5)	6	(3.3)	6	(2.5)	3	(5.1)		0	(4.1)		■
2	2	2	8	Litauen	-4	(3.0)	-4	(3.4)	-4	(2.2)	0	(3.7)		0	(4.0)		■
		2		Tschechien	7	(3.3)	8	(3.1)	6	(2.7)	-1	(4.3)		-2	(4.1)		■
				Deutschland	15	(2.7)	4	(2.8)	-2	(2.5)	-17	(3.7)	–	-6	(3.8)		■
2	2	2		Georgien	-10	(3.6)	5	(4.1)	-4	(3.1)	6	(4.8)		-9	(5.1)		■
			8	Iran	-14	(7.9)	2	(8.7)	-19	(7.0)	-5	(10.6)		-21	(11.2)		■
				Benchmark-Teilnehmer													
2 3 6	2		8	Dubai, VAE	-26	(8.3)	-2	(5.0)	5	(2.8)	31	(8.8)	+	7	(5.7)		■
2		2	8	Ontario, Kanada	7	(4.0)	4	(3.7)	6	(2.6)	-1	(4.8)		2	(4.5)		■
2	3	2 3	8	Québec, Kanada	2	(3.8)	6	(2.5)	6	(2.6)	4	(4.6)		0	(3.6)		■



Keine statistisch signifikanten Unterschiede zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Unterschiede zwischen 2019 und 2023 ($p \leq .05$)
▲ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
\+ = Mittelwert in 2023 signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

In vielen anderen Staaten waren die Geschlechterunterschiede bereits 2007 gering und blieben dann auf einem ebenfalls niedrigen Niveau (u. a. in England und Schweden).

6.6 Ergebnisse zu geschlechterspezifischen Unterschieden in Einstellungen und Selbstkonzepten

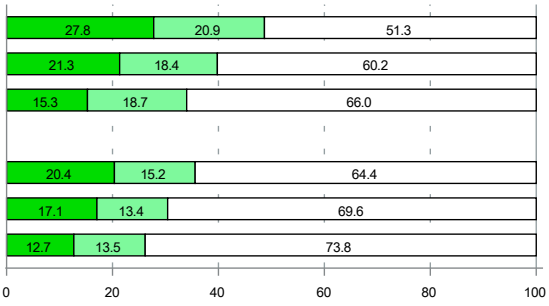
Wie bereits in den Ausführungen in Abschnitt 6.2 beschrieben, gehören die fachbezogenen Einstellungen und Selbstkonzepte von Schüler:innen zu den zentralen Prädiktoren erfolgreichen Lernens. Dabei gelten positive Einstellungen sowie positiv ausgeprägte Selbstkonzepte als zentrale Richtgrößen, die es insbesondere im Primarbereich zu fördern gilt. Im Folgenden werden geschlechterspezifische Unterschiede für diese lernförderlichen Merkmale sowohl für Mathematik als auch Naturwissenschaften für Deutschland beschrieben. Auch Veränderungen im Trend werden in den Blick genommen.

6.6.1 Geschlechterspezifische Unterschiede in Einstellungen und Selbstkonzepten in Mathematik

Die Erfassung der Einstellung zur Mathematik wird in Kapitel 3 in diesem Band ausführlich beschrieben. Entsprechend basieren die in diesem Kapitel vorgestellten Befunde auch auf den dort beschriebenen Instrumenten und Skalen. In Abbildung 6.13 werden sowohl der Mittelwert als auch die zugehörige Standardabweichung für Mädchen und Jungen getrennt berichtet. Zudem wurden für die vorliegenden Analysen die Angaben der Schüler:innen zu ihrer Einstellung in drei Gruppen unterteilt: niedrige positive Einstellung zu Mathematik (*niedrig*), mittlere positive Einstellung (*mittel*) und hohe positive Einstellung (*hoch*). Dabei werden jeweils die prozentualen Verteilungen von Mädchen und Jungen im Hinblick auf die Gruppenzugehörigkeit mit einer grafischen Darstellung berichtet (Abbildung 6.13).

Abbildung 6.13: Prozentuale Verteilung auf der Skala positive Einstellung zu Mathematik in TIMSS 2007, 2019 und 2023 – Mädchen und Jungen im Vergleich

		Mittlerer Skalenwert		Positive Einstellung zur Mathematik					
				niedrig		mittel		hoch	
		M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
Mädchen	2023	2.77 –	(0.02) ■	27.8	(1.1) ■	20.9	(0.8) ■	51.3	(1.2) ■
	2019	2.94 –	(0.03) ▲	21.3	(1.5) ▼	18.4	(1.1) ■	60.2	(1.6) ▲
	2007	3.07 –	(0.02) ▲	15.3	(0.8) ▼	18.7	(0.9) ■	66.0	(1.3) ▲
Jungen	2023	3.01	(0.02) ■	20.4	(0.9) ■	15.2	(0.7) ■	64.4	(1.1) ■
	2019	3.14	(0.03) ▲	17.1	(1.1) ▼	13.4	(1.0) ■	69.6	(1.4) ▲
	2007	3.24	(0.02) ▲	12.7	(0.9) ▼	13.5	(0.8) ■	73.8	(1.2) ▲



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen mit niedriger positiver Einstellung zur Mathematik
% der Schüler:innen mit mittlerer positiver Einstellung zur Mathematik
% der Schüler:innen mit hoher positiver Einstellung zur Mathematik
▲ = Wert im Vergleich zu 2023 statistisch signifikant höher ($p \leq .05$)
▼ = Wert im Vergleich zu 2023 statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu 2023 nicht statistisch signifikant ($p > .05$)
– = Mittlerer Skalenwert im Vergleich zu Jungen statistisch signifikant niedriger ($p \leq .05$)

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

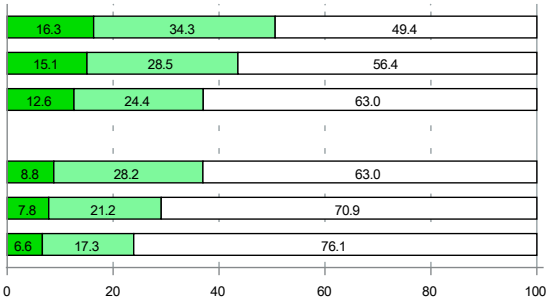
Im Vergleich von Mädchen und Jungen wird deutlich, dass der Mittelwert der Mädchen für die Einstellung zur Mathematik mit $M = 2.77$ signifikant unter dem Mittelwert der Jungen liegt ($M = 3.01$). Diese Differenz weist eine Effektstärke von $d = .25$ auf und ist damit als gering zu interpretieren. Zudem befinden sich deutlich mehr Mädchen in der Gruppe mit einer niedrigen positiven Einstellung (27.8 % versus 20.4 % der Jungen). Das bedeutet, dass mehr als jedes vierte Mädchen und etwas mehr als jeder fünfte Junge eine niedrige positive Einstellung gegenüber Mathematik haben. Demgegenüber steht ein relativ großer Anteil von 51.3 Prozent der Mädchen sowie 64.4 Prozent der Jungen, der über eine hohe positive Einstellung zu Mathematik verfügt. Obgleich dies erfreulich ist, zeigt sich doch im Trend, dass der Anteil in dieser Gruppe im Vergleich zu TIMSS 2019 und TIMSS 2007 sukzessive abgenommen hat. Diese Beobachtung ist bei Mädchen mit einer Abnahme von fast 15 Prozentpunkten im Vergleichs-

zeitraum von TIMSS 2007 zu TIMSS 2023 deutlicher ausgeprägt als bei Jungen (9 Punkte Differenz). Gleichzeitig hat der Anteil an Kindern mit einer niedrigen positiven Einstellung zu Mathematik zugenommen: bei Jungen im Vergleich zu TIMSS 2007 um fast 8 Prozentpunkte (von 12.7 % auf 20.4 %) und bei Mädchen um 12.5 Punkte (von 15.3 auf 27.8 %).

Neben der Einstellung zur Mathematik ist das fachbezogene Selbstkonzept als lernförderliche Variable ebenfalls von großem Interesse. Die Beschreibung des Instruments erfolgt in Kapitel 3 in diesem Band. Auch hier stehen hohe Werte für eine positive Ausprägung des Fähigkeitsselbstkonzepts. Aufgrund einer veränderten Operationalisierung des Konstruktes im Vergleich zu den vorherigen Zyklen (siehe ebenfalls Kapitel 3 in diesem Band) können Unterschiede im Trend nur eingeschränkt interpretiert werden. In Abbildung 6.14 werden erneut zunächst die Mittelwerte und Standardfehler für Mädchen und Jungen getrennt berichtet. Analog zur Einstellung zur Mathematik werden auch hier die Viertklässler:innen anhand ihrer Angaben in drei Gruppen unterteilt: Kinder mit einem niedrigen mathematikbezogenen Selbstkonzept (*niedrig*), einem mittleren mathematikbezogenen Selbstkonzept (*mittel*) und einem hohen mathematikbezogenen Selbstkonzept (*hoch*). Auch hier wird zur Visualisierung auf die Darstellung von gekippten Balkendiagrammen zurückgegriffen.

Abbildung 6.14: Prozentuale Verteilung auf der Skala mathematikbezogenes Selbstkonzept in TIMSS 2007, 2019 und 2023 – Mädchen und Jungen im Vergleich

		Mittlerer Skalenwert		Mathematikbezogenes Selbstkonzept niedrig		mittel		hoch	
		M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
Mädchen	2023[A]	2.86	(0.02)	16.3	(0.8)	34.3	(1.1)	49.4	(1.2)
	2019	2.95	(0.02)	15.1	(1.1)	28.5	(1.1)	56.4	(1.4)
	2007	3.05	(0.02)	12.6	(1.0)	24.4	(1.0)	63.0	(1.2)
Jungen	2023[A]	3.12	(0.02)	8.8	(0.6)	28.2	(0.9)	63.0	(1.1)
	2019	3.26	(0.02)	7.8	(0.8)	21.2	(1.2)	70.9	(1.4)
	2007	3.30	(0.02)	6.6	(0.6)	17.3	(1.0)	76.1	(1.1)



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen mit niedrigem mathematikbezogenen Selbstkonzept
% der Schüler:innen mit mittlerem mathematikbezogenen Selbstkonzept
% der Schüler:innen mit hohem mathematikbezogenen Selbstkonzept
▲ = Wert im Vergleich zu 2023 statistisch signifikant höher ($p \leq .05$)
▼ = Wert im Vergleich zu 2023 statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu 2023 nicht statistisch signifikant ($p > .05$)
– = Mittlerer Skalenwert im Vergleich zu Jungen statistisch signifikant niedriger ($p \leq .05$)
A = Aufgrund veränderter Operationalisierung sind die Ergebnisse nur eingeschränkt mit vorherigen Zyklen vergleichbar.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Der Mittelwertsunterschied im mathematikbezogenen Selbstkonzept von Mädchen ($M = 2.86$) und Jungen ($M = 3.12$) in TIMSS 2023 liegt bei rund 0.3 Skalenpunkten und ist signifikant. Mit einer Effektgröße von $d = .30$ handelt es sich, ähnlich wie für die mathematikbezogenen Einstellungen, um einen geringen Effekt. Es befinden sich mit 16.3 Prozent etwa doppelt so viele Mädchen wie Jungen (8.8 %) in der Gruppe mit einem niedrigen mathematikbezogenen Selbstkonzept. Weniger Mädchen als Jungen weisen zudem ein hohes mathematikbezogenes Selbstkonzept auf (rund 49 % im Vergleich zu 63 %). Bei Betrachtung der Veränderungen der prozentualen Anteile über die Zeit wird deutlich, dass

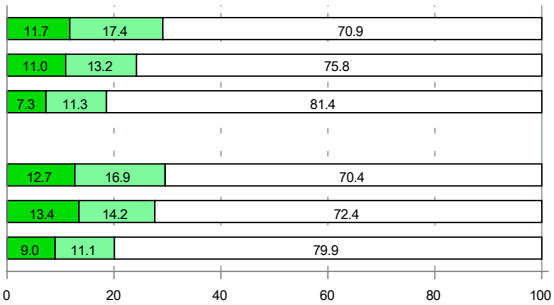
der Anteil an Viertklässler:innen mit einem hohen mathematikbezogenen Selbstkonzept über die Zyklen TIMSS 2007, 2019 bis 2023 sukzessive, aber in einem ähnlichen Ausmaß, abgenommen hat (bei den Mädchen insgesamt 13.6, bei den Jungen 13.1 Prozentpunkte). Diese Veränderung bildet sich auch bei Betrachtung der Veränderung des Mittelwertes im Trend ab. Auch hier zeigt sich sowohl bei Mädchen als auch bei Jungen eine Abnahme im durchschnittlichen mathematikbezogenen Selbstkonzept bei Mädchen und Jungen um rund 0.2 Skalenpunkte, obgleich hier aufgrund der abweichenden Operationalisierung im aktuellen Zyklus Vorsicht bei der Interpretation geboten ist. Dieser Trend findet sich bereits in den vergangenen TIMSS-Zyklen beschrieben, wo eine ähnliche Tendenz beobachtet werden konnte (Nonte et al., 2020).

6.6.2 Geschlechterspezifische Unterschiede in Einstellungen und Selbstkonzepten im Sachunterricht

Wie in den vorherigen Zyklen werden auch in TIMSS 2023 die Einstellung zum Sachunterricht und das sachunterrichtsbezogene Selbstkonzept erfasst. Abbildung 6.15 visualisiert die geschlechterspezifischen Unterschiede hinsichtlich der positiven Einstellung zum Sachunterricht. Neben direkten Vergleichen zwischen den Ausprägungen der Verteilungen und der Mittelwerte von Mädchen und Jungen für TIMSS 2023, können der Abbildung erneut auch Veränderungen im Trend (TIMSS 2007 und TIMSS 2019) entnommen werden.

Abbildung 6.15: Prozentuale Verteilung auf der Skala positive Einstellung zum Sachunterricht in TIMSS 2007, 2019 und 2023 – Mädchen und Jungen im Vergleich

		Mittlerer Skalenwert		Positive Einstellung zum Sachunterricht					
				niedrig		mittel		hoch	
		M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
Mädchen	2023	3.16	(0.02) ■	11.7	(0.8) ■	17.4	(0.8) ■	70.9	(1.1) ■
	2019	3.24	(0.03) ▲	11.0	(1.0) ■	13.2	(1.0) ▼	75.8	(1.3) ▲
	2007	3.38	(0.02) ▲	7.3	(0.7) ▼	11.3	(0.7) ▼	81.4	(1.0) ▲
Jungen	2023	3.15	(0.02) ■	12.7	(0.6) ■	16.9	(0.7) ■	70.4	(1.1) ■
	2019	3.20	(0.03) ■	13.4	(1.3) ■	14.2	(1.1) ▼	72.4	(1.8) ■
	2007	3.38	(0.02) ▲	9.0	(0.7) ▼	11.1	(0.8) ▼	79.9	(1.2) ▲



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen mit niedriger positiver Einstellung zum Sachunterricht
% der Schüler:innen mit mittlerer positiver Einstellung zum Sachunterricht
% der Schüler:innen mit hoher positiver Einstellung zum Sachunterricht
▲ = Wert im Vergleich zu 2023 statistisch signifikant höher ($p \leq .05$)
▼ = Wert im Vergleich zu 2023 statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu 2023 nicht statistisch signifikant ($p > .05$)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Zunächst ist der Anteil der Mädchen und Jungen, die über eine hohe positive Einstellung zum Sachunterricht verfügen mit rund 70 Prozent, insbesondere im direkten Vergleich mit Mathematik, sehr hoch. Die Anteile in der Gruppe mit einer niedrigen positiven Einstellung betragen 11.7 Prozent bei den Mädchen und 12.7 Prozent bei den Jungen und liegen ebenfalls nah beieinander. Im Vergleich zu den mathematikbezogenen Einstellungen fallen hier also die geschlechter-

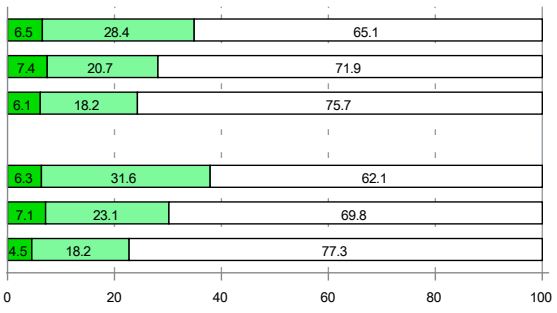
spezifischen Unterschiede eher moderat aus. Mädchen und Jungen unterscheiden sich weder im Hinblick auf den mittleren Skalenwert noch hinsichtlich ihrer Verteilung auf die Gruppen mit niedriger, mittlerer oder hoher positiver Einstellung zum Sachunterricht. Auch in den vergangenen Zyklen konnte dies bereits so beobachtet werden. Allerdings zeigt sich bei den Mädchen eine Abnahme der prozentualen Anteile in der Gruppe mit einer hohen positiven Einstellung zum Sachunterricht im Vergleich von TIMSS 2023 gegenüber TIMSS 2007 um 10.5 Prozentpunkte. Dieser Unterschied ist signifikant, ebenso wie die Zunahme in der Gruppe der Mädchen mit einer mittleren positiven Einstellung zum Sachunterricht im Vergleich zu TIMSS 2007 und TIMSS 2019. Auch bei den Jungen kann die Zunahme in der Gruppe mit einer mittleren positiven Einstellung beobachtet werden. Eine Tendenz zeichnet sich für die Abnahme der Anteile in der Gruppe mit hoher positiver Einstellung für Jungen ab. Diese ist im direkten Vergleich von TIMSS 2019 und TIMSS 2023 nicht signifikant, jedoch im Trend seit 2007.

Neben der Einstellung zum Sachunterricht ist die Einschätzung der eigenen fachbezogenen Fähigkeiten ebenfalls von Interesse. In Abbildung 6.16 finden sich der Skalenmittelwert für Mädchen und Jungen sowie die Verteilung auf die Gruppen mit einem niedrigen, mittleren und hohem Selbstkonzept.

Bei Betrachtung des Skalenmittelwertes fällt auf, dass Mädchen und Jungen sich in TIMSS 2023 bezüglich ihres sachunterrichtsbezogenen Selbstkonzepts nicht unterscheiden. Auch in TIMSS 2019 war der Geschlechterunterschied sehr gering und nicht signifikant, wobei auch für das sachunterrichtsbezogene Selbstkonzept Vergleiche im Trend nur eingeschränkt möglich sind (zur Operationalisierung des sachunterrichtsbezogenen Selbstkonzepts siehe Kapitel 4 in diesem Band). Lediglich im Jahr 2007 lag der Mittelwert der Jungen signifikant über dem der Mädchen. Innerhalb der Gruppen zeichnet sich hingegen eine negative Tendenz im Berichtszeitraum ab. So sinkt der mittlere Skalenwert der Mädchen

Abbildung 6.16: Prozentuale Verteilung auf der Skala sachunterrichtsbezogenes Selbstkonzept in TIMSS 2007, 2019 und 2023 – Mädchen und Jungen im Vergleich

		Mittlerer Skalenwert		Sachunterrichtsbezogenes Selbstkonzept					
				niedrig		mittel		hoch	
		M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
Mädchen	2023[A]	3.15	(0.02) ■	6.5	(0.4) ■	28.4	(1.0) ■	65.1	(1.1) ■
	2019	3.20	(0.02) ▲	7.4	(0.7) ■	20.7	(1.3) ▼	71.9	(1.2) ▲
	2007	3.25 –	(0.02) ▲	6.1	(0.6) ■	18.2	(0.8) ▼	75.7	(1.1) ▲
Jungen	2023[A]	3.12	(0.02) ■	6.3	(0.4) ■	31.6	(0.8) ■	62.1	(1.0) ■
	2019	3.23	(0.02) ▲	7.1	(0.7) ■	23.1	(1.3) ▼	69.8	(1.6) ▲
	2007	3.33	(0.02) ▲	4.5	(0.4) ▼	18.2	(0.9) ▼	77.3	(1.1) ▲



% der Schüler:innen mit niedrigem sachunterrichtsbezogenen Selbstkonzept
% der Schüler:innen mit mittlerem sachunterrichtsbezogenen Selbstkonzept
% der Schüler:innen mit hohem sachunterrichtsbezogenen Selbstkonzept
▲ = Wert im Vergleich zu 2023 statistisch signifikant höher ($p \leq .05$)
▼ = Wert im Vergleich zu 2023 statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu 2023 nicht statistisch signifikant ($p > .05$)
– = Mittlerer Skalenwert im Vergleich zu Jungen statistisch signifikant niedriger ($p \leq .05$)
A = Aufgrund veränderter Operationalisierung sind die Ergebnisse nur eingeschränkt mit vorherigen Zyklen vergleichbar.

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

von $M = 3.25$ im Jahr 2007 auf $M = 3.15$ im Jahr 2023. Dies kann ebenfalls für die Gruppe der Jungen beobachtet werden: von $M = 3.33$ in TIMSS 2007 auf $M = 3.12$ in TIMSS 2023. Bei der Betrachtung der Verteilung auf die Gruppen mit niedrigem, mittlerem und hohem sachunterrichtsbezogenen Selbstkonzept lassen sich erneut keine Geschlechterunterschiede für TIMSS 2023 beobachten. Es zeigt sich allerdings eine Abnahme in den Gruppen der Mädchen und Jungen mit einem hohen sachunterrichtsbezogenen Selbstkonzept im Vergleich zu TIMSS 2007 und TIMSS 2019. Befanden sich im Jahr 2007 noch jeweils mehr als 75 Prozent der Mädchen und Jungen in dieser Gruppe, sind es in TIMSS 2023 noch rund 65 Prozent der Mädchen und 62 Prozent der Jungen. Bei den Jungen beträgt die Differenz insgesamt 15 Prozentpunkte, bei den Mädchen rund 11 Punkte. Gleichzeitig kann eine Zunahme in der Gruppe der Mädchen und Jungen mit einem mittleren sachunterrichtsbezogenen Selbstkonzept betrachtet werden. Die Zunahme bei den Mädchen beträgt dabei 10.2 Prozentpunkte, bei den Jungen 13.4 Punkte.

Insgesamt zeigt sich anhand der hier berichteten Ausprägungen und Verteilungen bezogen auf das Fach Sachunterricht, dass der Großteil der Mädchen und Jungen über eine hohe positive Einstellung und ein hohes sachunterrichtsbezogenes Selbstkonzept verfügen.

6.7 Zusammenfassung

In diesem Kapitel wurde der Blick auf Geschlechterunterschiede in Mathematik und in den Naturwissenschaften gerichtet. Die Befunde liefern damit wichtige Hinweise und Kennzahlen zur Beschreibung der Beschaffenheit von geschlechterbezogenen Disparitäten und können somit als Impulsgeber für den Abbau von Ungleichheiten im deutschen Bildungssystem fungieren, da dies eine der zentralen Aufgaben des Schulsystems ist. So ist der Unterricht auf der „Wertebasis des Grundgesetzes mit Offenheit für die vielfältigen unterschiedlichen Wertvorstellungen zur Gestaltung der gesellschaftlichen Rollen von Frauen und Männern anzulegen“ (KMK, 2016, S. 2). Die Bedeutung von gesellschaftlichen Rollenverständnissen sowie den daraus resultierenden Einflüssen auf Motive, Einstellungen, Haltungen und Praxen wurden in den Abschnitten 6.1, 6.2 und 6.3 dargelegt. Basierend auf dem Grundgesetz ist die tatsächliche Durchsetzung der Gleichberechtigung von Jungen und Mädchen durch das staatliche Schulwesen zu fördern und etwaigen Benachteiligungen ist aktiv entgegenzuwirken.

Im Bereich Mathematik (Abschnitt 6.4) zeigt sich im aktuellen Zyklus erneut ein signifikanter Geschlechterunterschied von 13 Punkten auf der Gesamtskala zugunsten von Jungen. Diese Differenz liegt über dem internationalen Mittelwert, jedoch unterhalb der Vergleichswerte für die EU oder die OECD. In einigen Teilnehmerstaaten wie Frankreich, Australien, Italien und Portugal fallen die Geschlechterunterschiede zum Teil deutlich größer aus als in Deutschland, in einigen Ländern wie etwa in Lettland oder Irland fallen die Unterschiede geringer aus oder sind nicht signifikant. Zudem sind Jungen in Deutschland auf den oberen beiden Kompetenzstufen IV und V signifikant überrepräsentiert. Geschlechterunterschiede in den Kompetenzen können in allen drei Inhaltsbereichen (*Arithmetik, Messen und Geometrie, Daten*) beobachtet werden, wobei der Bereich *Arithmetik* die größte geschlechterspezifische Differenz zugunsten der Jungen aufweist. Im Vergleich mit den Befunden aus TIMSS 2019 (Nonte et al., 2020) haben die berichteten Geschlechterunterschiede für die Inhaltsbe-

reiche *Arithmetik*, *Messen und Geometrie* und *Daten* über alle Bildungssysteme hinweg zugenommen. Diese haben sich seit 2019 teilweise mehr als verdoppelt (*Arithmetik*), im Inhaltsbereich *Daten* sogar versiebenfacht. Auch für Deutschland zeigt sich im Vergleich zu TIMSS 2019 für den Bereich *Arithmetik* eine Zunahme der Geschlechterunterschiede, eine Abnahme hingegen für den Bereich *Messen und Geometrie* sowie eine leichte Zunahme im Bereich *Daten*. Auch in den Anforderungsbereichen *Reproduzieren*, *Anwenden* und *Problemlösen* sind die Geschlechterunterschiede sowohl im internationalen Durchschnitt als auch in Deutschland statistisch signifikant und fallen zugunsten der Jungen aus. Die größte Differenz kann in Deutschland in den Bereichen *Reproduzieren* und *Problemlösen* beobachtet werden. Dabei liegt Deutschland unterhalb des internationalen Mittelwertes und etwa im Bereich der Vergleichswerte der OECD und der EU. Im Vergleich mit den Befunden aus TIMSS 2019 zeigt sich für den Anwendungsbereich *Reproduzieren* eine ähnliche Ausprägung wie für den aktuellen Zyklus. Eine leichte Zunahme von Geschlechterunterschieden kann hingegen für die Anwendungsbereiche *Anwenden* und *Problemlösen* berichtet werden: Hier steigt die durchschnittliche Geschlechterdifferenz leicht an. Im Bereich *Problemlösen* ist ebenfalls eine Zunahme von Geschlechterunterschieden zu beobachten. Die Trendanalysen zur Geschlechterdifferenz auf der Gesamtskala Mathematik für die Jahre 2007, 2019 und 2023 verweisen für Mädchen und Jungen auf eine relativ stabile Entwicklung. Entsprechend haben sich die bestehenden Unterschiede zwischen den Geschlechtern im Trend nicht signifikant verändert. Dieser Befund verweist darauf, dass Geschlechterunterschiede in den mathematischen Leistungen von Viertklässler:innen in Deutschland bisher noch nicht effektiv abgebaut werden konnten.

Die Differenz in den naturwissenschaftlichen Leistungen von Mädchen und Jungen in Deutschland fällt im Vergleich zu den Mathematikleistungen eher gering aus (Abschnitt 6.4). Der Unterschied von 2 Punkten zugunsten der Mädchen ist zudem nicht statistisch signifikant. International betrachtet ist der Geschlechterunterschied in der naturwissenschaftlichen Leistung in vielen Ländern nicht signifikant, dennoch weisen einige Teilnehmerstaaten signifikante Unterschiede zugunsten der Jungen auf. Lediglich in Finnland erreichen Mädchen signifikant höhere Testleistungen. Deutschland liegt mit 2 Punkten Differenz etwa im internationalen Mittel. Die Betrachtung der Anteile von Mädchen und Jungen auf den Kompetenzstufen für Deutschland liefert ebenfalls keine signifikanten Unterschiede. Bei Betrachtung der unterschiedlichen Inhaltsbereiche können für Deutschland sowie auch international Unterschiede hinsichtlich der erreichten Testleistungen bei Mädchen und Jungen beobachtet werden. Vergleichsweise groß ist der Unterschied im Bereich *Biologie* ausgeprägt. Hier erreichen Mädchen in Deutschland signifikant höhere Testleistungen als Jungen. International zeigt sich hingegen kein einheitliches Muster. In Finnland erreichen Mädchen ebenfalls höhere Testleistungen als Jungen, in der Republik Korea (Südkorea) erreichen Jungen hingegen mehr Punkte als Mädchen. Auch für den Inhaltsbereich *Physik/Chemie* sind die Geschlechterunterschiede stark unterschiedlich ausgeprägt. Die Differenz für Deutschland beträgt lediglich 2 Punkte. Dieser Unterschied ist nicht signifikant, wie es für viele andere Teilnehmerstaaten ebenfalls zu beobachten ist. Anders sieht es für den Bereich *Geografie* aus. Hier erreichen in einer Vielzahl der Teilnehmerstaaten Jungen bessere Testleistungen als Mädchen. Obgleich Mädchen im Inhaltsbereich *Biologie* im Mittel besser abschneiden, schneiden Jungen im Inhaltsbereich *Geografie* besser ab. Im Vergleich zu TIMSS 2019 (Nonte et al., 2020) zeigt sich für den Inhalts-

bereich *Biologie* eine Zunahme der Geschlechterunterschiede von insgesamt 7 Punkten zugunsten von Mädchen. Im Inhaltsbereich *Physik/Chemie* fällt der Geschlechterunterschied deutlich geringer aus, im Inhaltsbereich *Geografie* hingegen kann ebenfalls eine leichte Zunahme von Geschlechterunterschieden verzeichnet werden. In den Anforderungsbereichen unterscheiden sich Mädchen und Jungen in Deutschland lediglich im Bereich *Problemlösen*. Hier erreichen Mädchen in Deutschland mehr Punkte als Jungen. Auch international finden sich im Bereich *Problemlösen* in den meisten Staaten Tendenzen in Richtung höherer Testleistungen von Mädchen. Im Vergleich mit den Befunden für Deutschland aus TIMSS 2019 zeigen sich für den Anforderungsbereich *Reproduzieren* keine bedeutsamen Unterschiede. Im Bereich *Anwenden* fallen die Befunde sehr ähnlich aus (ebenfalls ohne signifikanten Unterschied). Im Bereich *Problemlösen* findet sich eine deutliche Zunahme von Geschlechterunterschieden in TIMSS 2023 zugunsten von Mädchen. Die Trendanalysen hingegen zeigen keine signifikanten Unterschiede der Testleistungen von Mädchen für die Gesamtskala Naturwissenschaften im Vergleich der Jahre 2007, 2019 und 2023, jedoch eine deutliche Abnahme der Testleistungen der Jungen im Zeitraum von 2007 bis 2023. Diese Abnahme beläuft sich auf 20 Punkte im Vergleich zu TIMSS 2007. Entsprechend muss hier konstatiert werden, dass Geschlechterunterschiede hinsichtlich ihrer Ausprägung möglicherweise stabil geblieben sind, da insbesondere für Jungen eine starke Abnahme der Testleistungen im Trend zu verzeichnen ist und diese entsprechend nicht auf eine Zunahme der Testleistungen der Mädchen zurückgeführt werden kann. So findet über die Zeit eine Angleichung der Leistungen statt. Dieser Befund wurde erstmals im Berichtsband von TIMSS 2019 dokumentiert (Nonte et al., 2020).

Hinweise zu geschlechterspezifischen Unterschieden in Einstellungen und Selbstkonzepten wurden getrennt für Mathematik und den Sachunterricht berichtet (Abschnitt 6.6). Erwartungskonform weisen Mädchen eine signifikant geringere positive Einstellung zu Mathematik auf. Mehr als jedes vierte Mädchen verfügt über eine niedrige positive Einstellung zu Mathematik, bei Jungen ist es jeder fünfte. Insbesondere im Vergleich zu vorherigen Zyklen zeigt der Trend eine deutliche Abnahme von Mädchen und Jungen, die über eine hohe positive Einstellung zu Mathematik verfügen. Auch hinsichtlich der Ausprägung des mathematikbezogenen Selbstkonzepts wird deutlich, dass sich etwa doppelt so viele Mädchen wie Jungen in der Gruppe mit einem niedrigen mathematikbezogenen Selbstkonzept befinden. Analog dazu weist nur jedes zweite Mädchen ein hohes mathematikbezogenes Selbstkonzept auf. Im Trend wird deutlich, dass die Gruppe der Schüler:innen mit einem hohen mathematikbezogenen Selbstkonzept, unabhängig vom Geschlecht, um mehr als 13 Prozentpunkte seit dem Jahr 2007 abgenommen hat. Diesem negativen Trend in der Entwicklung lernförderlicher, motivationaler Haltungen und Einstellungen gilt es zwingend entgegenzuwirken. Auch wenn Mädchen hier besonders stark von einer niedrigen positiven Einstellung und einem niedrigen mathematikbezogenen Selbstkonzept betroffen zu sein scheinen, so ist die Abnahme positiver Einstellungen zu Mathematik auch bei Jungen zu beobachten. Damit scheinen Hinweise auf eine geschlechtersensible Gestaltung des Mathematikunterrichts (u. a. Budde, 2008; Jahnke-Klein, 2012) nur dann vielversprechend zu sein, wenn sie nicht mit einer zu starken Fokussierung auf das weibliche Geschlecht einhergehen. Vielmehr braucht es ein ganzheitliches und alle Geschlechter inkludierendes Konzept zur Förderung lernförderlicher Einstellungen und Selbstkonzepte im Mathematikunterricht.

Für den Sachunterricht können keine Unterschiede zwischen Mädchen und Jungen hinsichtlich ihrer Einstellung zum Sachunterricht beobachtet werden. Darüber hinaus verfügt die Mehrzahl der Mädchen und Jungen über ein hohes sachunterrichtsbezogenes Selbstkonzept. Im Trend zeigen sich hingegen einige Veränderungen. So kann sowohl bei Mädchen als auch bei Jungen eine Abnahme der prozentualen Anteile in der Gruppe mit einer hohen positiven Einstellung zum Sachunterricht seit dem Jahr 2007 beobachtet werden. Ein ähnliches Muster findet sich für das sachunterrichtsbezogene Selbstkonzept. Mädchen und Jungen unterscheiden sich auch hier nicht signifikant voneinander. Dennoch kann für beide eine signifikante Abnahme des Skalenmittelwertes seit dem Jahr 2007 beobachtet werden, bei Jungen ist dies noch deutlicher ausgeprägt als bei Mädchen. Auch im Hinblick auf die drei Gruppen mit niedrigem, mittlerem und hohem sachunterrichtsbezogenen Selbstkonzept liegen keine Unterschiede zwischen Mädchen und Jungen vor. Allerdings kann auch hier eine Abnahme in der Gruppe mit hohem sachunterrichtsbezogenen Selbstkonzept sowohl bei Mädchen als auch Jungen beobachtet werden, wobei der Rückgang bei den Jungen auch hier deutlicher ausfällt.

Erneut zeigt sich im Vergleich der Domänen Mathematik und Sachunterricht, dass die positive Einstellung und das Selbstkonzept von Mädchen und Jungen im Sachunterricht höher ausgeprägt sind als in Mathematik. Können für den Sachunterricht keine signifikanten Geschlechterunterschiede dokumentiert werden, so zeigt sich für das Fach Mathematik ein deutlicher Unterschied in den Einstellungen zu Mathematik und dem mathematikbezogenen Selbstkonzept zugunsten der Jungen. Obgleich die Mehrzahl der Viertklässler:innen über positive Einstellungen zur Mathematik und zum Sachunterricht verfügt, nimmt dieser Anteil seit 2007 kontinuierlich ab. Einschränkend sei darauf hingewiesen, dass die hier berichteten Befunde rein deskriptiver Natur sind und Wirkungs- und Ursachenanalysen auf dieser Basis nicht erfolgen können. Hierfür benötigt es eine multikriteriale Betrachtung, zumindest aber eine Analyse unter Berücksichtigung weiterer relevanter Hintergrundmerkmale, wie sie in Kapitel 9 in diesem Band vorgenommen wird. Des Weiteren wird die Geschlechtervariable nach wie vor binär kodiert, u. a. auch, um die Vergleichbarkeit zu bisherigen Berichtsbänden herzustellen. Darüber hinaus können tiefergehende Analysen unter Berücksichtigung von Geschlechterrollen(selbst)konzepten sowie Geschlechterrollenstereotypen an dieser Stelle nicht erfolgen. Hierzu bedarf es weiterer Forschung.

Tatsächlich erscheinen die berichteten Befunde besorgniserregend und verweisen auf einen deutlichen Handlungsbedarf im Hinblick auf die Stärkung von mathematikbezogenen Einstellungen und Leistungen, insbesondere, aber nicht nur, von Mädchen. In den Naturwissenschaften fallen die Befunde zu den Geschlechterdisparitäten hingegen sehr unterschiedlich aus. So erreichen Mädchen in *Biologie* höhere Testleistungen als Jungen, in *Geografie* sind es hingegen die Jungen, die signifikant besser abschneiden. Dieser Befund verdeutlicht die Notwendigkeit einer differenzierten Betrachtung und geschlechtersensiblen Förderung dieser unterschiedlichen Inhaltsbereiche. Einmal mehr kann darauf verwiesen werden, dass sich neben fachspezifischen Ansatzpunkten weitere im schulischen Alltag, aber auch strukturell im Bildungssystem finden, die eine übergreifende, systematische Gendersensibilität ermöglichen. Die *Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland* (KMK) hat hierzu Leitlinien zur Sicherung der Chancengleichheit durch geschlechtersensible schulische Bildung und Erziehung im Jahr 2016 herausgegeben (KMK, 2016), in denen unterschiedliche Handlungsbereiche angesprochen werden. Dazu

gehören neben Unterrichtsvorgaben und differenzierten Lehr- und Lernmitteln u. a. auch strukturelle Maßnahmen auf Schulebene sowie die Lehramtsaus- und -fortbildung. Die Handreichung der KMK liefert wichtige und konkrete Ansatzpunkte zur systematischen und nachhaltigen Sicherung von Chancengleichheit im Bildungswesen. Die Befunde von TIMSS 2023 deuten für Deutschland erneut darauf hin, dass eine stärkere Integration geschlechtersensibler Perspektiven im Mathematik- und Sachunterricht von Nöten ist und entsprechend verstärkt Handreichungen wie die der KMK im Rahmen der Lehrkräfteaus- und -fortbildung herangezogen werden sollten.

Literatur

Budde, J. (2008). *Mathematikunterricht und Geschlecht: Empirische Ergebnisse und pädagogische Ansätze*. https://doi.org/10.2314/GBV:601293495

Costache, O., Edelsbrunner, P. A., Becker, E. S., Sticca, F., Staub, F. C., & Götz, T. (2022). Wachstumsverläufe der intrinsischen Wertüberzeugungen in Mathematik und Französisch: Zusammenhänge mit Berufsorientierungen. *Zeitschrift für Erziehungswissenschaft*, *25*, 269–291. https://doi.org/10.1007/s11618-022-01095-y

Diedrich, J., Patzl, S., Seßler, S., & Reinhold, F. (2023). Motivational-emotionale Orientierungen Fünfzehnjähriger bezüglich Mathematik in PISA 2022: Zwischen Anstrengung und Selbstbild. In D. Lewalter, J. Diedrich, F. Goldhammer, O. Köller & K. Reiss (Hrsg.), *PISA 2022. Analyse der Bildungsergebnisse in Deutschland* (S. 87–112). Waxmann. https://doi.org/10.31244/9783830998488

Diedrich, J., Reinhold, F., Heinze, A., & Reiss, K. (2023). Mathematische Kompetenz in PISA 2022. Von Leistungsunterschieden und ihren Entwicklungen. In D. Lewalter, J. Diedrich, F. Goldhammer, O. Köller & K. Reiss (Hrsg.), *PISA 2022. Analyse der Bildungsergebnisse in Deutschland* (S. 53–85). Waxmann. https://doi.org/10.31244/9783830998488

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes. Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology, 23*, 1–74. https://doi.org/10.1016/S0065-2601(08)60317-2

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotyped domains? A meta-analysis. *Journal of School Psychology*, *53*(1), 25–44. https://doi.org/10.1016/j.jsp.2014.10.002

Gentrup, S., Schipolowski, S., & Wittig, J. (2022). Geschlechtsbezogene Disparitäten. In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (S. 127–150). Waxmann. https://doi.org/10.31244/9783830996064

Heyder, A., Weidinger, A. F., & Steinmayr, R. (2021). Only a burden for females in math? Gender and domain differences in the relation between adolescents' fixed mindsets and motivation. *Journal of Youth and Adolescence*, *50*(1), 177–188. https://doi.org/10.1007/s10964-020-01345-4

Holder, K., & Kessels, U. (2017). Gender and ethnic stereotypes in student teachers' judgments: a new look from a shifting standards perspective. *Social Psychology of Education*, *20*, 471–490. https://doi.org/10.1007/s11218-017-9384-z

IEA (International Association for the Evaluation of Educational Achievement). (o. J.). *Early IEA Studies. FIMS*. https://www.iea.nl/studies/iea/earlier#section-170

Jahnke-Klein, S. (2012). Der Wunsch, zu verstehen. Geschlechtersensibler Mathematikunterricht. *Praxis Schule 5-10*, *23*(6), 43–49.

Kastorff, T., Rönnebeck, S., Neumann, K., Seßler, S., Diedrich, J., & Schiepe-Tiska, A. (2023). Naturwissenschaftliche Kompetenz in PISA 2022. Entwicklungen und mögliche Herausforderungen. In D. Lewalter, J. Diedrich, F. Goldhammer, O. Köller & K.

Reiss (Hrsg.), *PISA 2022. Analyse der Bildungsergebnisse in Deutschland* (S. 113–137). Waxmann. https://doi.org/10.31244/9783830998488

Keeves, J. (1973). Differences between the sexes in mathematics and science courses. *International Review of Education*, *19*, 47–63. https://doi.org/10.1007/bf00597775

Keller, L., Preckel, F., Eccles, J. S., & Brunner, M. (2022). Top-performing math students in 82 countries: An integrative data analysis of gender differences in achievement, achievement profiles, and achievement motivation. *Journal of Educational Psychology*, *114*(5), 966–991. https://doi.org/10.1037/edu0000685

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2016). *Leitlinien zur Sicherung der Chancengleichheit durch geschlechtersensible schulische Bildung und Erziehung.* https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2016/2016_10_06-Geschlechtersensible-schulische_Bildung.pdf

Kuster, F. (2019). Mann – Frau: die konstitutive Differenz der Geschlechterforschung. In B. Kortendiek, B. Riegraf & K. Sabisch (Hrsg.), *Handbuch Interdisziplinäre Geschlechterforschung* (S. 3–12). Springer. https://doi.org/10.1007/978-3-658-12496-0_3

Lazarides, R., Dicke, A.-L., Rubach, C., & Eccles, J. S. (2020). Profiles of motivational beliefs in math: Exploring their development, relations to student-perceived classroom characteristics, and impact on future career aspirations and choices. *Journal of Educational Psychology*, *112*(1), 70–92. https://doi.org/10.1037/edu0000368

McCoy, S., Byrne, D., & O'Connor, P. (2022). Gender stereotyping in mothers' and teachers' perceptions of boys' and girls' mathematics performance in Ireland. *Oxford Review of Education*, *48*(3), 341–363. https://doi.org/10.1080/03054985.2021.1987208

Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research*, *90*(3), 376–419. https://doi.org/10.3102/0034654320919354

Nielsen, M. W., Stefanick, M. L., Peragine, D., Neilands, T. B., Ioannidis, J. P. A., Pilote, L., Prochaska, J. J., Cullen, M. R., Einstein, G., Klinge, I., LeBlanc, H., Paik, H. Y., & Schiebinger, L. (2021). Gender-related variables for health research. *Biology of Sex Differences*, *12*(23), 1–16. https://doi.org/10.1186/s13293-021-00366-3

Nonte, S., Steinmayr, R., & Scholz, L. A. (2020). Geschlechterunterschiede in mathematischen und naturwissenschaftlichen Kompetenzen. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 223–262). Waxmann. https://doi.org/10.31244/9783830993193

OECD (Organization for Economic Co-operation and Development). (2023). *PISA 2022 Ergebnisse. Lernstände und Bildungsgerechtigkeit.* https://doi.org/10.3278/6004956w

Olson, K. R., Key, A. C., & Eaton, N. R. (2015). Gender cognition in transgender children. *Psychological Science*, *26*(4), 467–474. https://doi.org/10.1177/0956797614568156

Pauletti, R. E., Menon, M., Cooper, P. J., Aults, C. D., & Perry, D. G. (2017). Psychological androgyny and children's mental health: A new look with new measures. *Sex Roles*, *76*, 705–718. https://doi.org/10.1007/s11199-016-0627-9

Schwippert, K., Kasper, D., Köller, O., McElvany, N., Selter, C., Steffensky, M., & Wendt, H. (Hrsg.). (2020). *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Waxmann. https://doi.org/10.31244/9783830993193

Stanat, P., Schipolowski, S., Schneider, R., Sachse, K. A., Weirich, S., & Henschel, S. (Hrsg.). (2022). *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich*. Waxmann. https://doi.org/10.31244/9783830996064

Stanat, P., Schipolowski, S., Schneider, R., Weirich, S., Henschel, S., & Sachse, K. A. (Hrsg.). (2023). *IQB-Bildungstrend 2022. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im dritten Ländervergleich*. Waxmann. https://doi.org/10.31244/9783830997771

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. https://doi.org/10.1037//0022-3514.69.5.797

van Mier, H. I., Schleepen, T. M. J., & Van den Berg, F. C. G. (2018). Gender differences regarding the impact of math anxiety on arithmetic performance in second and fourth graders. *Frontiers in Psychology*, *9*, 2690. https://doi.org/10.3389/fpsyg.2018.02690

Wendt, H., Steinmayr, R., & Kasper, D. (2016). Geschlechterunterschiede in mathematischen und naturwissenschaftlichen Kompetenzen. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 257–297). Waxmann.

Young, C. B., Wu, S. S., & Menon, V. (2012). The neurodevelopmental basis of math anxiety. *Psychological Science*, *23*(5), 492–501. https://doi.org/10.1177/0956797611429134

Kapitel 7
Soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässler:innen

Tobias C. Stubbe, Leonhard Schulz und Christin Beese

7.1 Einleitung

Seit vielen Jahren belegen sowohl nationale als auch internationale Befunde der empirischen Bildungsforschung, dass der Bildungserfolg von Schüler:innen eng mit dem sozioökonomischen Status ihrer Familien verknüpft ist (im Überblick Berkemeyer et al., 2017). Ein niedriger sozioökonomischer Status ist für den Bildungserfolg von Kindern insofern von Nachteil, da ein damit verbundenes geringeres Haushaltseinkommen häufig mit einem nicht förderlichen Lernumfeld und fehlenden Mitteln, um beispielsweise Nachhilfe zu finanzieren, einhergeht (Blossfeld, 2018). Da sich der Anteil der armutsgefährdeten Bevölkerung in Deutschland in den letzten Jahren stabil auf einem hohen Niveau gehalten hat (DESTATIS, 2024d), ist mit einer hohen sozialen Benachteiligung von Kindern aus entsprechenden Familien zu rechnen. Die sich aus dieser Ungleichheit ergebenden Konsequenzen im Hinblick auf Teilhabechancen, Bildungsverlauf sowie Kompetenzerwerb sind in der Literatur vielfach dokumentiert (z. B. Laubstein et al., 2016; Stubbe & Krieg, 2023; Tophoven et al., 2017).

Hinsichtlich der Ungleichverteilung von Bildungschancen lassen sich nach Boudon (1974) zwei Arten sogenannter Herkunftseffekte unterscheiden. Primäre Herkunftseffekte beschreiben die Ursachen für den schulischen Leistungsvorsprung von Kindern aus bildungsnahen Familien gegenüber Kindern aus bildungsfernen Familien. Sekundäre Herkunftseffekte beziehen sich auf Bildungsentscheidungen, die – auch unter Kontrolle der primären Herkunftseffekte – durch die soziale Herkunft der Schüler:innen beeinflusst werden. Mögliche Ursachen primärer Herkunftseffekte sind ein Mangel an ökonomischen Ressourcen bei bildungsfernen Familien, beispielsweise für den Kauf von Lernmaterialien oder der Finanzierung von Nachhilfe (siehe Kapitel 11 in diesem Band), sowie ein anregungsarmes soziales Umfeld, in dem kulturelle Aktivitäten (z. B. Theater- oder Museumsbesuche) selten stattfinden. Sekundäre Herkunfts-

effekte resultieren z. B. aus Schulformentscheidungen beim Übergang von der Grundschule auf eine weiterführende Schule (siehe Kapitel 12 in diesem Band).

Das vorliegende Kapitel stellt zunächst zentrale Befunde zum Zusammenhang von sozialer Herkunft und Bildungserfolg in Deutschland vor (Abschnitt 7.2). Anschließend werden Indikatoren beschrieben, die im Rahmen der *Trends in International Mathematics and Science Study* (TIMSS) 2023 zur Operationalisierung der sozialen Herkunft verwendet werden und die übergeordneten Fragestellungen dieses Kapitels vorgestellt (Abschnitt 7.3). Im Ergebnisteil werden zunächst internationale Vergleiche zu sozialen Disparitäten in Hinblick auf die mathematischen und naturwissenschaftlichen Kompetenzen präsentiert (Abschnitt 7.4). Auf nationaler Ebene werden anschließend vertiefend soziale Disparitäten beleuchtet (Abschnitt 7.5). Abschließend werden die zentralen Befunde dieses Kapitels zusammenfassend diskutiert (Abschnitt 7.6).

7.2 Soziale Herkunft und Bildungserfolg

Mit der Bildungsexpansion der vergangenen Jahrzehnte – also dem deutlich gestiegenen Anteil an höheren Bildungsabschlüssen – war die Hoffnung verbunden, dass so auch soziale Disparitäten reduziert werden könnten, da alle Schulformen Kindern aus allen sozialen Lagen grundsätzlich offenstehen (Geißler, 2014). Dass soziale Ungleichheiten dennoch weiterhin festzustellen sind, belegen empirisch fundiert vor allem die groß angelegten Schulleistungsstudien wie TIMSS, die *Internationale Grundschul-Lese-Untersuchung* (IGLU), die *International Computer and Information Literacy Study* (ICILS) und das *Programme for International Student Assessment* (PISA), an denen die Bundesrepublik Deutschland seit Ende der 1990er-Jahre regelmäßig teilnimmt (Mang et al., 2023; Senkbeil et al., 2019; Stubbe et al., 2023; Stubbe et al., 2020).

Während bei PISA im Sekundarschulbereich im Zeitraum von 2000 bis 2022 der Zusammenhang zwischen sozialer Herkunft und Bildungserfolg von Fünfzehnjährigen konstant auf hohem Niveau ist, sich über die Zeit jedoch leicht verringerte (Mang et al., 2023), zeigt sich für den Grundschulbereich über die Kohorten hinweg eine relativ stabile Koppelung zwischen sozialer Herkunft und Bildungserfolg. Weder zwischen IGLU 2001 und IGLU 2021 noch zwischen TIMSS 2007 und TIMSS 2019 konnten signifikante Veränderungen festgestellt werden (Stubbe et al., 2023; Stubbe et al., 2020).

Der regelmäßig vom *Institut zur Qualitätsentwicklung im Bildungswesen* (IQB) durchgeführte IQB-Bildungstrend (bis 2012: Ländervergleich) ermöglicht es, den Blick auf soziale Disparitäten im Vergleich der Länder der Bundesrepublik Deutschland zu richten. Die Befunde des IQB-Bildungstrends 2021 (Sachse et al., 2022) zeigen in der Primarstufe für das Fach Mathematik bundesweit einen sozialen Gradienten von 45. Dabei versteht man unter dem sozialen Gradienten den Regressionskoeffizient (b_1) einer Regression mit der jeweiligen Kompetenz als abhängige Variable und dem z-standardisierten HISEI (siehe Abschnitt 7.3) als unabhängige Variable.[1] Die entsprechenden Werte der einzelnen Länder liegen zwischen 38 in Niedersachsen und 57 in Bremen, wobei sich nur der Bremer Wert signifikant vom deutschen Wert unterscheidet. Im Vergleich zur ersten Erhebung im Jahr 2012 hat sich der soziale Gradient bundesweit signifikant um 8 Punkte vergrößert. Eine signifikante Zunahme dieses Wertes fin-

1 Abweichend davon wird im Rahmen von TIMSS und IGLU der unstandardisierte HISEI zur Berechnung des sozialen Gradienten verwendet.

det sich auf Länderebene in Berlin, Brandenburg, Bremen, Nordrhein-Westfalen, Sachsen und Schleswig-Holstein.

7.3 Indikatoren der sozialen Herkunft

Erhebungsinstrumente, die sich in der empirischen Bildungsforschung zur Erfassung der sozialen Herkunft von Schüler:innen etabliert haben, basieren weitreichend auf der theoretischen Grundlage soziologischer Arbeiten von Bourdieu (1983). Dieser unterscheidet zwischen ökonomischem, kulturellem und sozialem Kapital und beschreibt die damit verbundenen Ressourcen, auf die Familien zurückgreifen können, um in bestimmten Kontexten (z. B. im schulischen Kontext) erfolgreich handeln zu können (Stubbe & Krieg, 2023; ausführlich: Bos et al., 2010).

Eine direkte Möglichkeit zur Erfassung des kulturellen Kapitals stellt die Frage nach dem höchsten Bildungsniveau der Erziehungsberechtigten dar. In TIMSS wird dieses Merkmal mithilfe der internationalen Bildungsskala *International Standard Classification of Education* (ISCED) erhoben, die von der UNESCO entwickelt wurde (Schroedter et al., 2006; UNESCO, 2003).

Bei der *Anzahl der Bücher zu Hause* handelt es sich um einen weiteren Indikator, der sowohl das objektivierte Kulturkapital als auch das ökonomische Kapital erfasst. Indirekt misst die Variable zudem das inkorporierte Kulturkapital, da ein Zusammenhang zwischen der Anzahl der Bücher im Haushalt und dem Bildungsniveau sowie der Bildungsnähe einer Familie besteht. Die Frage nach den Büchern im Haushalt weist sowohl für die Befragung von Schüler:innen als auch Erziehungsberechtigten gute Messeigenschaften auf, ist einfach und ökonomisch in der Durchführung. Auch im Zeitalter zunehmender Digitalisierung hat sich die Anzahl der Bücher im Haushalt als besonders aussagekräftiger Indikator bewährt (Schwippert, 2019).

Der Berufsstatus von Personen liefert neben Informationen über das kulturelle Kapital zusätzlich Erkenntnisse hinsichtlich des ökonomischen Kapitals, da einerseits gewisse Bildungsniveaus notwendig für die Ausübung bestimmter Berufe sind und andererseits die Einkommensverhältnisse durch die jeweiligen Berufe meist klar definiert sind. Ausgehend von der offenen Frage nach dem Beruf werden zwei Indizes des Berufsstatus der Erziehungsberechtigten gebildet:

Der *International Socio-Economic Index of Occupational Status* (ISEI) bewertet Berufe hinsichtlich ihres sozioökonomischen Status auf einer Skala von 10 bis 90 Punkten. Während Berufe mit einem geringen sozioökonomischen Status durch niedrige Werte gekennzeichnet sind, weisen Berufe mit einem hohen sozioökonomischen Status entsprechend hohe Werte auf (Ganzeboom et al., 1992; Ganzeboom & Treiman, 1996, 2010). In der empirischen Bildungsforschung wird häufig der höchste ISEI im Haushalt (HISEI) genutzt, um den sozioökonomischen Status der Familie abzubilden.

Die zweite Abbildung des Berufsstatus ist die *Erikson-Goldthorpe-Portocarero-Klasseneinteilung* (EGP-Klassen; Erikson et al., 1979). Hierbei werden Berufe in distinkte Gruppen eingeteilt, die sich nach Art der Tätigkeit, der Weisungsbefugnis, der Art der Beschäftigung (selbstständig oder angestellt) sowie der erforderlichen Qualifikation unterscheiden. Die Gruppen zeichnen sich hinsichtlich dieser Merkmale durch eine hohe Homogenität innerhalb einer EGP-

Klasse und einer hohen Heterogenität zwischen den EGP-Klassen aus.[2] Diese nominalskalierte – und somit nicht hierarchische – Einteilung wird häufig in s*ervice class* (EGP-Klassen I und II), *intermediate class* (EGP-Klassen III und IV) und *working class* (EGP-Klassen V, VI und VII) zusammengefasst.

Zur direkten Erfassung des ökonomischen Kapitals, den monetären und materiellen Ressourcen von Familien, wird das Haushaltseinkommen verwendet. Im Zusammenschluss mit der Haushaltsgröße lässt sich zusätzlich ermitteln, ob eine Familie armutsgefährdet ist. Als armutsgefährdet gelten nach Definition der EU (Bardone & Guio, 2005) Haushalte, die zum jeweiligen Zeitpunkt über weniger als 60 Prozent des medianen Nettoäquivalenzeinkommens in einem Staat verfügen (vgl. zur Berechnung: Stubbe et al., 2020). Das mediane Nettoäquivalenzeinkommen im Jahr 2023 betrug in Deutschland 26 274 Euro (DESTATIS, 2024c). Daraus ergibt sich eine Armutsgefährdungsgrenze (60 % des Medians) von 15 765 Euro pro Jahr bzw. 1 314 Euro pro Monat. Die Armutsgefährdungsquote lag in Deutschland 2023 bei 14.4 Prozent (DESTATIS, 2024b). Dem sechsten Armuts- und Reichtumsbericht des *Bundesministeriums für Arbeit und Soziales* (BMAS) zufolge, liegt die Armutsgefährdungsquote von Kindern in Deutschland derzeit bei 69 Prozent, wenn kein erwachsenes Haushaltsmitglied erwerbstätig ist. Sind hingegen beide Erziehungsberechtigte erwerbstätig, mindestens eine oder einer in Vollzeit, sinkt die Armutsgefährdungsquote auf 4 Prozent (BMAS, 2021). Empirisch besteht folglich ein enger Zusammenhang zwischen Armut und weiteren Indikatoren der sozialen Herkunft. Da viele Familien trotz (relativ) hohem Kultur- und Sozialkapital von Armut bedroht sein können, liefert die Berücksichtigung der Armutsgefährdung zusätzliche Informationen zur sozialen Lage von Familien (z. B. Bosch & Kalina, 2007; Butterwegge, 2000; DESTATIS, 2024a).

Im Folgenden werden deshalb die übergeordneten Forschungsfragen untersucht:

1. Wie sind soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässler:innen in Deutschland im internationalen Vergleich ausgeprägt?
2. Wie haben sich soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässler:innen in Deutschland im Trend seit 2007 entwickelt?

2 Erläuterung: EGP I – Obere Dienstklasse: höhere Beamt:innen, Selbstständige mit mehr als zehn Mitarbeitenden, Hochschul- und Gymnasiallehrkräfte; EGP II – Untere Dienstklasse: mittleres Management, Beamt:innen im mittleren und gehobenen Dienst; EGP III – Routinedienstleistungen: Büro- und Verwaltungsberufe, Verkaufs- und Servicetätigkeiten; EGP IV – Selbstständige: Selbstständige ohne Mitarbeitende, Freiberufler:innen in nicht hoch qualifizierten Berufen; EGP V, VI – (Fach-)Arbeiter:innen: Vorarbeiter:innen, Meister:innen, Techniker:innen, Aufsichtskräfte; EGP VII – Un- und angelernte Arbeiter:innen: un- und angelernte Berufe mit manuellem Charakter oder geringem Anforderungsniveau (Baumert & Schümer, 2001).

7.4 Soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen im internationalen Vergleich

Im Folgenden werden zunächst Befunde zu sozialen Disparitäten in der 4. Jahrgangsstufe im internationalen Vergleich präsentiert. Als Indikatoren werden dabei zum einen die Anzahl der Bücher im Haushalt und zum anderen das Bildungsniveau der Erziehungsberechtigten (ISCED) genutzt (siehe Abschnitt 7.3). Dazu werden in den Abbildungen 7.1 und 7.2 Familien, in denen mindestens ein Elternteil einen (Fach-)Hochschulabschluss (d. h. tertiären Bildungsabschluss) erreicht hat (mindestens ISCED-Level 6), mit Familien verglichen, in denen kein Elternteil einen tertiären Bildungsabschluss besitzt (maximal ISCED-Level 5). Für die einzelnen Teilnehmerstaaten, für den Durchschnitt aller EU- bzw. OECD-Staaten sowie für den internationalen Mittelwert werden zum einen die Anteile der beiden Gruppen in Prozent und zum anderen die erreichte Mathematikkompetenz (Abbildung 7.1) bzw. die erreichte Naturwissenschaftskompetenz (Abbildung 7.2) dargestellt. Die Differenz in den Kompetenzen zwischen den beiden Gruppen wird zudem grafisch veranschaulicht.

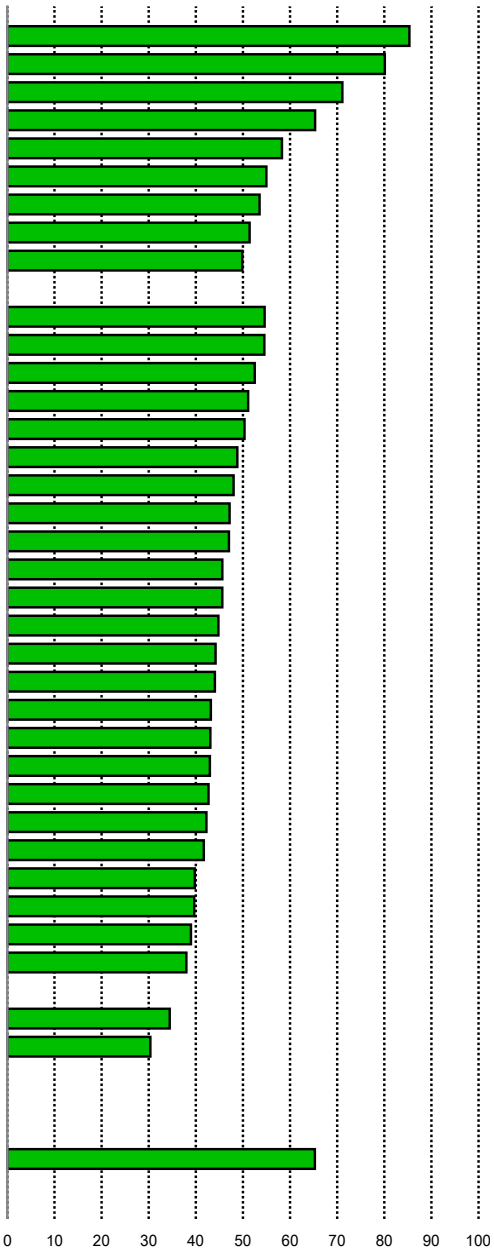
In Deutschland gehören 40.6 Prozent der Familien in die Gruppe mit tertiärem Bildungsabschluss. Dieser Wert liegt im Durchschnitt aller OECD-Staaten bzw. im internationalen Mittelwert etwas höher (jeweils 47.1 %). Noch deutlicher ist der Unterschied zu Deutschland bei der Betrachtung des Durchschnitts aller EU-Staaten. Dort gilt für 52.7 Prozent der Familien, dass mindestens ein Elternteil einen tertiären Abschluss hat. Abgesehen von dem Benchmark-Teilnehmer finden sich die höchsten Werte in Norwegen (71.2 %), Dänemark (69.6 %) und Singapur (67.4 %) aber auch in Schweden und den anderen ostasiatischen Staaten sowie in Kanada und einigen weiteren europäischen Staaten. Geringer als in Deutschland ist der Anteil der Familien mit mindestens einem tertiären Bildungsabschluss nur in der Türkei (26.3 %) sowie in Chile (33.5 %), Italien (34.3 %) und Serbien (34.6 %).

Im Bereich Mathematik (Abbildung 7.1) erreichen Schüler:innen in Deutschland einen Kompetenzwert von 564 Punkten, wenn mindestens ein Elternteil einen tertiären Bildungsabschluss besitzt. Ist dies nicht der Fall, liegt der Wert bei 520 Punkten, was einer Differenz von 44 Punkten entspricht. Diese Differenz ist, wie in allen anderen Teilnehmerstaaten, signifikant. Im internationalen Vergleich fällt diese Differenz der mathematischen Kompetenzwerte jedoch relativ gering aus. Zum einen weisen nur Dänemark und Macau signifikant geringere Werte auf und zum anderen zeigen sich im Durchschnitt der EU-Staaten sowie im Durchschnitt aller Teilnehmerstaaten signifikant größere Disparitäten.

Ein ähnliches Bild ergibt sich für den Kompetenzbereich Naturwissenschaften (Abbildung 7.2). Die Leistungsdifferenz in Deutschland beträgt 48 Punkte (maximal ISCED-Level 5: 513 Punkte; mindestens ISCED-Level 6: 561 Punkte). In allen Teilnehmerstaaten finden sich signifikante Unterschiede zwischen den beiden Gruppen. Nur in Macau sind die Disparitäten signifikant geringer als in Deutschland. Signifikant größer fallen die Unterschiede in Bulgarien, der Slowakei, der Türkei, Ungarn, Singapur und Zypern aus. Zu den beiden Vergleichsgruppen sowie zum internationalen Mittelwert gibt es einen signifikanten Unterschied.

Abbildung 7.1: Leistungsvorsprung in Mathematik von Kindern aus Familien, in denen mindestens ein Elternteil einen tertiären Bildungsabschluss erreicht hat (ISCED-Level 6, 7 und 8), vor Kindern aus Familien, in denen dies nicht der Fall ist

	Kein Elternteil mit (Fach-)Hochschulabschluss			Min. ein Elternteil mit (Fach-)Hochschulabschluss			Differenz[B]		
Teilnehmer[A]	%	M_1	*(SE)*	%	M_2	*(SE)*	$M_2 - M_1$	*(SE)*	
123 *Türkei (5. Jgst.)*	73.7	534	(4.3)	26.3	620	(4.3)	85	(5.4)	▲
Bulgarien	53.4	494	(5.6)	46.6	574	(2.7)	80	(6.3)	▲
2 *Ungarn*	52.2	501	(5.6)	47.8	572	(2.9)	71	(6.0)	▲
2 *Singapur*	32.6	584	(2.8)	67.4	649	(1.7)	65	(2.4)	▲
Slowakei	55.3	491	(4.3)	44.7	549	(2.3)	58	(4.4)	▲
2 *Frankreich*	57.6	466	(2.8)	42.4	521	(3.0)	55	(3.3)	▲
2 *Litauen*	35.4	533	(3.5)	64.6	587	(2.6)	54	(4.1)	▲
Internationaler Mittelwert	**52.9**	**483**	**(0.5)**	**47.1**	**534**	**(0.5)**	**51**	**(0.6)**	▲
VG EU	**47.3**	**503**	**(0.7)**	**52.7**	**553**	**(0.6)**	**50**	**(0.9)**	▲
2 *Serbien*	65.4	505	(4.3)	34.6	560	(3.1)	55	(4.9)	■
23 *Rumänien*	56.7	533	(5.1)	43.3	588	(4.3)	55	(6.8)	■
Portugal	54.1	495	(3.2)	45.9	548	(2.7)	52	(3.6)	■
2 *Zypern*	34.8	487	(3.3)	65.2	538	(2.0)	51	(3.3)	■
2 *Belgien (Frz. Gem.)*	43.0	467	(2.8)	57.0	517	(2.3)	50	(3.2)	■
VG OECD	**52.9**	**507**	**(0.7)**	**47.1**	**555**	**(0.6)**	**49**	**(0.8)**	■
23 *Belgien (Fläm. Gem.)*	37.7	497	(2.7)	62.3	545	(2.3)	48	(2.8)	■
2 *Polen*	37.7	519	(2.7)	62.3	566	(2.1)	47	(3.4)	■
Republik Korea (Südkorea)	39.0	567	(2.9)	61.0	614	(2.5)	47	(3.3)	■
2 *Tschechien*	55.6	518	(2.1)	44.4	563	(2.6)	46	(2.6)	■
2 *Spanien*	54.4	482	(2.1)	45.6	528	(1.9)	46	(2.3)	■
2 *Schweden*	39.6	511	(3.2)	60.4	556	(2.5)	45	(3.7)	■
23 *Kanada*	35.9	496	(2.8)	64.1	540	(2.3)	44	(3.3)	■
Deutschland	59.4	520	(1.9)	40.6	564	(2.1)	44	(2.5)	■
2 *Slowenien*	42.0	493	(2.2)	58.0	536	(2.3)	43	(2.9)	■
Lettland	45.4	517	(3.3)	54.6	560	(3.6)	43	(4.1)	■
Japan	49.9	571	(2.3)	50.1	614	(2.7)	43	(3.0)	■
Finnland	35.5	507	(2.3)	64.5	550	(2.2)	43	(2.7)	■
Irland	46.8	527	(3.2)	53.2	569	(3.0)	42	(3.7)	■
23 *Chile*	66.5	433	(3.4)	33.5	475	(3.4)	42	(4.0)	■
2 *Italien*	65.7	503	(3.4)	34.3	543	(2.6)	40	(3.4)	■
3 *Hongkong*	56.0	578	(4.4)	44.0	618	(3.9)	40	(4.9)	■
12 *Norwegen (5. Jgst.)*	28.8	516	(4.8)	71.2	554	(2.3)	39	(4.8)	■
Taiwan	38.4	584	(1.7)	61.6	622	(1.5)	38	(1.9)	■
23 *Dänemark*	30.4	503	(2.7)	69.6	537	(2.2)	34	(3.2)	▼
Macau	52.3	569	(1.1)	47.7	599	(1.2)	30	(1.6)	▼
Benchmark-Teilnehmer									
Dubai, VAE	18.3	513	(4.7)	81.7	579	(1.9)	65	(4.9)	▲

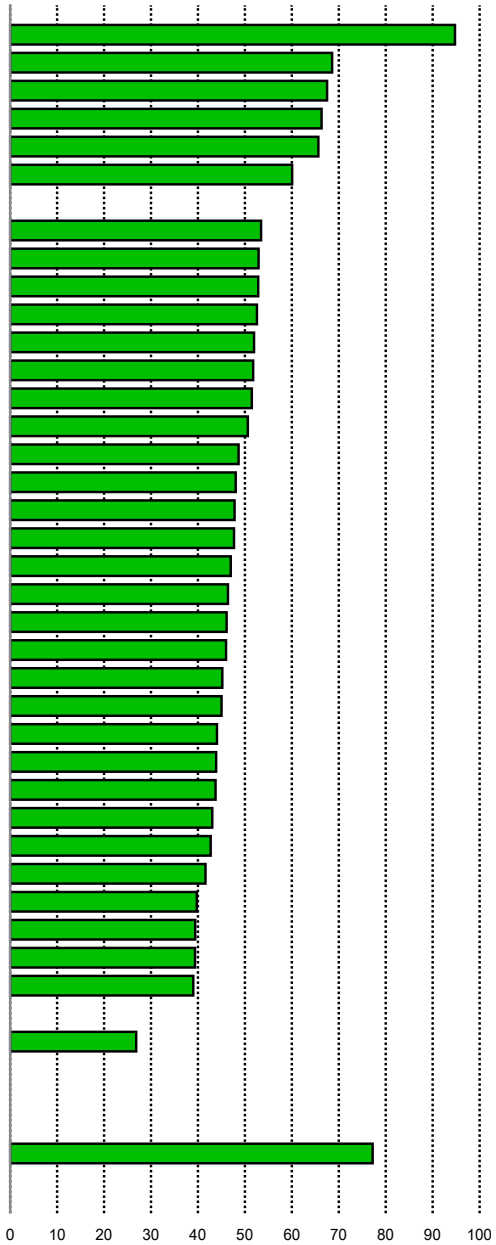


□ Keine statistisch signifikanten Unterschiede ($p > .05$)
■ Statistisch signifikante Unterschiede ($p \leq .05$)
▲ = Differenzwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Da für Australien, England, die Niederlande und die USA keine Ergebnisse vorliegen und die Rücklaufquoten von Neuseeland und Ontario, Kanada unter 40 Prozent liegen, sind diese Teilnehmer nicht Teil der Darstellung und der berichteten Mittelwerte (International, VG EU, VG OECD).
B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Abbildung 7.2: Leistungsvorsprung in Naturwissenschaften von Kindern aus Familien, in denen mindestens ein Elternteil einen tertiären Bildungsabschluss erreicht hat (ISCED-Level 6, 7 und 8), vor Kindern aus Familien, in denen dies nicht der Fall ist

	Kein Elternteil mit (Fach-)Hochschulabschluss			Min. ein Elternteil mit (Fach-)Hochschulabschluss			Differenz[B]		
Teilnehmer[A]	%	M_1	*(SE)*	%	M_2	*(SE)*	$M_2 - M_1$	*(SE)*	
Bulgarien	53.4	488	(7.6)	46.6	582	(2.9)	95	(8.1)	▲
Slowakei	55.3	492	(4.7)	44.7	561	(2.4)	69	(4.9)	▲
123 *Türkei (5. Jgst.)*	73.7	555	(3.9)	26.3	623	(3.6)	68	(4.9)	▲
2 *Ungarn*	52.2	506	(5.6)	47.8	572	(3.2)	66	(5.8)	▲
2 *Singapur*	32.6	575	(2.9)	67.4	641	(1.6)	66	(2.5)	▲
2 *Zypern*	34.8	452	(4.0)	65.2	512	(2.7)	60	(3.5)	▲
2 *Belgien (Frz. Gem.)*	43.0	456	(3.7)	57.0	510	(2.0)	53	(3.6)	■
2 *Schweden*	39.6	510	(3.3)	60.4	563	(2.9)	53	(3.5)	■
Internationaler Mittelwert	**52.9**	**472**	**(0.5)**	**47.1**	**524**	**(0.5)**	**53**	**(0.6)**	■
2 *Frankreich*	57.6	471	(3.0)	42.4	524	(2.6)	53	(3.1)	■
23 *Belgien (Fläm. Gem.)*	37.7	462	(3.6)	62.3	514	(2.5)	52	(3.7)	■
VG EU	**47.3**	**497**	**(0.8)**	**52.7**	**549**	**(0.6)**	**52**	**(0.9)**	■
2 *Slowenien*	42.0	500	(3.1)	58.0	551	(2.4)	51	(3.0)	■
2 *Litauen*	35.4	511	(4.0)	64.6	561	(3.0)	51	(4.6)	■
VG OECD	**52.9**	**505**	**(0.7)**	**47.1**	**553**	**(0.6)**	**49**	**(0.8)**	■
2 *Serbien*	65.4	494	(4.0)	34.6	542	(3.0)	48	(4.5)	■
2 *Polen*	37.7	523	(2.9)	62.3	571	(2.1)	48	(3.1)	■
Deutschland	59.4	513	(3.1)	40.6	561	(2.7)	48	(3.4)	■
12 *Norwegen (5. Jgst.)*	28.8	510	(5.2)	71.2	557	(2.6)	47	(4.7)	■
Portugal	54.1	491	(2.5)	45.9	537	(2.2)	46	(2.7)	■
23 *Rumänien*	56.7	522	(4.7)	43.3	568	(4.2)	46	(6.2)	■
3 *Hongkong*	56.0	524	(4.6)	44.0	570	(3.7)	46	(4.8)	■
Finnland	35.5	519	(3.0)	64.5	564	(2.3)	45	(2.7)	■
Taiwan	38.4	546	(1.9)	61.6	591	(1.3)	45	(2.0)	■
Republik Korea (Südkorea)	39.0	558	(2.6)	61.0	602	(2.4)	44	(2.9)	■
Lettland	45.4	509	(4.0)	54.6	553	(4.3)	44	(4.7)	■
2 *Tschechien*	55.6	514	(2.2)	44.4	557	(3.0)	44	(2.9)	■
23 *Chile*	66.5	469	(2.6)	33.5	512	(3.7)	43	(4.0)	■
2 *Spanien*	54.4	489	(1.9)	45.6	532	(2.3)	43	(2.5)	■
2 *Italien*	65.7	500	(3.1)	34.3	542	(2.3)	42	(3.3)	■
23 *Dänemark*	30.4	499	(3.0)	69.6	539	(2.7)	40	(3.6)	■
Irland	46.8	514	(3.3)	53.2	554	(3.5)	39	(3.6)	■
Japan	49.9	537	(2.0)	50.1	576	(3.4)	39	(3.6)	■
23 *Kanada*	35.9	512	(2.7)	64.1	551	(2.0)	39	(2.9)	■
Macau	52.3	524	(1.2)	47.7	551	(2.0)	27	(2.0)	▼
Benchmark-Teilnehmer									
Dubai, VAE	18.3	512	(4.8)	81.7	590	(2.2)	77	(5.1)	▲



Keine statistisch signifikanten Unterschiede ($p > .05$)
Statistisch signifikante Unterschiede ($p \leq .05$)
▲ = Differenzwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Da für Australien, England, die Niederlande und die USA keine Ergebnisse vorliegen und die Rücklaufquoten von Neuseeland und Ontario, Kanada unter 40 Prozent liegen, sind diese Teilnehmer nicht Teil der Darstellung und der berichteten Mittelwerte (International, VG EU, VG OECD).
B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

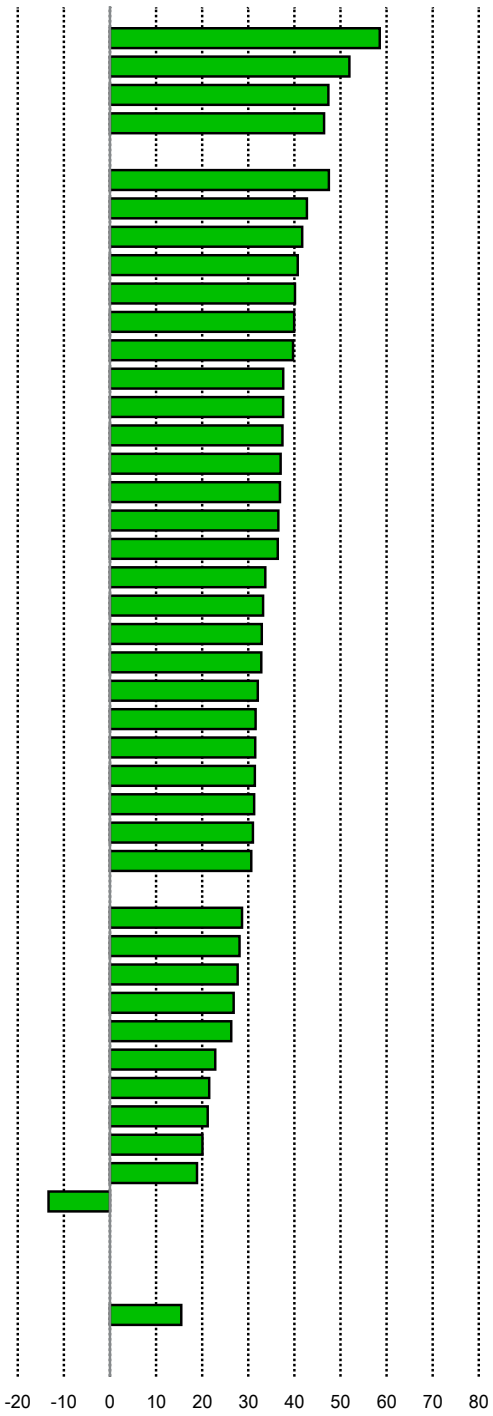
Die Abbildungen 7.3 und 7.4 sind ähnlich aufgebaut wie die beiden vorangegangenen. Indikator für den sozialen Status der Familien ist die Anzahl der im Haushalt vorhandenen Bücher, wobei die ursprünglich fünfstufige Variable dichotomisiert wurde. In Deutschland besitzen etwas mehr als zwei Drittel (71.1 %) der Familien maximal 100 Bücher und entsprechend etwas weniger als ein Drittel (28.9 %) mehr als 100 Bücher. Im internationalen Vergleich ist der Anteil von Familien mit mehr als 100 Büchern in Deutschland durchschnittlich. Im Mittel der EU-Staaten gehören 29.0 Prozent zur Gruppe mit mehr als 100 Büchern im Haushalt und im Mittel der OECD-Staaten 26.1 Prozent. Die höchsten Anteile von Familien mit vielen Büchern finden sich in Irland (40.6 %), Singapur (39.4 %), Hongkong (38.9 %), Australien (36.4 %), Frankreich (35.5 %), Finnland (34.8 %) und Neuseeland (34.7 %). Ein Ausreißer nach oben ist wie schon in den vergangenen Zyklen die Republik Korea (Südkorea), wo in 62.5 Prozent der Haushalte mehr als 100 Bücher vorhanden sind. Besonders gering ist der Anteil von Familien mit vielen Büchern innerhalb der EU in Litauen (19.6 %), Lettland (19.8 %), Rumänien (20.5 %), Bulgarien (20.7 %), Italien (23.5 %) und Portugal (24.5 %).

In Deutschland erreichen Viertklässler:innen durchschnittlich eine Mathematikkompetenz (Abbildung 7.3) von 552 Punkten, wenn es bei ihnen zu Hause mehr als 100 Bücher gibt. Ist dies nicht der Fall, beträgt der Wert nur 515 Punkte und die Differenz zwischen beiden Gruppen damit 37 Punkte. In allen teilnehmenden Staaten fallen diese Differenzen signifikant aus, wobei der Wert in Chile negativ ist, was bedeutet, dass dort der Besitz von vielen Büchern mit geringeren Kompetenzwerten einhergeht. Der Leistungsunterschied in Deutschland weicht nicht signifikant vom EU- (34 Punkte) und vom OECD-Mittelwert (33 Punkte) ab. Die Leistungsdifferenz über alle Staaten hinweg liegt mit 19 Punkten jedoch signifikant unter dem Wert für Deutschland. In vier Staaten finden sich signifikant größere Differenzen als in Deutschland: in Bulgarien, der Türkei, Ungarn und der Republik Korea (Südkorea). Die Mehrzahl der Staaten unterscheidet sich nicht signifikant von Deutschland und in zehn Staaten fallen die Disparitäten signifikant geringer aus, darunter in den EU-Staaten Italien, Lettland, Niederlande, Slowenien und Litauen.

Im Bereich Naturwissenschaften (Abbildung 7.4) fallen die sozialen Disparitäten in Deutschland etwas größer aus. Die Viertklässler:innen aus Haushalten mit vielen Büchern erreichen eine Kompetenz von 552 Punkten. Sind maximal 100 Bücher vorhanden, fällt dieser Wert 48 Punkte niedriger aus (504 Punkte). Diese Differenz liegt signifikant oberhalb der Mittelwerte der EU (40 Punkte), der OECD (37 Punkte) sowie aller Staaten (21 Punkte). Nur in Bulgarien fällt der Leistungsunterschied signifikant größer aus als in Deutschland, während sich in mehreren EU-Staaten signifikant geringere Disparitäten zeigen. Dies betrifft Italien, Lettland, Portugal, Slowenien, Litauen, die Niederlande, Spanien, Rumänien, die Französische Gemeinschaft in Belgien, Polen, Finnland und Tschechien. In allen Teilnehmerstaaten sind die Differenzen zwischen den beiden Gruppen signifikant, wobei der Wert in Chile wie im Bereich Mathematik negativ ausfällt.

Abbildung 7.3: Leistungsvorsprung in Mathematik von Kindern aus Familien mit mehr als 100 Büchern vor denen mit maximal 100 Büchern

Teilnehmer	maximal 100 Bücher %	$M_{\leq 100}$	(SE)	mehr als 100 Bücher %	$M_{>100}$	(SE)	Differenz[A] $M_{>100} - M_{\leq 100}$	(SE)	
Bulgarien	79.3	519	(4.1)	20.7	578	(4.1)	59	(5.6)	▲
123*Türkei (5. Jgst.)*	70.7	538	(4.0)	29.3	590	(4.7)	52	(4.2)	▲
2 *Ungarn*	66.6	506	(3.8)	33.4	553	(3.6)	47	(4.0)	▲
Republik Korea (Südkorea)	37.5	565	(3.6)	62.5	612	(2.5)	46	(3.6)	▲
123*England*	67.6	538	(3.1)	32.4	585	(4.1)	47	(4.7)	■
Irland	59.4	529	(3.3)	40.6	571	(3.0)	43	(3.5)	■
3*Hongkong*	61.1	579	(4.0)	38.9	621	(5.1)	42	(4.7)	■
2 *Schweden*	67.5	517	(3.2)	32.5	558	(2.5)	41	(3.7)	■
Australien	63.6	512	(2.0)	36.4	552	(2.4)	40	(2.5)	■
2 *Singapur*	60.6	599	(2.5)	39.4	639	(2.2)	40	(2.3)	■
123*Neuseeland*	65.3	477	(2.9)	34.7	517	(3.0)	40	(3.2)	■
2 *Frankreich*	64.5	471	(3.1)	35.5	509	(3.3)	38	(3.7)	■
Japan	72.6	581	(2.2)	27.4	619	(3.2)	38	(2.7)	■
23*USA*	71.5	507	(2.5)	28.5	544	(3.7)	37	(3.0)	■
Deutschland	71.1	515	(1.8)	28.9	552	(2.9)	37	(2.7)	■
Slowakei	70.3	505	(3.6)	29.7	542	(2.8)	37	(3.6)	■
2 *Zypern*	67.3	506	(2.4)	32.7	542	(2.5)	37	(2.5)	■
2 *Polen*	68.9	535	(2.0)	31.1	572	(2.8)	36	(3.0)	■
VG EU	**71.0**	**515**	**(0.6)**	**29.0**	**548**	**(0.7)**	**34**	**(0.8)**	■
VG OECD	**73.9**	**515**	**(0.5)**	**26.1**	**548**	**(0.6)**	**33**	**(0.6)**	■
23*Belgien (Fläm. Gem.)*	74.0	514	(2.3)	26.0	547	(3.0)	33	(3.3)	■
2 *Tschechien*	66.6	520	(2.1)	33.4	553	(2.9)	33	(2.9)	■
23*Dänemark*	67.2	514	(2.3)	32.8	546	(2.2)	32	(2.7)	■
2 *Spanien*	71.7	490	(2.0)	28.3	522	(2.1)	32	(2.0)	■
Portugal	75.5	510	(3.0)	24.5	541	(3.4)	32	(3.4)	■
12 *Norwegen (5. Jgst.)*	68.7	522	(2.1)	31.3	553	(3.0)	31	(3.0)	■
23*Rumänien*	79.5	537	(4.9)	20.5	568	(7.2)	31	(7.0)	■
Finnland	65.2	519	(2.4)	34.8	550	(2.8)	31	(2.6)	■
2 *Belgien (Frz. Gem.)*	66.8	480	(2.5)	33.2	510	(2.8)	31	(3.3)	■
Taiwan	70.2	599	(1.4)	29.8	628	(2.2)	29	(2.2)	▼
2 *Litauen*	80.4	555	(2.4)	19.6	584	(3.9)	28	(3.6)	▼
23*Kanada*	67.9	497	(1.6)	32.1	524	(1.9)	28	(1.9)	▼
Macau	68.4	574	(1.0)	31.6	601	(1.8)	27	(2.0)	▼
2 *Serbien*	76.7	518	(3.4)	23.3	544	(4.7)	26	(4.7)	▼
2 *Slowenien*	70.0	508	(2.0)	30.0	531	(2.9)	23	(3.2)	▼
3*Niederlande*	73.8	532	(2.1)	26.2	553	(3.2)	22	(3.5)	▼
Lettland	80.2	531	(2.5)	19.8	552	(4.6)	21	(4.2)	▼
2 *Italien*	76.5	510	(3.1)	23.5	530	(3.2)	20	(3.5)	▼
Internationaler Mittelwert	**73.9**	**498**	**(0.4)**	**26.1**	**517**	**(0.6)**	**19**	**(0.6)**	▼
23*Chile*	85.2	447	(2.5)	14.8	434	(5.6)	-13	(4.9)	▼
Benchmark-Teilnehmer									
Dubai, VAE	71.2	553	(1.6)	28.8	569	(2.2)	15	(2.4)	▼

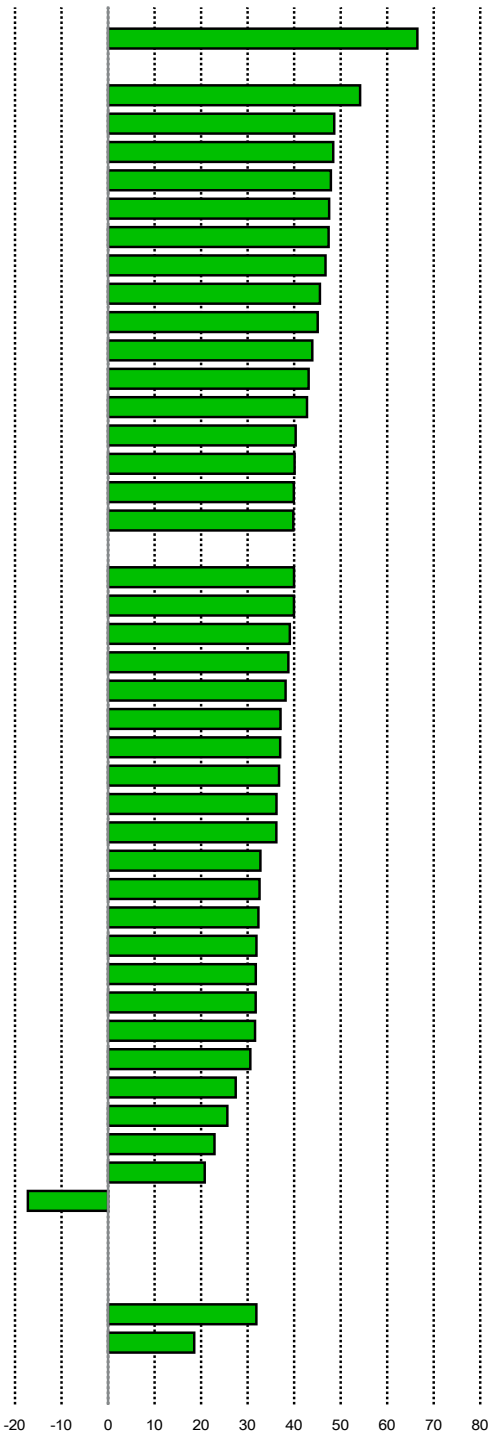


Keine statistisch signifikanten Unterschiede ($p > .05$)
Statistisch signifikante Unterschiede ($p \leq .05$)
▲ = Differenzwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Abbildung 7.4: Leistungsvorsprung in Naturwissenschaften von Kindern aus Familien mit mehr als 100 Büchern vor denen mit maximal 100 Büchern

	maximal 100 Bücher			mehr als 100 Bücher			Differenz[A]		
Teilnehmer	%	$M_{\leq 100}$	*(SE)*	%	$M_{>100}$	*(SE)*	$M_{>100} - M_{\leq 100}$	*(SE)*	
Bulgarien	79.3	517	(5.6)	20.7	584	(4.2)	66	(6.9)	▲
2 *Schweden*	67.5	516	(3.6)	32.5	570	(3.2)	54	(4.6)	■
Slowakei	70.3	507	(3.7)	29.7	556	(3.7)	49	(4.1)	■
2 *Zypern*	67.3	473	(2.8)	32.7	521	(3.6)	48	(3.0)	■
Deutschland	71.1	504	(2.5)	28.9	552	(3.6)	48	(3.0)	■
3 *Hongkong*	61.1	527	(3.8)	38.9	575	(5.0)	48	(5.0)	■
23 *England*	67.6	542	(2.8)	32.4	590	(3.7)	47	(4.0)	■
Republik Korea (Südkorea)	37.5	554	(3.6)	62.5	601	(2.4)	47	(3.7)	■
2 *Ungarn*	66.6	510	(3.5)	33.4	556	(3.6)	46	(3.9)	■
2 *Frankreich*	64.5	473	(3.4)	35.5	518	(3.4)	45	(3.8)	■
Irland	59.4	514	(3.6)	40.6	558	(3.5)	44	(3.8)	■
2 *Singapur*	60.6	591	(2.6)	39.4	634	(2.0)	43	(2.2)	■
23 *Dänemark*	67.2	509	(2.5)	32.8	552	(3.0)	43	(2.9)	■
2 *Norwegen (5. Jgst.)*	68.7	520	(2.6)	31.3	560	(3.1)	40	(2.7)	■
23 *Türkei (5. Jgst.)*	70.7	559	(3.4)	29.3	599	(4.0)	40	(3.8)	■
23 *Neuseeland*	65.3	504	(3.3)	34.7	544	(3.1)	40	(3.6)	■
23 *Belgien (Fläm. Gem.)*	74.0	479	(2.5)	26.0	519	(3.1)	40	(3.2)	■
VG EU	**71.0**	**507**	**(0.7)**	**29.0**	**547**	**(0.8)**	**40**	**(0.8)**	▼
23 *USA*	71.5	522	(2.2)	28.5	562	(3.4)	40	(2.5)	▼
2 *Tschechien*	66.6	513	(2.3)	33.4	552	(2.9)	39	(2.8)	▼
Finnland	65.2	529	(3.0)	34.8	568	(2.8)	39	(2.9)	▼
2 *Polen*	68.9	538	(2.3)	31.1	576	(2.8)	38	(3.0)	▼
VG OECD	**73.9**	**514**	**(0.5)**	**26.1**	**551**	**(0.6)**	**37**	**(0.7)**	▼
2 *Belgien (Frz. Gem.)*	66.8	469	(3.0)	33.2	506	(2.8)	37	(3.1)	▼
Australien	63.6	537	(2.0)	36.4	574	(2.2)	37	(2.5)	▼
23 *Kanada*	67.9	511	(1.7)	32.1	547	(1.9)	36	(1.7)	▼
Japan	72.6	546	(2.2)	27.4	582	(3.3)	36	(2.8)	▼
Taiwan	70.2	564	(1.6)	29.8	596	(2.0)	33	(2.6)	▼
23 *Rumänien*	79.5	521	(4.9)	20.5	554	(7.0)	33	(7.0)	▼
2 *Spanien*	71.7	495	(1.8)	28.3	528	(2.4)	32	(2.0)	▼
3 *Niederlande*	73.8	509	(3.0)	26.2	541	(3.9)	32	(3.8)	▼
2 *Litauen*	80.4	531	(2.3)	19.6	563	(4.0)	32	(3.5)	▼
Macau	68.4	526	(1.3)	31.6	558	(2.3)	32	(2.3)	▼
2 *Slowenien*	70.0	517	(2.4)	30.0	549	(3.0)	32	(3.0)	▼
Portugal	75.5	504	(2.4)	24.5	534	(3.3)	31	(3.3)	▼
Lettland	80.2	521	(2.7)	19.8	548	(5.2)	27	(4.4)	▼
2 *Italien*	76.5	505	(2.8)	23.5	531	(2.9)	26	(3.5)	▼
2 *Serbien*	76.7	505	(3.2)	23.3	528	(4.4)	23	(4.0)	▼
Internationaler Mittelwert	**73.9**	**488**	**(0.4)**	**26.1**	**509**	**(0.6)**	**21**	**(0.6)**	▼
23 *Chile*	85.2	483	(2.2)	14.8	466	(5.7)	-17	(5.0)	▼
Benchmark-Teilnehmer									
2 *Ontario, Kanada*	67.6	517	(2.5)	32.4	549	(3.0)	32	(2.8)	▼
Dubai, VAE	71.2	558	(1.8)	28.8	577	(2.6)	19	(2.7)	▼



□ Keine statistisch signifikanten Unterschiede ($p > .05$)
■ Statistisch signifikante Unterschiede ($p \leq .05$)
▲ = Differenzwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Die Leistungsunterschiede in Abhängigkeit von den im Haushalt vorhandenen Büchern werden in den Abbildungen 7.5 und 7.6 für TIMSS 2007, TIMSS 2019 und TIMSS 2023 dargestellt (für die entsprechenden Werte aus TIMSS 2011 und 2015 siehe Stubbe et al., 2020). Neben den Mittelwertsdifferenzen für die drei Studienzyklen finden sich in den Abbildungen auch die Veränderungen zwischen TIMSS 2023 und den beiden vorangegangenen Erhebungen, wobei die Vier-Jahres-Differenz (Vergleich mit TIMSS 2019) zudem grafisch dargestellt wird. Die Auswahl der Staaten beschränkt sich auf diejenigen, die seit 2007 an allen TIMSS-Zyklen mit der Primarstufen-Population teilgenommen haben.

Im Bereich Mathematik (Abbildung 7.5) liegen die Leistungsdifferenzen zwischen Schüler:innen aus Familien mit mehr als 100 Büchern und aus Familien mit maximal 100 Büchern in Deutschland im Zeitverlauf zwischen 37 und 41 Punkten, wobei zu keinem Zeitpunkt eine signifikante Veränderung festgestellt werden kann. In Litauen, Neuseeland, der Slowakei und Australien lässt sich im Vergleich zur vorangegangenen Erhebung eine signifikante Verringerung der sozialen Disparitäten beobachten. In den übrigen Staaten haben sich in den vergangenen vier Jahren keine signifikanten Veränderungen der sozialen Disparitäten hinsichtlich der Mathematikkompetenz ergeben.

Für die Naturwissenschaftskompetenz (Abbildung 7.6) zeigen sich im Trend ähnliche Ergebnisse wie für die Differenzen in der Mathematikkompetenz zwischen Viertklässler:innen aus Familien mit mehr als 100 Büchern und aus Familien mit maximal 100 Büchern. Für Deutschland (zwischen 44[3] und 51 Punkten Differenz) und die meisten anderen Teilnehmer finden sich keine signifikanten Veränderungen zwischen 2019 und 2023. In keinem Staat haben sich die sozialen Disparitäten in diesem Zeitraum signifikant vergrößert, es gibt aber vier Teilnehmer, die eine signifikante Verringerung der Mittelwertsdifferenzen zu verzeichnen haben: Iran (2019: 42 Punkte; 2023: 21 Punkte), Litauen (2019: 50 Punkte; 2023: 32 Punkte), Neuseeland (2019: 52 Punkte; 2023: 40 Punkte) und Australien (2019: 45 Punkte; 2023: 37 Punkte).

Mithilfe von Regressionsanalysen kann für alle Teilnehmerstaaten der Varianzanteil der Mathematik- bzw. Naturwissenschaftskompetenz bestimmt werden, der durch die fünfstufige Variable Bücher im Haushalt erklärt wird. Die Abbildungen 7.7 und 7.8 zeigen für TIMSS 2023 den Zusammenhang zwischen dieser Varianzaufklärung und der Ausprägung der mittleren Mathematik- bzw. Naturwissenschaftskompetenz in den Teilnehmerstaaten.

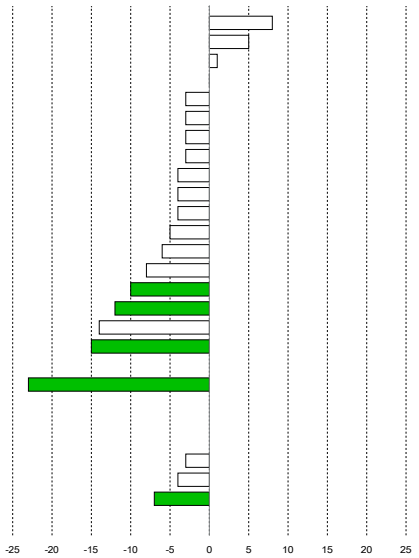
Die mittlere erklärte Varianz der Mathematikkompetenz (VG EU) beträgt 10.3 Prozent und die mittlere Mathematikkompetenz 524 Punkte. Für die Naturwissenschaftskompetenz werden im Durchschnitt (VG EU) 13.0 Prozent der Varianz erklärt und der Mittelwert liegt bei 518 Punkten.

In den oberen linken Quadranten der beiden Abbildungen finden sich diejenigen Staaten, die es schaffen, an ihren Grundschulen ein hohes Leistungsniveau zu erreichen und bei denen gleichzeitig die durch die Anzahl der Bücher im Haushalt erklärbaren Unterschiede in den erreichten Kompetenzen relativ gering ausfallen. In Abhängigkeit von der jeweiligen Situation in den Staaten kann dies als ein Hinweis auf geringe soziale Disparitäten in den Leistungen der Schüler:innen interpretiert werden. In beiden Kompetenzbereichen gehören Australien, Dänemark, Japan, Lettland, Litauen, Macau, Norwegen, Rumänien, Singapur, Taiwan, Tschechien sowie der OECD-Durchschnitt in diese Gruppe. Einen ebenfalls unterdurchschnittlichen Zusammenhang zwischen der Anzahl

3 Dieser Wert wurde für TIMSS 2011 und TIMSS 2015 beobachtet und ist aufgrund der Fokussierung in diesem Berichtsband in Abbildung 7.6 nicht dargestellt.

Abbildung 7.5: Unterschiede im Leistungsvorsprung in Mathematik von Kindern aus Familien mit mehr als 100 Büchern vor denen mit maximal 100 Büchern im Vergleich von TIMSS 2007, 2019 und 2023

Fußnoten 2007	2019	2023	T	Teilnehmer	2007 M_{07}	(SE)	2019 M_{19}	(SE)	2023 M_{23}	(SE)	Veränderungen[A] M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)	
	3		8	Japan	41	(3.5)	30	(3.3)	38	(2.7)	-3	(4.4)		8	(4.3)	■
2 3	3	3	8	Hongkong	22	(3.7)	37	(4.2)	42	(4.7)	20	(6.0)	+	5	(6.3)	■
2		2	8	Italien	15	(3.8)	19	(3.3)	20	(3.5)	5	(5.2)		1	(4.8)	■
	2	2	8	Singapur	40	(3.9)	40	(3.4)	40	(2.3)	0	(4.5)		0	(4.1)	■
		2		Tschechien	32	(3.7)	36	(3.2)	33	(2.9)	1	(4.7)		-3	(4.3)	■
		2		Ungarn	57	(3.7)	50	(3.7)	47	(4.0)	-10	(5.4)		-3	(5.4)	■
2 3	2 3	2 3		USA	35	(2.5)	40	(3.2)	37	(3.0)	2	(3.9)		-3	(4.4)	■
1 3	1 2	1 2 3	8	England	48	(4.3)	50	(4.7)	47	(4.7)	-1	(6.4)		-3	(6.6)	■
3	3	2 3	8	Dänemark	30	(3.5)	36	(2.8)	32	(2.7)	2	(4.4)		-4	(3.9)	■
				Deutschland	41	(3.1)	41	(3.6)	37	(2.7)	-4	(4.1)		-4	(4.5)	■
3	3	3	8	Niederlande	21	(3.8)	26	(3.4)	22	(3.5)	1	(5.2)		-4	(4.9)	■
	2	2		Schweden	31	(2.8)	46	(3.2)	41	(3.7)	10	(4.6)	+	-5	(4.9)	■
				Taiwan	39	(2.8)	35	(2.9)	29	(2.2)	-10	(3.6)	–	-6	(3.6)	■
2	2	2		Georgien	13	(5.5)	19	(4.8)	11	(3.8)	-2	(6.7)		-8	(6.1)	■
				Australien	33	(4.0)	50	(3.3)	40	(2.5)	7	(4.7)		-10	(4.1)	– ■
	2			Slowakei	31	(3.9)	49	(4.3)	37	(3.6)	6	(5.3)		-12	(5.6)	– ■
		4	8	Iran	44	(7.3)	35	(7.3)	21	(5.7)	-23	(9.3)	–	-14	(9.3)	■
1 2	1 2	1 2 3	8	Neuseeland	44	(3.2)	55	(3.9)	40	(3.2)	-4	(4.5)		-15	(5.0)	– ■
2	2	2	8	Litauen	22	(4.7)	51	(4.3)	28	(3.6)	6	(5.9)		-23	(5.6)	– ▼
				Benchmark-Teilnehmer												
2	3	2 3	8	Québec, Kanada	18	(3.9)	26	(3.4)	23	(2.2)	5	(4.5)		-3	(4.0)	■
2		2	8	Ontario, Kanada	28	(3.7)	28	(5.8)	24	(3.0)	-4	(4.8)		-4	(6.5)	■
2 3 6	2		8	Dubai, VAE	33	(5.0)	22	(2.4)	15	(2.4)	-18	(5.5)	–	-7	(3.4)	– ■

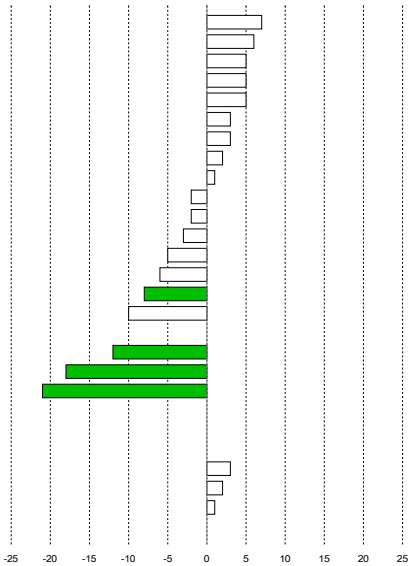


□ Keine statistisch signifikanten Unterschiede ($p > .05$)
■ Statistisch signifikante Unterschiede ($p \leq .05$)
▲ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
\+ = Mittelwert in 2023 signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
4 = Sehr hoher Anteil an Schüler:innen mit nicht skalierbaren Leistungswerten
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Abbildung 7.6: Unterschiede im Leistungsvorsprung in Naturwissenschaften von Kindern aus Familien mit mehr als 100 Büchern vor denen mit maximal 100 Büchern im Vergleich von TIMSS 2007, 2019 und 2023

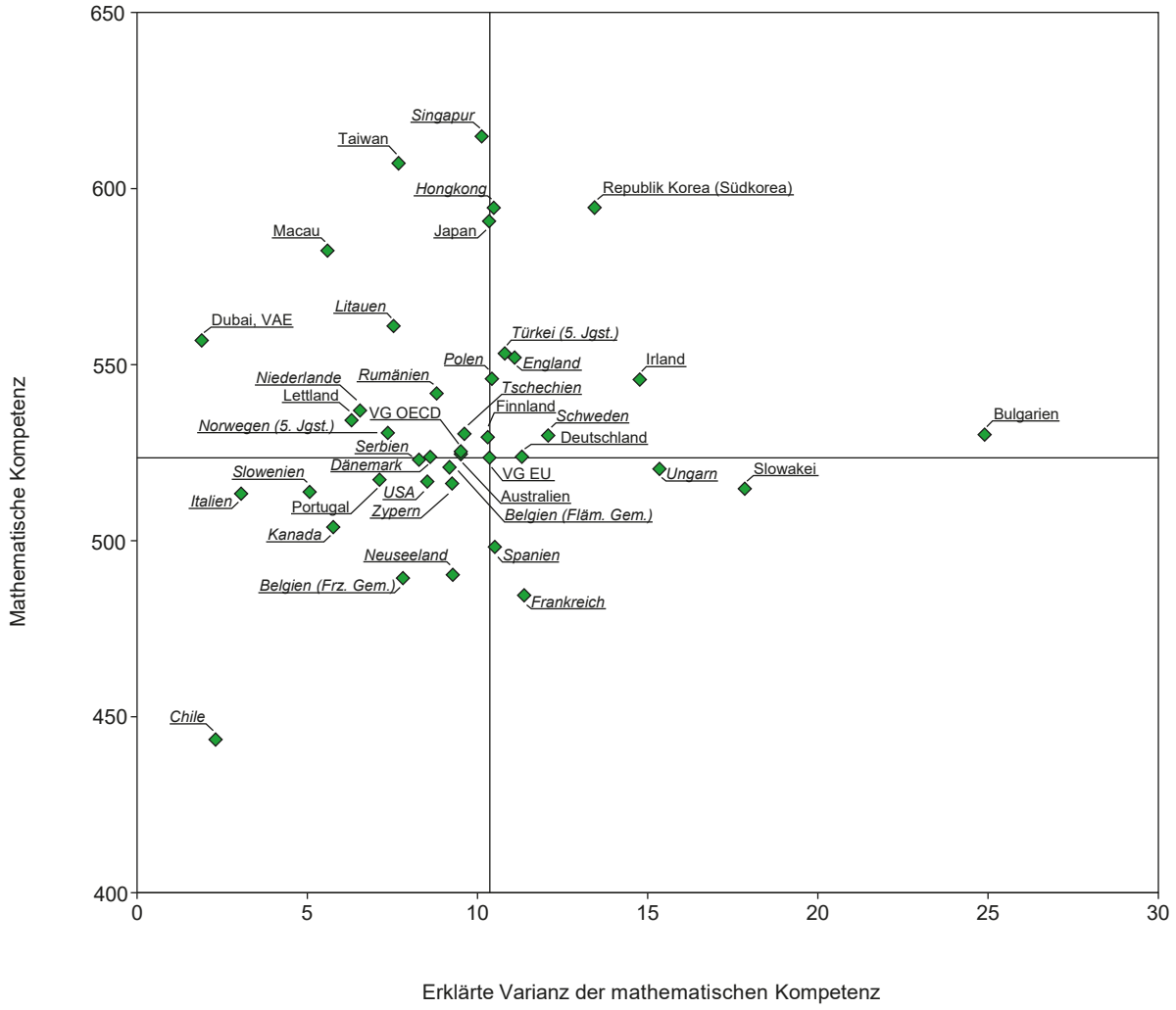
Fußnoten 2007	2019	2023	T	Teilnehmer	2007 M_{07}	(SE)	2019 M_{19}	(SE)	2023 M_{23}	(SE)	Veränderungen[A] M_{23}-M_{07}	(SE)		M_{23}-M_{19}	(SE)	
2 3	3	3	8	Hongkong	16	(4.1)	41	(4.0)	48	(5.0)	32	(6.5)	+	7	(6.4)	■
	2			Tschechien	33	(3.8)	33	(3.0)	39	(2.8)	6	(4.7)		6	(4.1)	■
3	3	2 3	8	Dänemark	37	(4.1)	38	(3.2)	43	(2.9)	6	(5.0)		5	(4.3)	■
2		2	8	Italien	22	(3.8)	21	(3.4)	26	(3.5)	4	(5.2)		5	(4.9)	■
3	3	3	8	Niederlande	30	(3.6)	27	(4.2)	32	(3.8)	2	(5.2)		5	(5.7)	■
	3		8	Japan	34	(3.6)	33	(3.3)	36	(2.8)	2	(4.6)		3	(4.3)	■
	2	2		Schweden	41	(3.5)	51	(3.4)	54	(4.6)	13	(5.8)	+	3	(5.7)	■
1 3	1 2	1 2 3	8	England	47	(3.9)	45	(3.9)	47	(4.0)	0	(5.6)		2	(5.6)	■
				Deutschland	51	(3.7)	47	(3.5)	48	(3.0)	-3	(4.8)		1	(4.6)	■
				Taiwan	43	(3.5)	35	(2.9)	33	(2.6)	-10	(4.4)	–	-2	(3.9)	■
2 3	2 3	2 3		USA	37	(3.0)	42	(3.6)	40	(2.5)	3	(3.9)		-2	(4.4)	■
	2			Ungarn	53	(3.5)	49	(3.5)	46	(3.9)	-7	(5.2)		-3	(5.2)	■
	2	2	8	Singapur	48	(4.4)	48	(3.1)	43	(2.2)	-5	(4.9)		-5	(3.8)	■
	2			Slowakei	34	(4.0)	55	(5.0)	49	(4.1)	15	(5.7)	+	-6	(6.5)	■
				Australien	39	(3.5)	45	(3.2)	37	(2.5)	-2	(4.3)		-8	(4.1)	– ■
2	2	2		Georgien	24	(5.2)	22	(5.1)	12	(4.0)	-12	(6.6)		-10	(6.5)	■
1 2	1 2	1 2 3	8	Neuseeland	51	(3.3)	52	(3.1)	40	(3.6)	-11	(4.9)	–	-12	(4.8)	– ▼
2	2	2	8	Litauen	18	(4.4)	50	(4.2)	32	(3.5)	14	(5.6)	+	-18	(5.5)	– ▼
			8	Iran	54	(7.7)	42	(6.8)	21	(7.2)	-33	(10.5)	–	-21	(9.9)	– ▼
				Benchmark-Teilnehmer												
2		2	8	Ontario, Kanada	36	(4.3)	29	(4.5)	32	(2.8)	-4	(5.1)		3	(5.3)	■
2 3 6	2		8	Dubai, VAE	26	(6.4)	17	(3.0)	19	(2.7)	-7	(6.9)		2	(4.0)	■
2	3	2 3	8	Québec, Kanada	19	(3.4)	30	(3.9)	31	(2.0)	12	(3.9)	+	1	(4.4)	■



□ Keine statistisch signifikanten Unterschiede ($p > .05$)
■ Statistisch signifikante Unterschiede ($p \leq .05$)
▲ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwerte im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
\+ = Mittelwert in 2023 signifikant höher als 2007 bzw. 2019 ($p \leq .05$)
– = Mittelwert in 2023 signifikant niedriger als 2007 bzw. 2019 ($p \leq .05$)
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
6 = Abweichender Testzeitpunkt
8 = Eingeschränkte Vergleichbarkeit aufgrund veränderter Teilnahmebedingungen seit 2007
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Abbildung 7.7: Zusammenhang zwischen der erklärten Varianz der Mathematikkompetenz auf Individualebene durch die Anzahl der Bücher im Haushalt und den mittleren Mathematikkompetenzen in den Teilnehmerstaaten



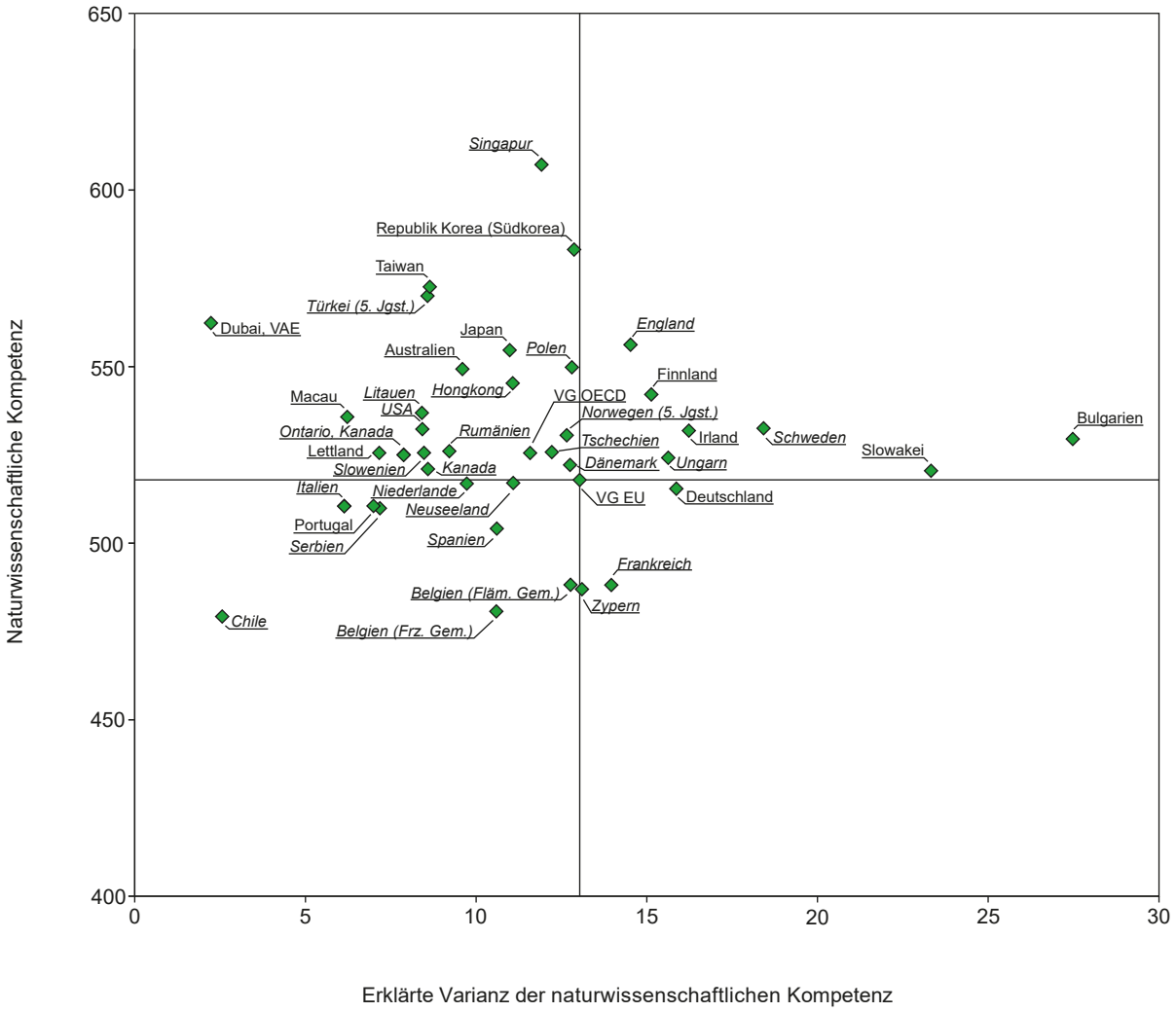
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

der Bücher im Haushalt und den individuellen Kompetenzen, aber gleichzeitig unterdurchschnittlichen Leistungen, weisen in beiden Domänen Chile, die Flämische Gemeinschaft in Belgien, die Französische Gemeinschaft in Belgien, Italien, Neuseeland, Portugal und Serbien auf (untere linke Quadranten).

In den jeweils rechten Quadranten befinden sich die Teilnehmer mit einem überdurchschnittlichen Anteil an erklärter Varianz, was auf eine enge Koppelung zwischen sozialer Herkunft und Leistung hinweisen kann. Besonders hohe Werte erreichen diesbezüglich Bulgarien und die Slowakei, wobei Bulgarien in beiden Domänen und die Slowakei in Naturwissenschaften überdurchschnittliche Leistungen aufweist. Im oberen rechten Quadranten befinden sich in beiden Kompetenzbereichen zudem England, Irland und Schweden. Die Kombination aus unterdurchschnittlichen Kompetenzen und überdurchschnittlichen Disparitäten (unterer rechter Quadrant) findet sich erfreulicherweise nur bei wenigen Teilnehmern. Für Mathematik sind dies Frankreich, die Slowakei, Spanien sowie Ungarn und für die Naturwissenschaften Deutschland, Frankreich und Zypern.

Abbildung 7.8: Zusammenhang zwischen der erklärten Varianz der Naturwissenschaftskompetenz auf Individualebene durch die Anzahl der Bücher im Haushalt und den mittleren Naturwissenschaftskompetenzen in den Teilnehmerstaaten



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

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

7.5 Soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen in Deutschland

Wie in Abschnitt 7.3 beschrieben sind die unterschiedlichen Variablen zum Berufsstatus von Personen geeignete Indikatoren für den sozialen Status von Familien. In Deutschland wurden die Erziehungsberechtigten daher in offener Form nach ihren Berufen, ihrer Tätigkeit, ihrer Weisungsbefugnis und der Betriebsgröße gefragt, um anschließend EGP-Klassen und den ISEI berechnen zu können. Auch das Merkmal *Armutsgefährdung* (siehe Abschnitt 7.3) wurde nur in Deutschland erhoben. Daher werden in diesem Abschnitt Befunde zu diesen drei Hintergrundmerkmalen der Viertklässler:innen auf nationaler Ebene berichtet.

Der Zusammenhang zwischen dem höchsten ISEI der Erziehungsberechtigten und den erreichten Leistungen lässt sich mithilfe von Regressionsanalysen

Tabelle 7.1: Soziale Gradienten für die mathematischen und naturwissenschaftlichen Kompetenzen in Deutschland bei TIMSS 2007, 2019 und 2023

	Mathematik		Naturwissenschaften	
	sozialer Gradient	(*SE*)	sozialer Gradient	(*SE*)
2023	1.40	(0.067)	1.58	(0.083)
2019	1.27	(0.096)	1.39	(0.087)
2007	1.41	(0.088)	1.63	(0.101)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

darstellen. In Tabelle 7.1 ist der soziale Gradient (Steigung der Regressionsgeraden b_1) für die abhängigen Variablen Mathematik- und Naturwissenschaftskompetenz in den Jahren 2007, 2019 und 2023 angegeben (für die entsprechenden Werte aus TIMSS 2011 und 2015 siehe Stubbe et al., 2020).

Relativ zu der Ausgangserhebung im Jahr 2007 waren die beiden sozialen Gradienten bei TIMSS 2011 und 2015 signifikant kleiner, während für TIMSS 2019 und 2023 keine signifikanten Differenzen festgestellt werden konnten. Auch zwischen TIMSS 2019 und 2023 gab es keine signifikante Veränderung. Nominell fallen die Werte in der aktuellen Erhebung allerdings noch einmal höher aus als vor vier Jahren und liegen ziemlich exakt auf dem Niveau der Ausgangserhebung. In den vergangenen 16 Jahren konnte somit zunächst ein leichter Rückgang der sozialen Disparitäten beobachtet werden und anschließend ein Anstieg im selben Umfang.

Für den zweiten Indikator des Berufsstatus – die EGP-Klassen – muss zunächst auf das Erhebungsinstrument im Rahmen von TIMSS 2023 eingegangen werden, da dieses von dem Vorgehen der vorangegangenen Zyklen abweicht. Bis einschließlich TIMSS 2019 wurde im Elternfragebogen nach dem Beruf des Vaters und der Mutter gefragt. Seit TIMSS 2023 wird der Beruf von Elternteil A und Elternteil B erhoben. Für die Auswertung des ISEI hat diese Änderung keine Auswirkungen, da in der Bildungsforschung üblicherweise der höchste ISEI im Haushalt (HISEI) betrachtet wird, wobei unerheblich ist, um welches konkrete Elternteil es sich dabei handelt (siehe Abschnitt 7.3).

Anders verhält es sich bei den EGP-Klassen. Da diese Variable nominalskaliert ist, ist es nicht zulässig, von einer höheren (oder niedrigeren) Ausprägung zu sprechen. Um dennoch einen einzelnen Messwert pro Familie nutzen zu können, hat sich in der empirischen Bildungsforschung das Vorgehen etabliert, dass zunächst die EGP-Klasse des Vaters berücksichtigt wird und nur, wenn diese Angabe fehlt, die EGP-Klasse der Mutter ergänzt wird (vgl. aktuell Mang et al., 2023; Stubbe et al., 2023). Da im Rahmen von TIMSS 2023 nicht festgestellt werden kann, welche Berufsangabe von welchem Elternteil stammt, wird im Folgenden die EGP-Klasse von Elternteil A berichtet und nur wenn diese Angabe fehlt die EGP-Klasse von Elternteil B. Dies führt zu nicht unerheblichen Verschiebungen hinsichtlich der Verteilung der Schüler:innen auf die einzelnen Gruppen. Um dies zu verdeutlichen, werden in Tabelle 7.2 für die bisherigen TIMSS-Zyklen die Anteile der einzelnen EGP-Klassen sowie der mittlere HISEI pro EGP-Klasse berichtet.

Tabelle 7.2: Anteile der einzelnen EGP-Klassen in den Familien und mittlerer HISEI pro EGP-Klasse bei TIMSS 2007, 2011, 2015, 2019 und 2023

Sozioökonomische Stellung der Familie (EGP-Klasse)	Anteile in Prozent									
	2007		2011		2015		2019		2023	
	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
Obere Dienstklasse (I)	20.9	(1.0)	25.1	(1.2)	25.5	(1.1)	25.9	(1.2)	20.7	(0.8)
Untere Dienstklasse (II)	24.2	(0.8)	20.8	(0.7)	19.6	(0.9)	20.4	(1.0)	35.0	(0.7)
Routinedienstleistungen (III)	6.1	(0.4)	6.4	(0.5)	7.6	(0.5)	8.1	(0.7)	23.1	(0.7)
Selbstständige (IV)	11.6	(0.6)	8.2	(0.6)	7.6	(0.6)	7.7	(0.7)	4.6	(0.4)
(Fach-)Arbeiter:innen (V, VI)	20.9	(0.8)	19.8	(0.9)	21.6	(1.0)	18.1	(0.9)	7.3	(0.4)
Un- und angelernte Arbeiter:innen (VII)	16.3	(0.9)	19.6	(1.0)	18.1	(1.2)	19.8	(1.1)	9.4	(0.4)
Sozioökonomische Stellung der Familie (EGP-Klasse)	**Berufsstatus (mittlerer HISEI)**									
	2007		2011		2015		2019		2023	
	M	(*SE*)	*M*	(*SE*)	*M*	(*SE*)	*M*	(*SE*)	*M*	(*SE*)
Obere Dienstklasse (I)	72.4	(0.4)	72.2	(0.6)	72.7	(0.5)	73.4	(0.5)	73.6	(0.5)
Untere Dienstklasse (II)	55.1	(0.4)	60.4	(0.6)	60.7	(0.8)	63.1	(0.7)	62.2	(0.6)
Routinedienstleistungen (III)	46.2	(0.8)	45.0	(1.4)	44.6	(1.3)	46.3	(1.1)	48.7	(0.5)
Selbstständige (IV)	47.4	(0.7)	40.7	(1.1)	44.4	(1.5)	49.0	(1.9)	43.1	(1.5)
(Fach-)Arbeiter:innen (V, VI)	41.1	(0.4)	40.3	(0.6)	43.6	(0.8)	43.0	(1.0)	40.9	(0.9)
Un- und angelernte Arbeiter:innen (VII)	34.7	(0.5)	32.8	(0.7)	35.5	(0.8)	35.9	(0.8)	30.2	(0.8)

Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (weiterführende Informationen dazu finden sich im Text).
Rücklaufquoten des Elternfragebogens: TIMSS 2007: 77 %; TIMSS 2011: 80 %; TIMSS 2015: 63 %; TIMSS 2019: 67 %; TIMSS 2023: 60 %

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Zwischen TIMSS 2007 und 2019 schwanken die Anteile in den einzelnen EGP-Klassen um maximal 5 Prozentpunkte. Diese Schwankungen begründen sich zum Teil durch die Stichprobenziehung und in geringerem Maße möglicherweise durch tatsächliche Veränderungen in der Gesellschaft. Eine wichtige Ursache sind darüber hinaus die im Zeitverlauf deutlich gesunkenen Rücklaufquoten der Elternfragebögen. Da die Erfahrungen zeigen, dass Fragebögen überproportional häufig von Personen mit niedrigem sozialen Status nicht ausgefüllt werden (Menold & Züll, 2010) und dieser Ausfall somit systematisch ist, führt dies entsprechend auch zu Verzerrungen in den in diesem Kapitel betrachteten Herkunftsvariablen.

Für TIMSS 2023 zeigt sich eine deutlich veränderte Häufigkeitsverteilung der EGP-Klassen. Verglichen mit dem Mittel der vier vorangegangenen Erhebungen ergibt sich für die EGP-Klasse III ein Anstieg um rund 16 Prozentpunkte und für die EGP-Klasse II ein Anstieg um knapp 14 Prozentpunkte. In den übrigen Gruppen sind die Anteile 2023 entsprechend geringer als in den Vorjahren (V, VI: knapp 13 Prozentpunkte; VII: rund 9 Prozentpunkte; IV: rund 4 Prozentpunkte; I: knapp 4 Prozentpunkte). Ausgehend von dem veränderten Erhebungsinstrument sind diese Veränderungen allerdings erwartungskonform. In den EGP-Klassen II und III ist der Frauenanteil deutlich größer als der Männeranteil. In den übrigen EGP-Klassen gibt es insgesamt mehr Männer, wobei der Unterschied

in den EGP-Klassen V, VI und VII besonders groß ist.[4] Dadurch, dass in der Vergangenheit im Regelfall die Angaben der Väter in die Analysen eingeflossen sind, ist für die Jahre 2007 bis 2019 von einer systematischen Verzerrung zugunsten der EGP-Klassen I, IV, V, VI und VII auszugehen. Eine detaillierte Analyse der aktuellen Daten führt allerdings zu dem Schluss, dass die Verteilung für TIMSS 2023 keineswegs unverzerrt ist, sondern systematisch zugunsten der EGP-Klassen II und III verzerrt ist, da bei Elternteil A die meisten Eltern offenbar den Beruf der Mutter angegeben haben.

Um abschätzen zu können, welche Auswirkungen diese veränderte Verteilung auf die weiteren Analysen hat, kann der mittlere HISEI pro Klasse im Zeitverlauf betrachtet werden (siehe Tabelle 7.2). Verglichen mit den Schwankungen in den vorangegangenen Zyklen sind die Veränderungen der Mittelwerte bei TIMSS 2023 nicht gravierend. Dennoch muss bei allen folgenden Trendanalysen mit den EGP-Klassen bedacht werden, dass die Daten nur eingeschränkt vergleichbar sind.

Tabelle 7.3 zeigt die mittlere Mathematik- und Naturwissenschaftskompetenz der Schüler:innen in Abhängigkeit von der EGP-Klassenzugehörigkeit der Erziehungsberechtigten. Kinder, deren Erziehungsberechtigte der oberen Dienstklasse (I) angehören, erreichen in Mathematik 50 Punkte und in den Naturwissenschaften 56 Punkte mehr als ihre Mitschüler:innen, deren Erziehungsberechtigte der Gruppe der (Fach-)Arbeiter:innen (V & VI) angehören. Dies entspricht in Mathematik einem Vorsprung von rund einem Lernjahr und in den Naturwissenschaften von mehr als anderthalb Lernjahren (Wendt et al., 2017). Die Differenz zu Kindern von un- und angelernten Arbeiter:innen (VII) beträgt sogar 68 Punkte (Mathematik) bzw. 75 Punkte (Naturwissenschaften).

Tabelle 7.3: Mittlere mathematische und naturwissenschaftliche Kompetenzen pro EGP-Klasse in Deutschland

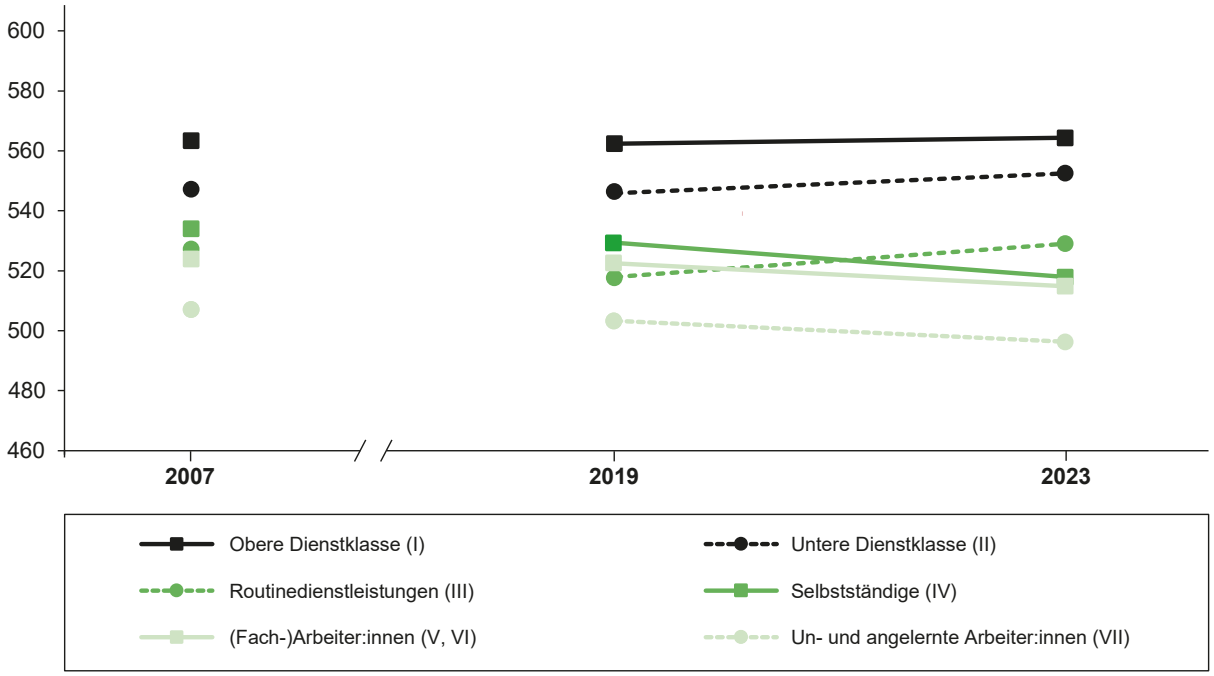
Sozioökonomische Stellung der Familie (EGP-Klasse)	Mathematik				Naturwissenschaften			
	M	*(SE)*	*SD*	*(SE)*	*M*	*(SE)*	*SD*	*(SE)*
Obere Dienstklasse (I)	565	(2.9)	68	(2.0)	562	(3.7)	72	(2.9)
Untere Dienstklasse (II)	553	(2.4)	68	(1.6)	551	(2.8)	72	(1.8)
Routinedienstleistungen (III)	529	(3.0)	72	(2.0)	524	(3.4)	76	(2.1)
Selbstständige (IV)	518	(6.8)	67	(4.6)	512	(7.5)	81	(5.4)
(Fach-)Arbeiter:innen (V, VI)	515	(4.9)	76	(4.3)	506	(5.8)	85	(4.3)
Un- und angelernte Arbeiter:innen (VII)	497	(4.8)	78	(3.5)	487	(6.6)	84	(3.4)
Gesamt	540	(1.7)	73	(1.1)	536	(2.6)	79	(1.3)

Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (weiterführende Informationen dazu finden sich im Text).

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

4 Im Rahmen der Berichtslegung zu TIMSS 2023 wurden vor dem Hintergrund der geänderten Abfrage der Elternberufe die Anteile der einzelnen EGP-Klassen für die vergangenen TIMSS-Zyklen getrennt für Mütter und Väter berechnet. Dabei haben sich die berichteten Unterschiede gezeigt.

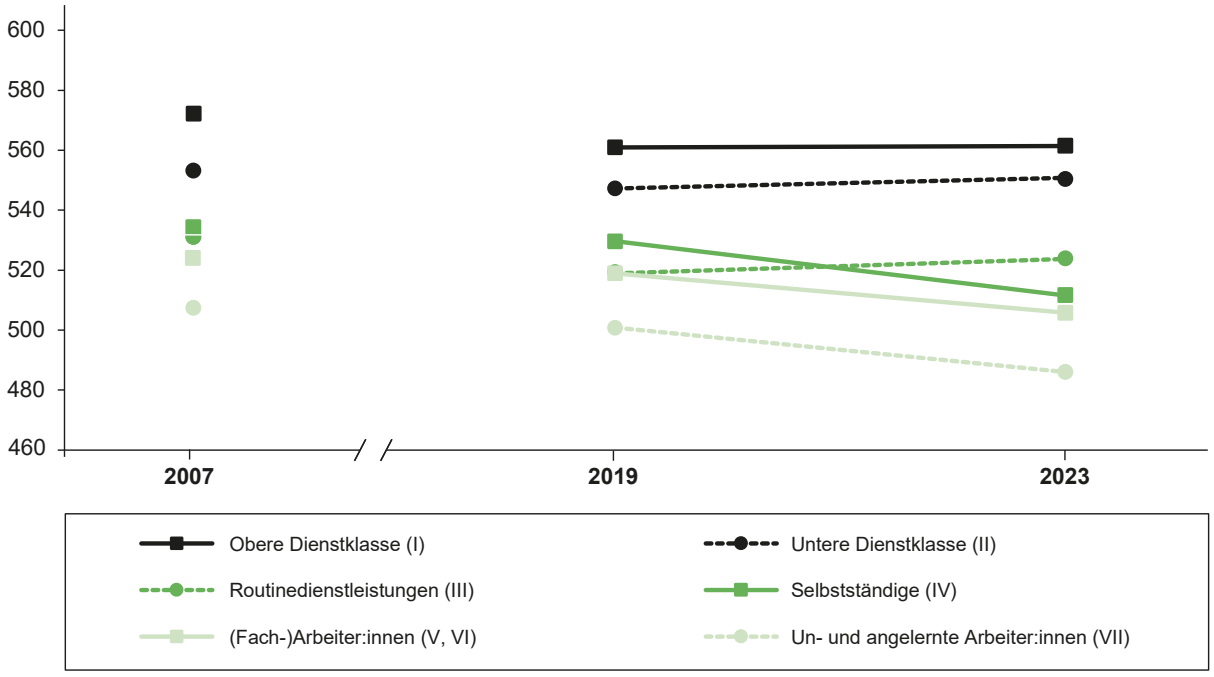
Abbildung 7.9: Mittlere mathematische Kompetenzen pro EGP-Klasse in Deutschland bei TIMSS 2007, 2019 und 2023



Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (weiterführende Informationen dazu finden sich im Text).

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Abbildung 7.10: Mittlere naturwissenschaftliche Kompetenzen pro EGP-Klasse in Deutschland bei TIMSS 2007, 2019 und 2023



Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (weiterführende Informationen dazu finden sich im Text).

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

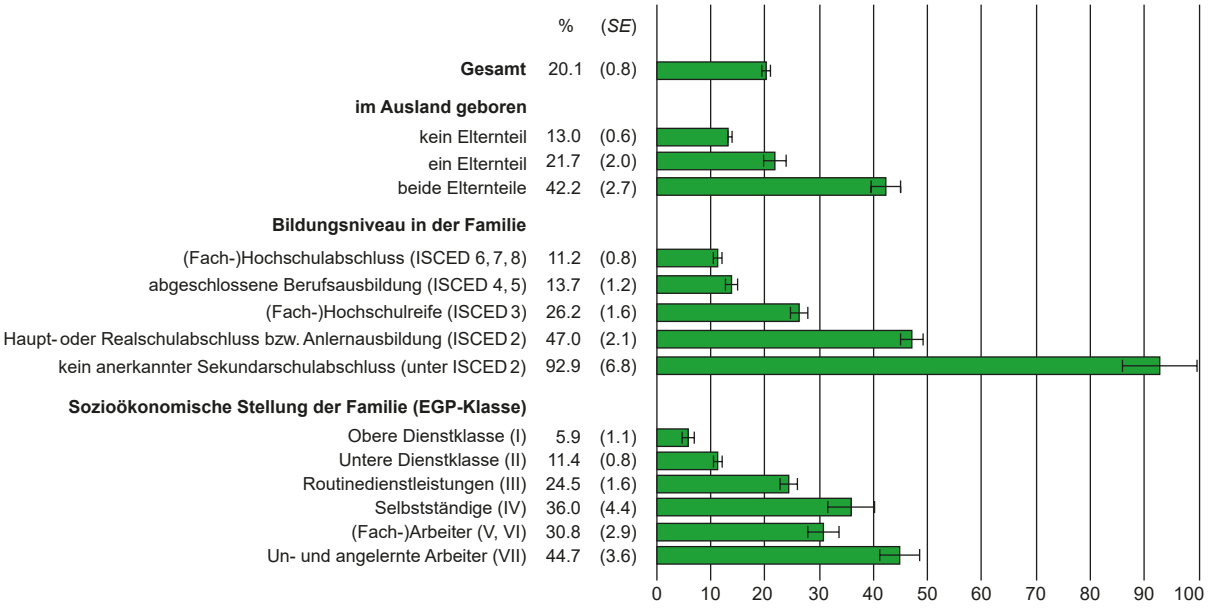
Die Veränderungen in den Kompetenzen pro EGP-Klasse zwischen TIMSS 2019 und 2023 werden in den Abbildungen 7.9 (Mathematik) und 7.10 (Naturwissenschaften) visualisiert. Als Referenz finden sich in den beiden Abbildungen auch die Werte für TIMSS 2007 (für die entsprechenden Werte aus TIMSS 2011 und 2015 siehe Stubbe et al., 2020). In jedem Zyklus zeigt sich eindeutig, dass die mathematischen und naturwissenschaftlichen Kompetenzen in der oberen Dienstklasse (I) am höchsten und in der unteren Dienstklasse (II) am zweithöchsten sind. In der Gruppe der un- und angelernten Arbeiter:innen (VII) sind die Kompetenzen in jeder TIMSS-Erhebung am geringsten. In den vergangenen vier Jahren haben sich die Leistungsdifferenzen zwischen den Dienstklassen und den un- und angelernten Arbeiter:innen etwas vergrößert. In den Dienstklassen ist ein sehr geringer Anstieg der Kompetenzen zu verzeichnen, während die Werte bei den un- und angelernten Arbeiter:innen um 7 Punkte (Mathematik) bzw. 15 Punkte (Naturwissenschaften) gesunken sind. Keine dieser Veränderungen ist allerdings statistisch signifikant.

Zwischen den Gruppen Routinedienstleistungen, Selbstständige und (Fach-) Arbeiter:innen konnten in der Vergangenheit in beiden Kompetenzbereichen nur sehr geringe Leistungsunterschiede festgestellt werden. Tendenziell waren die Werte bei den Selbstständigen am größten und bei den (Fach-)Arbeiter:innen am geringsten. Auch bei TIMSS 2023 können zwischen diesen drei Gruppen keine signifikanten Differenzen festgestellt werden und auch die Veränderungen zu TIMSS 2019 fallen nicht signifikant aus. Nominell zeigt sich allerdings eine positive Entwicklung in der Gruppe der Routinedienstleistungen, während die mittleren Kompetenzen in den beiden anderen EGP-Klassen niedriger ausfallen als in TIMSS 2019. Diese Unterschiede beruhen höchstwahrscheinlich auf dem geänderten Erhebungsverfahren, sodass insgesamt keine bedeutsamen Veränderungen im Zeitverlauf zu verzeichnen sind.

Wie Abbildung 7.11 zu entnehmen ist, beträgt die Armutsgefährdungsquote für TIMSS 2023 insgesamt 20.1 Prozent und liegt damit geringfügig – aber nicht signifikant – unter dem Wert aus TIMSS 2019 (21.5 %). Theoriekonform unterscheiden sich die Armutsgefährdungsquoten deutlich zwischen verschiedenen gesellschaftlichen Gruppen. Während die Quote für Kinder, deren Eltern beide in Deutschland geboren wurden, nur 13.0 Prozent beträgt, sind 42.2 Prozent armutsgefährdet, wenn beide Elternteile im Ausland geboren wurden. Weist mindestens ein Elternteil eine abgeschlossene Berufsausbildung oder einen (Fach-)Hochschulabschluss auf (mindestens ISCED-Level 4), so liegt die Quote nur knapp über 10 Prozent. Verfügen die Erziehungsberechtigten über einen Bildungsabschluss auf ISCED-Level 2, beträgt die Quote hingegen fast 50 Prozent. Werden ergänzend die EGP-Klassen betrachtet, gelten in Familien der *service class* nur 5.9 Prozent (Obere Dienstklasse) bzw. 11.4 Prozent (Untere Dienstklasse) als armutsgefährdet, während dies in den übrigen EGP-Klassen für 24.5 bis 44.7 Prozent der Familien zutrifft.

Wie Tabelle 7.4 entnommen werden kann, liegt die Mathematikkompetenz von Kindern aus nicht armutsgefährdeten Elternhäusern 48 Punkte (rund ein Lernjahr) über der Mathematikkompetenz von armutsgefährdeten Schüler:innen. Für den Bereich Naturwissenschaften fällt diese Differenz mit 63 Punkten (rund zwei Lernjahre) größer aus. Im Vergleich mit TIMSS 2019 (Stubbe et al., 2020) haben sich diese Differenzen signifikant ($p < .05$) vergrößert. Auch innerhalb der einzelnen EGP-Klassen gibt es einige Veränderungen hinsichtlich des Leistungsvorsprungs von nicht armutsgefährdeten Viertklässler:innen. Diese fallen bedingt durch die großen Standardfehler allerdings nur in der Gruppe der

Abbildung 7.11: Armutsgefährdungsquoten nach Migrationshintergrund, höchstem Bildungsabschluss und EGP-Klasse in Deutschland (in Prozent)



Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (weiterführende Informationen dazu finden sich im Text).

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

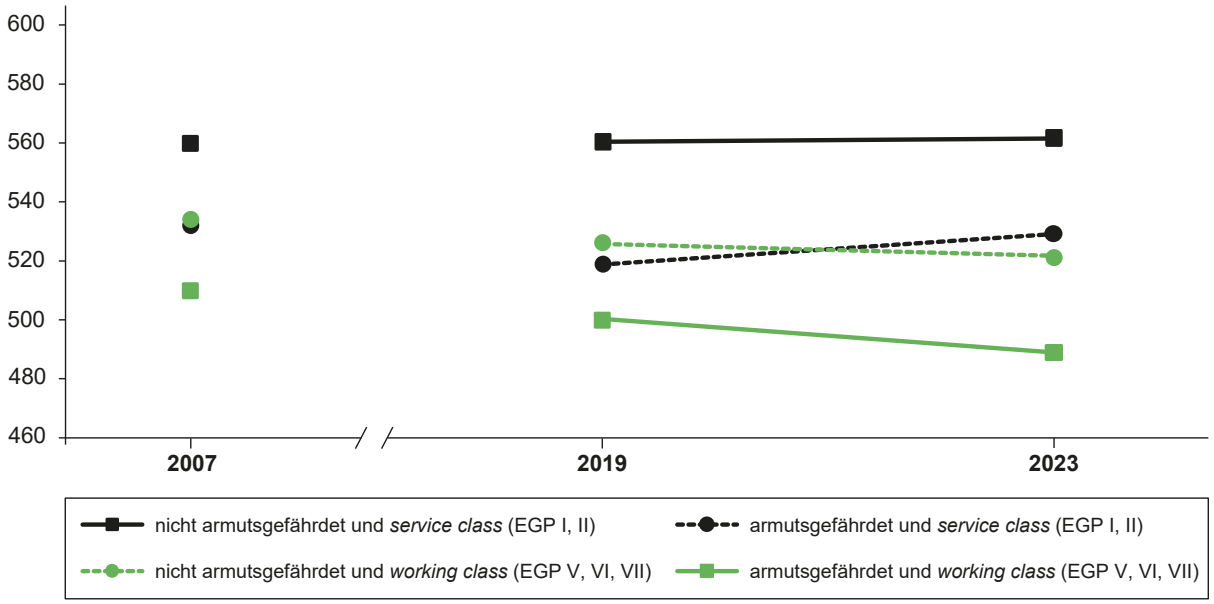
Tabelle 7.4: Mittlere mathematische und naturwissenschaftliche Kompetenzen nach *Armutsgefährdung* und EGP-Klasse in Deutschland

Sozioökonomische Stellung der Familie (EGP-Klasse)	Mathematik									
	nicht armutsgefährdet				armutsgefährdet				Differenz[A]	
	M_n	(*SE*)	*SD*	(*SE*)	M_a	(*SE*)	*SD*	(*SE*)	M_n-M_a	(*SE*)
Obere Dienstklasse (I)	568	(3.1)	65	(2.1)	520	(12.6)	83	(9.8)	48	(13.0)
Untere Dienstklasse (II)	557	(2.5)	66	(1.5)	534	(7.8)	73	(5.2)	23	(8.2)
Routinedienstleistungen (III)	545	(3.2)	64	(2.2)	504	(8.3)	79	(5.5)	41	(8.9)
Selbstständige (IV)	527	(7.1)	63	(5.3)	503	(12.6)	71	(7.8)	25	(14.5)
(Fach-)Arbeiter:innen (V, VI)	531	(6.6)	73	(4.1)	485	(11.5)	82	(10.7)	46	(13.3)
Un- und angelernte Arbeiter:innen (VII)	513	(7.0)	72	(6.5)	488	(7.5)	82	(5.6)	25	(10.3)
Gesamt	552	(1.7)	68	(0.9)	503	(3.9)	79	(2.4)	48	(4.3)
Sozioökonomische Stellung der Familie (EGP-Klasse)	**Naturwissenschaften**									
	nicht armutsgefährdet				armutsgefährdet				Differenz[A]	
	M_n	(*SE*)	*SD*	(*SE*)	M_a	(*SE*)	*SD*	(*SE*)	M_n-M_a	(*SE*)
Obere Dienstklasse (I)	569	(3.7)	67	(2.9)	489	(14.4)	87	(11.1)	80	(14.9)
Untere Dienstklasse (II)	557	(2.9)	67	(2.0)	527	(9.3)	88	(5.4)	30	(9.7)
Routinedienstleistungen (III)	539	(3.1)	70	(2.2)	495	(8.8)	83	(6.0)	44	(9.3)
Selbstständige (IV)	530	(8.0)	65	(4.4)	481	(16.2)	102	(9.2)	49	(18.0)
(Fach-)Arbeiter:innen (V, VI)	527	(6.8)	79	(4.7)	471	(11.1)	92	(8.0)	55	(13.1)
Un- und angelernte Arbeiter:innen (VII)	509	(8.6)	81	(6.8)	468	(10.9)	90	(6.2)	41	(13.9)
Gesamt	551	(2.3)	72	(1.4)	488	(5.3)	92	(2.3)	63	(5.8)

Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (weiterführende Informationen dazu finden sich im Text).
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

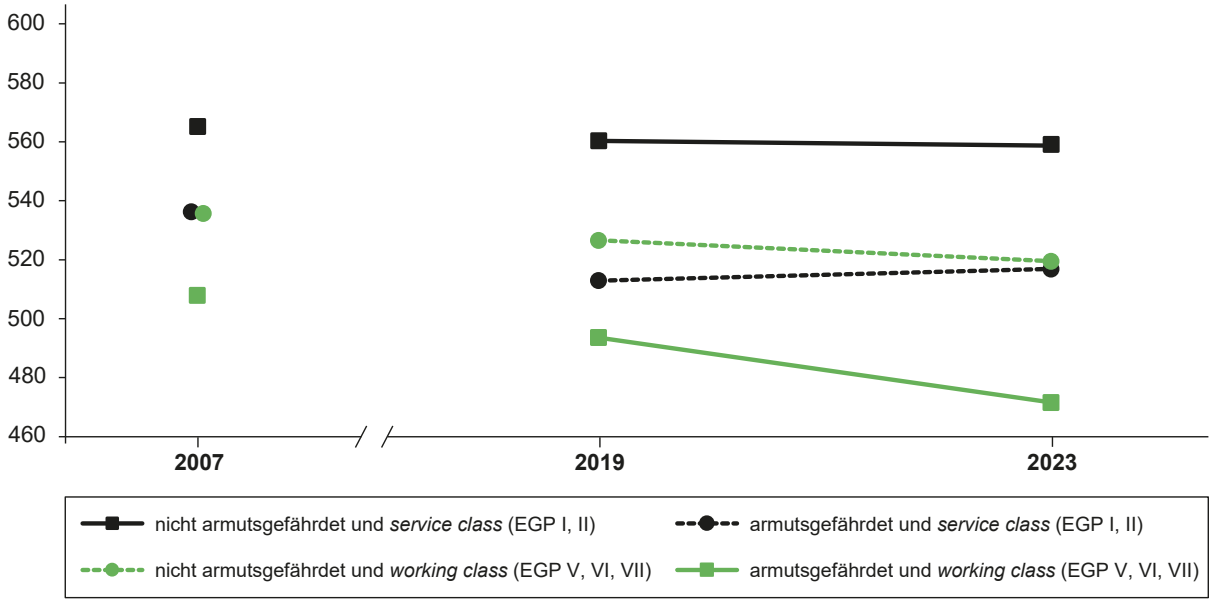
Abbildung 7.12: Mittlere mathematische Kompetenzen nach Armutsgefährdung und EGP-Klasse in Deutschland bei TIMSS 2007, 2019 und 2023



Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (weiterführende Informationen dazu finden sich im Text).

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Abbildung 7.13: Mittlere naturwissenschaftliche Kompetenzen nach Armutsgefährdung und EGP-Klasse in Deutschland bei TIMSS 2007, 2019 und 2023



Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (weiterführende Informationen dazu finden sich im Text).

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

(Fach-)Arbeiter:innen (V, VI) signifikant ($p < .05$) aus, was höchstwahrscheinlich an dem geänderten Erhebungsverfahren liegt.

Die Abbildungen 7.12 (Mathematikkompetenz) und 7.13 (Naturwissenschaftskompetenz) zeigen für vier gesellschaftliche Gruppen die Veränderungen in den Kompetenzen zwischen 2019 und 2023. Als Referenz finden sich in den beiden Abbildungen auch die Werte für TIMSS 2007 (für die entsprechenden Werte aus TIMSS 2011 und 2015 siehe Stubbe et al., 2020). Die mit Abstand höchsten Kompetenzwerte werden von Viertklässler:innen erreicht, deren Erziehungsberechtigte der *service class* (EGP-Klassen I und II) angehören und nicht armutsgefährdet sind (zwischen 557 und 561 Punkten in Mathematik und zwischen 562 und 568 Punkten in Naturwissenschaften). Schüler:innen, deren Erziehungsberechtigte der *working class* (EGP-Klassen V, VI und VII) angehören und armutsgefährdet sind, schneiden gegenüber der vorherigen Gruppe in der Mathematikkompetenz zwischen TIMSS 2007 und 2019 rund 50 bis 60 Punkte schlechter ab (2023 sogar ca. 70 Punkte). Im Bereich Naturwissenschaften betragen diese Differenzen in den Jahren 2007 bis 2019 ca. 60 bis 70 Punkte (2023 sogar ca. 90 Punkte). Zwischen den beiden anderen Gruppen – ‚armutsgefährdet und *service class*' sowie ‚nicht armutsgefährdet und *working class*' – zeigen sich in den Jahren 2007 und 2015 in beiden Kompetenzbereichen praktisch keine Leistungsunterschiede. In den Jahren 2011 (beide Domänen) und 2019 (Mathematikkompetenz) liegen die Werte für die Gruppe ‚armutsgefährdet und *service class*' nominell rund 8 Punkte unter denjenigen für die Gruppe ‚nicht armutsgefährdet und *working class*' (nicht signifikant). Für die Naturwissenschaftskompetenz bei TIMSS 2019 zeigt sich mit einer Differenz von 14 Punkten sogar ein signifikanter Unterschied zugunsten der Gruppe ‚nicht armutsgefährdet und *working class*'. Im Bereich Naturwissenschaften weisen die beiden Gruppen 2023 erneut praktisch identische Kompetenzwerte auf. Bei der Mathematikkompetenz zeigt sich erstmals ein nomineller Vorsprung der Gruppe ‚armutsgefährdet und *service class*' (nicht signifikant). Ein systematischer Trend im Zeitverlauf ist somit nicht festzustellen.

7.6 Zusammenfassung

Die in diesem Kapitel präsentierten Ergebnisse stehen im Einklang mit den vorangegangenen Erhebungen TIMSS 2007, 2011, 2015 und 2019 und lassen weiterhin sowohl im internationalen Vergleich als auch in Deutschland auf eine enge Kopplung zwischen dem sozioökonomischen Status einer Familie und dem Bildungserfolg ihrer Kinder schließen. Schüler:innen, die mindestens ein Elternteil mit tertiärem Bildungsabschluss haben, weisen sowohl eine signifikant höhere Mathematik- als auch Naturwissenschaftskompetenz auf als Schüler:innen mit Eltern ohne tertiären Bildungsabschluss. Während diese Disparität in nahezu jedem Teilnehmerstaat wiederzufinden ist, fällt die Differenz beider Gruppen in Deutschland (bis auf wenige Ausnahmen) im internationalen Vergleich signifikant kleiner aus.

Eine im internationalen Vergleich größere Leistungsdifferenz zeigt sich hingegen in den Naturwissenschaften zwischen Familien mit maximal 100 Büchern und Familien mit mehr als 100 Büchern im Haushalt, wobei der Wert signifikant größer als die mittlere Differenz der anderen Teilnehmerstaaten ist. Bei den Mathematikkompetenzen fällt die Differenz innerhalb Deutschlands geringer

aus als für Naturwissenschaftskompetenzen und ähnelt denen der teilnehmenden OECD- und EU-Staaten.

Auf nationaler Ebene wird die bestehende soziale Disparität durch den sozialen Gradienten deutlich, welcher sich seit der letzten Erhebung im Jahr 2019 leicht – jedoch nicht signifikant – erhöht hat. Damit gleicht sich eine vorangegangene leichte Reduzierung aus, sodass sich der Wert wieder dem Niveau der Ausgangslage von 2007 nähert.

Der Anteil armutsgefährdeter Familien bleibt ebenfalls relativ konstant und liegt mit 20.1 Prozent nur knapp unter dem Wert aus TIMSS 2019 mit 21.5 Prozent, wobei insbesondere Familien, in denen beide Elternteile im Ausland geboren sind, als armutsgefährdet gelten. Auch an dieser Stelle wird der Zusammenhang zwischen der sozioökonomischen Herkunft der Kinder und ihren schulischen Leistungen deutlich. Kinder aus nicht armutsgefährdeten Familien sind in ihren Leistungskompetenzen ihren Mitschüler:innen aus armutsgefährdeten Familien in der Mathematik etwa ein Jahr und in den Naturwissenschaften sogar zwei Jahre voraus. Die Differenz hat sich im Vergleich zu 2019 leicht vergrößert, der Unterschied ist jedoch nicht signifikant.

Da seit TIMSS 2023 im Erhebungsinstrument der EGP-Klassen nicht mehr nach den Berufen von Mutter und Vater, sondern der Elternteile A und B gefragt wird, finden sich bei der Häufigkeitsverteilung der EGP-Klassen im Trendvergleich für diesen Erhebungszeitpunkt erwartbare Unterschiede. Insbesondere in den Klassen II und III gab es im Vergleich zur vorherigen Erhebung jeweils einen substanziellen Anstieg, während der Anteil an Familien, die den EGP-Klassen V, VI und VII zugeordnet werden, entsprechend gesunken ist. Diese Veränderung ist darauf zurückzuführen, dass zuvor (wenn vorhanden) die EGP-Klasse des Vaters und seit 2023 die EGP-Klasse des Elternteils A betrachtet wird und die meisten Eltern an dieser Stelle die Mutter anzugeben scheinen. Der mittlere HISEI der Klassen ist jedoch trotz abweichender Verteilung relativ konstant geblieben. Hinsichtlich ihrer Kompetenzen sind Schüler:innen, deren Eltern der oberen Dienstklasse angehören, ihren Mitschüler:innen, die der Gruppe (Fach-) Arbeiter:innen angehören, in Mathematik etwa ein Lernjahr und in den Naturwissenschaften sogar mehr als anderthalb Lernjahre voraus. Die Leistungsdifferenzen zwischen den EGP-Klassen haben sich dabei in den letzten vier Jahren leicht vergrößert.

Werden die Leistungen von Schüler:innen sowohl unter Berücksichtigung ihrer Armutsgefährdung als auch der EGP-Klasse betrachtet, weisen nicht armutsgefährdete Kinder der *service class* sowohl in der Mathematik als auch in den Naturwissenschaften die höchsten Werte auf. Ein systematischer Trend ist im Zeitverlauf jedoch nicht zu erkennen.

Sowohl die vorherigen als auch aktuellen Ergebnisse verdeutlichen, dass nach wie vor die Notwendigkeit besteht, soziale Disparitäten einerseits im Hinblick auf den Bildungsverlauf sowie den Kompetenzerwerb zu berücksichtigen und andererseits Maßnahmen zu ergreifen, sozialer Disparität entgegenzuwirken.

Literatur

Bardone, L., & Guio, A.-C. (2005). In-work poverty. New commonly agreed indicators at EU level. *Statistics in Focus, 5*, 1–11.

Baumert, J., & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323–407). Leske + Budrich.

Berkemeyer, N., Bos, W., Hermstein, B., Abendroth, S., & Semper, I. (Hrsg.). (2017). *Chancenspiegel – eine Zwischenbilanz. Zur Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme seit 2002*. Bertelsmann Stiftung. https://www.ganztaegig-lernen.de/sites/default/files/Chancenspiegel_2017_WEB.pdf

Blossfeld, P. N. (2018). *Changes in inequality of educational opportunity. The long-term development in Germany*. Springer. https://doi.org/10.1007/978-3-658-22522-3

BMAS (Bundesministerium für Arbeit und Soziales). (2021). *Lebenslagen in Deutschland. Der sechste Armuts- und Reichtumsbericht der Bundesregierung*. https://www.armuts-und-reichtumsbericht.de/SharedDocs/Downloads/Berichte/sechster-armuts-reichtums-bericht.pdf?__blob=publicationFile&v=6

Bos, W., Stubbe, T. C., & Buddeberg, M. (2010). Gibt es eine armutsbedingte Bildungsbenachteiligung? Die Operationalisierung verschiedener Indikatoren der sozialen Herkunft in der empirischen Bildungsforschung. In D. H. Rost (Hrsg.), *Intelligenz, Hochbegabung, Vorschulerziehung, Bildungsbenachteiligung* (S. 165–208). Waxmann.

Bosch, G., & Kalina, T. (2007). Niedriglöhne in Deutschland – Zahlen, Fakten, Ursachen. In G. Bosch & C. Weinkopf (Hrsg.), *Arbeiten für wenig Geld. Niedriglohnbeschäftigung in Deutschland* (S. 20–105). Campus.

Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in western society.* Wiley.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (S. 183–198). Schwartz.

Butterwegge, C. (Hrsg.). (2000). *Kinderarmut in Deutschland. Ursachen, Erscheinungsformen und Gegenmaßnahmen.* Campus.

DESTATIS (Statistisches Bundesamt). (2024a). *Armutsgefährdungsquote (monetäre Armut) nach überwiegendem Erwerbsstatus im Vorjahr und Bildungsstand*. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Lebensbedingungen-Armutsgefaehrdung/Tabellen/armutsgef-quote-bildungsstand-mz-silc.html

DESTATIS (Statistisches Bundesamt). (2024b). *Armutsgefährdungsschwelle und Armutsgefährdung (monetäre Armut) im Zeitvergleich*. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Lebensbedingungen-Armutsgefaehrdung/Tabellen/armutsschwelle-gefaehrdung-mz-silc.html?nn=210056

DESTATIS (Statistisches Bundesamt). (2024c). *Einkommensverteilung (Nettoäquivalenzeinkommen) im Zeitvergleich*. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Lebensbedingungen-Armutsgefaehrdung/Tabellen/einkommensverteilung-mz-silc.html

DESTATIS (Statistisches Bundesamt). (2024d). *Weiterhin gut ein Fünftel der Bevölkerung von Armut oder sozialer Ungleichheit bedroht* (Pressemitteilung Nr. 147 vom 10. April 2024). Statistisches Bundesamt.

Erikson, R., Goldthorpe, H. J., & Portocarero, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *British Journal of Sociology, 30*(4), 415–441. https://doi.org/10.2307/589632

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. https://doi.org/10.1016/0049-089X(92)90017-B

Ganzeboom, H. B. G., & Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 international standard classification of occupations. *Social Science Research, 25*(3), 201–239. https://doi.org/10.1006/ssre.1996.0010

Ganzeboom, H. B. G., & Treiman, D. J. (2010). *Occupational status measures for the new International Standard Classification of Occupations ISCO-08; with a discussion of the new classification*. http://www.harryganzeboom.nl/isol/isol2010c2-ganzeboom.pdf

Geißler, R. (2014). Bildungsexpansion und Bildungschancen. In Bundeszentrale für politische Bildung (Hrsg.), *Sozialer Wandel in Deutschland*. http://www.bpb.de/izpb/198031/bildungsexpansion-und-bildungschancen? p=all

Laubstein, C., Holz, G., & Seddig, N. (2016). *Armutsfolgen für Kinder und Jugendliche. Erkenntnisse aus empirischen Studien in Deutschland.* Bertelsmann Stiftung. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/Studie_WB_Armutsfolgen_fuer_Kinder_und_Jugendliche_2016.pdf

Mang, J., Müller, K., Lewalter, D., Kastorff, T., Müller, M., Ziernwald, L., Tupac-Yupanqui, A., Heine, J.-H., & Köller, O. (2023). Herkunftsbezogene Ungleichheiten im Kompetenzerwerb. In D. Lewalter, J. Diedrich, F. Goldhammer, O. Köller & K. Reiss (Hrsg.), *PISA 2022. Analyse der Bildungsergebnisse in Deutschland* (S. 163–197). Waxmann. https://doi.org/10.31244/9783830998488

Menold, N., & Züll, C. (2010). *Codierung von Gründen der Verweigerung der Teilnahme an Interviews: ein Kategorienschema*. GESIS – Leibniz-Institut für Sozialwissenschaften. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-207210

Sachse, K. A., Jindra, C., Schumann, K., & Schipolowski, S. (2022). Soziale Disparitäten, In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (S. 152–180). Waxmann. http://doi.org/10.31244/9783830996064

Schroedter, J. H., Lechert, Y., & Lüttinger, P. (2006). *Die Umsetzung der Bildungsskala ISCED-1997 für die Volkszählung 1970, die Mikrozensus-Zusatzerhebung 1971 und die Mikrozensen 1976–2004 (Version 1). ZUMA-Methodenbericht 2006/08.* https://www.ssoar.info/ssoar/bitstream/handle/document/26239/ssoar-2006-schroedter_et_al-die_umsetzung_der_bildungsskala_isced-1997.pdf

Schwippert, K. (2019). Was wird aus den Büchern? Sozialer Hintergrund von Lernenden und Bildungsungleichheit aus Sicht der international vergleichenden Erziehungswissenschaft. *Journal for Educational Research Online, 11*(1), 92–117.

Senkbeil, M., Drossel, K., Eickelmann, B., & Vennemann, M. (2019). Soziale Herkunft und computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich. In B. Eickelmann, W. Bos, J. Gerick, F. Goldhammer, H. Schaumburg, K. Schwippert, M. Senkbeil & J. Vahrenhold (Hrsg.), *ICILS 2018 Deutschland. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking* (S. 301–333). Waxmann. https://doi.org/10.25656/01:18166

Stubbe, T. C., Kleinkorres, R., Krieg, M., Schaufelberg, R., & Schlitter, T. (2023). Soziale und migrationsbedingte Disparitäten in der Lesekompetenz von Viertklässlerinnen und Viertklässlern. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher & T. C. Stubbe (Hrsg.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (S. 151–177). Waxmann. https://doi.org/10.31244/9783830997009

Stubbe, T. C., & Krieg, M. (2023). Soziale Disparitäten im deutschen Bildungssystem. Theoretische Modelle und empirische Befunde. In N. McElvany, A. Grecu, R. Lorenz, M. Becker, C. Diegnath, H. Gaspard & F. Lauermann (Hrsg.), *Jahrbuch der Schulentwicklung 2023. 50 Jahre Schulentwicklung – Leitthemen der empirischen Bildungsforschung* (S. 14–44). Beltz Juventa.

Stubbe, T. C., Krieg, M., Beese, C., & Jusufi, D. (2020). Soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässlerinnen und Viertklässlern. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 263–289). Waxmann. https://doi.org/10.31244/9783830993193

Tophoven, S., Lietzmann, T., Reiter, S., & Wenzig, C. (2017). *Armutsmuster in Kindheit und Jugend. Längsschnittbetrachtungen von Kinderarmut.* https://www.bertelsmann-

stiftung.de/fileadmin/files/Projekte/Familie_und_Bildung/Studie_WB_Armutsmuster_in_Kindheit_und_Jugend_2017.pdf

UNESCO (United Nations Educational, Scientific and Cultural Organization). (2003). International Standard Classification of Education, ISCED 1997. In J. H. Hoffmeyer-Zlotnik & C. Wolf (Hrsg.), *Advances in cross-national comparison. A European working book for demographic and socio-economic variables* (S. 195–220). Plenum Press. https://doi.org/10.1007/978-1-4419-9186-7_10

Wendt, H., Kasper, D., Bos, W., Vennemann, M., & Goy, M. (2017). Wie viele Punkte auf der TIMSS-Metrik entsprechen einem Lernjahr? Leistungszuwächse in Mathematik und Naturwissenschaften am Ende der Grundschulzeit. In T. Eckert & B. Gniewosz (Hrsg.), *Bildungsgerechtigkeit* (S. 121–153). Springer. https://doi.org/10.1007/978-3-658-15003-7_8

Kapitel 8
Mathematische und naturwissenschaftliche Kompetenzen von Schüler:innen mit und ohne Migrationshintergrund

Knut Schwippert, Donieta Jusufi, Martje Lott und Tobias C. Stubbe

8.1 Einleitung[1]

Das Lernen von Schüler:innen hängt nicht nur von institutionellen Lehr- und Lerngelegenheiten ab, auch persönliche oder familiäre Merkmale und damit zusammenhängende gesellschaftliche Zuschreibungen können (Bildungs-)Biografien beeinflussen. Dabei rücken eben diese Merkmale bereits seit Jahrzehnten zunehmend in den Fokus bildungspolitischer Diskurse. Diese beziehen sich beispielsweise auf Fragen zum Abbau von bildungsbenachteiligenden Ausgangslagen von (Grund-)Schüler:innen bzw. auf die Stärkung der Bildungsgerechtigkeit sowie Chancengleichheit im deutschen Bildungssystem. Dabei hat sich national wie international gezeigt, dass Unterscheidungsmerkmale der Schüler:innen in Bezug auf Migrationserfahrungen in der Familie und den sozialen, ökonomischen, aber auch kulturellen Hintergrund der Familien einen Zusammenhang mit dem Lernen von Kindern aufweisen (Bonsen et al., 2008; Stubbe et al., 2020; Tarelli et al., 2012; Wendt et al., 2016; Wendt et al., 2020). In diesem Kontext treten die wechselseitigen Auswirkungen von Bildung, Migration und Integration zunehmend in den Vordergrund nationaler Bildungsdebatten. Dies geschieht nicht zuletzt, da Deutschland inzwischen zu den Staaten zählt, die international den höchsten Anteil an Personen mit Migrationshintergrund in der Schülerschaft aufweisen (Wendt et al., 2020). So konnte für das Jahr 2022 ein Rekord an Zuwanderungen verzeichnet werden, der laut dem Bildungsbericht 2024 der höchste seit dem Beginn der Aufzeichnungen in den 1950er-Jahren ist (Autor:innengruppe Bildungsberichterstattung, 2024).

Eine klare Definition der Gruppe von Menschen mit Migrationshintergrund ist jedoch herausfordernd, da diese vielfältige Herkunftsstaaten, (Einwanderungs-)Generationen, Sprachen, Religionen und zum Teil Fluchterfahrun-

1 Die Autor:innen danken Frau Melanie Ahrens (Universität Osnabrück) für ihre kritischen und konstruktiven Hinweise zur theoretischen Rahmung dieses Kapitels.

gen sowie Lebensformen und Familienzusammensetzungen in sich vereint und somit hinsichtlich zahlreicher Merkmale als heterogen anzusehen ist. Dennoch ist eine nähere Untersuchung von migrationsbezogenen Bildungsdisparitäten als äußerst relevant einzuschätzen, da die Sichtbarmachung der damit einhergehenden Benachteiligungen als Anlass genommen werden sollte, diese zu reduzieren. Der Zugang zu Bildung ist – unabhängig von Herkunft und sozialem Status – ein unantastbares und grundlegendes Grundrecht (Art. 26, GG). Dabei galt schon früher (Klafki, 2007) und gilt heute umso mehr, dass ein hohes Bildungsniveau kritische und konstruktive Selbstbestimmung in und für eine heterogene Gesellschaft gewährleistet und damit ein wichtiges Fundament für die zukünftige Übernahme von Verantwortung für sich, aber auch für die Gesellschaft bildet. Gleichzeitig bestimmt Klafki Bildung als Allgemeinbildung für alle und damit verbunden als Recht für alle Individuen und unabhängig von bestehenden Ungleichheiten in einer Gesellschaft (Verba et al., 1995). Entsprechend finden sich auch in aktuellen politischen Programmen und Zielsetzungen, wie beispielsweise in den globalen Zielen für nachhaltige Entwicklung (engl. *sustainable development goals*; SDGs) unter Punkt 4 (Deutsche UNESCO-Kommission, 2017), Forderungen nach einer chancengerechten Bildung für alle. Eine Realisierung dieser Forderungen hätte nicht nur positive Auswirkungen für die einzelnen Individuen, sondern auch für die gesamte Gesellschaft. Hinsichtlich der Teilhabe und Integration von Menschen mit Migrationshintergrund bedeutet dies, vor allem die Bildung und Erziehung in den Fokus zu stellen, da nur auf dieser Basis eine „dauerhafte schulische, soziale und berufliche Integration" (KMK 2002, S. 4) möglich ist. Die Integration von Kindern mit Migrationshintergrund kann als eine der wichtigsten Aufgaben sowohl für die Gesellschaft als auch für die Schule angesehen werden und stellt ein zentrales nationales bildungspolitisches Ziel zur Reduzierung von Bildungsungleichheiten und der Möglichkeit zur gesellschaftlichen Teilhabe dar (Autorengruppe Bildungsberichterstattung, 2016; KMK, 2002; Wendt et al., 2020).

Insbesondere vor dem Hintergrund, dass Deutschland vor extremen Herausforderungen wie Fachkräftemangel, zunehmenden Umweltproblemen oder gestiegener sozialer Ungleichheit steht, ist die Vermittlung von Bildung zur Bewältigung dieser Herausforderung von zentraler Bedeutung (UNESCO, 2018). Ein Staat, der vor solchen Herausforderungen steht, kann und darf es sich nicht erlauben, Kindern und Jugendlichen in den Schulen nicht die Möglichkeit zu geben, ihre Begabungsreserven voll auszuschöpfen.

In verschiedenen groß angelegten Schulleistungsstudien in der Grundschule für das Leseverständnis sowie für mathematische und naturwissenschaftliche Kompetenzen (*Internationale Grundschul-Lese-Untersuchung*; IGLU seit 2001, *Trends in International Mathematics and Science Study*; TIMSS seit 2007, IQB-Bildungstrend seit 2008/09) ist deutlich geworden, dass bereits eine einfache Gegenüberstellung des Lernerfolges mit Heterogenitätsmerkmalen der Schüler:innen eindeutige Befunde liefert. Zusammengefasst zeigt sich für Deutschland seit Jahrzehnten ein kaum verändertes Bild: Kinder mit Migrationshintergrund erreichen im Vergleich zu Kindern ohne Migrationshintergrund im Durchschnitt niedrigere Kompetenzstände in zentralen Kulturtechniken wie Lesen und Rechnen oder in den Naturwissenschaften (Henschel et al., 2022; Stubbe et al., 2023; Wendt et al., 2020). Dabei wird die Brisanz dieser Befunde immer wieder deutlich: Schon am Anfang der schulischen Laufbahn hätte ein substanzieller Anteil der getesteten Kinder mit Migrationshintergrund rund ein bis zwei Schuljahre aufzuholen, um an die erreichten Kompetenzen von Kindern ohne Migrationshin-

tergrund in der gleichen Klassenstufe anzuschließen (Wendt et al., 2020). In bisherigen Untersuchungen hat sich zudem gezeigt, dass der Migrationshintergrund nicht nur mit schulischen Leistungen zusammenhängt, sondern auch motivational-affektive Merkmale der Schüler:innen, wie beispielsweise die Einstellung und das Selbstkonzept, beeinflusst (de Boer & Merklinger, 2021). Im Rahmen von TIMSS 2019 konnte zudem zwar gezeigt werden, dass Kinder – unabhängig vom Migrationsstatus – mehrheitlich eine hohe positive Selbsteinschätzung angeben. Kinder, deren Eltern beide in Deutschland geboren wurden, geben dabei allerdings eine etwas positivere Selbsteinschätzung an (Wendt et al., 2020).

Damit im Zusammenhang steht, inwiefern sich Kinder mit Migrationshintergrund sozial eingebunden oder angenommen fühlen (als Facetten des globaleren Konzepts von *wellbeing*; Walther et al., 2020). So zeigen die Ergebnisse aus TIMSS 2019, dass Kinder mit Migrationshintergrund ungünstigere Ausprägungen bezüglich der motivationalen Merkmale aufweisen als Kinder ohne Migrationshintergrund. Auch hinsichtlich der sozialen Eingebundenheit fällt die Befundlage zuungunsten der Kinder mit Migrationshintergrund aus (Wendt et al., 2020). Eine konsistente Beobachtung der empirischen Befunde zu zuwanderungsbezogenen Disparitäten sowohl bei Veränderungen als auch bei gleichbleibender Befundlage ist eine zentrale Aufgabe und Pflicht einer internationalen Schulleistungsstudie wie TIMSS, ebenso wie auf diese problematische Situation hinzuweisen.

8.2 Weitere relevante Heterogenitätsmerkmale der Schüler:innen und ihrer Familien

Für den schulischen Lernerfolg gehen mit dem Migrationshintergrund weitere relevante Familienmerkmale einher, wie z. B. das soziale, ökonomische und kulturelle Kapital (Bourdieu, 1983; Bourdieu & Passerson, 1971). In diesem Kapitel erfolgt eine erste Einordnung dieser unterschiedlichen Heterogenitätsmerkmale in Bezug auf die Bedeutung des Migrationshintergrundes von Schüler:innen.

Insbesondere die Heterogenitätsmerkmale *sozioökonomischer Hintergrund, sonderpädagogischer Förderbedarf* und *Geschlecht* werden in der bisherigen Forschung häufig in Kombination mit dem Migrationshintergrund betrachtet und als mögliche Erklärungsfaktoren für beobachtbare Leistungsunterschiede herangezogen. In Kapitel 9 in diesem Band werden die Effekte des Zusammenwirkens dieser Merkmale gebündelt und im Trend analysiert.

Als grundlegende Basis für einen positiven Bildungsverlauf ist in diesem Zusammenhang die Sprache zu nennen, die von Esser (2006) als Schlüsselfaktor einer erfolgreichen Integration von Menschen mit Migrationshintergrund in die Aufnahmegesellschaft angesehen wird. Die Beherrschung der Sprache ist „eine Bedingung für die Weiterentwicklung des eigenen Wissens und der eigenen Fähigkeiten“ (Artelt et al., 2001, S. 11) und stellt somit die Basis u. a. für die Kommunikation im Unterricht und die schulische Kompetenzvermittlung dar (Gogolin, 1994). Sprachkenntnissen in der Unterrichts- oder auch der Alltagssprache kommt somit eine hohe Relevanz zu. Lernende, die den Erklärungen der Lehrkraft nicht folgen oder im Unterricht genutzte Texte und Aufgaben nicht verstehen können, sind somit auch nicht in der Lage, Bildungsangebote optimal zu nutzen (Wendt et al., 2020). Repräsentative Schulleistungsstudien verdeutlichen, dass der heimische Sprachgebrauch zur Klärung der Varianz – auch nach Kontrolle des sozioökonomischen Status und des Bildungshintergrundes der Eltern – beiträgt. Kinder, deren Eltern angeben, zu Hause immer Deutsch zu

sprechen, erzielen signifikant höhere mathematische und naturwissenschaftliche Leistungen als Kinder, die nie oder nur manchmal zu Hause Deutsch sprechen (Wendt et al., 2020).

Eng gekoppelt mit dem sozioökonomischen Status der Familie stellt die *Armutsgefährdung* ein weiteres Heterogenitätsmerkmal und einen Indikator der sozialen Herkunft der Familie dar (siehe Kapitel 7 in diesem Band). Mit dem jeweiligen Status sind unterschiedliche Privilegien und Benachteiligungen verbunden, die zu unterschiedlich verteilten Ressourcen und Möglichkeiten führen (Ditton & Krüsken, 2006). Im Rahmen von Grundschulstudien konnte gezeigt werden, dass der höchste Anteil der Kinder, die armutsgefährdet sind, in Familien aufwächst, in denen beide Elternteile im Ausland geboren wurden. Zudem erzielen Kinder, die in einem bildungsnahen Haushalt mit bildungsrelevanten Ressourcen aufwachsen und damit verbunden in der Familie einen höheren sozioökonomischen Status haben, bessere mathematische und naturwissenschaftliche Leistungen (Stubbe et al., 2023; Stubbe et al., 2020).

Aufbauend auf den vorangegangenen Erläuterungen werden in diesem Kapitel der Migrationshintergrund der Schüler:innen in TIMSS 2023, ihre mathematischen und naturwissenschaftlichen Kompetenzen sowie weitere soziale Hintergrundmerkmale vorgestellt. Dazu werden deskriptive sowie gegenüberstellende Analysen genutzt. Für weitergehende Betrachtungen sowie vertiefende Analysen zum Zusammenspiel der verschiedenen Merkmale der Schüler:innen im Trend siehe Kapitel 9 in diesem Band. Im folgenden Abschnitt wird die Erfassung des Migrationshintergrundes von Grundschüler:innen in TIMSS 2023 dargelegt.

8.3 Erfassung des Migrationshintergrundes in TIMSS 2023

Wie eingangs erläutert gibt es unterschiedliche Ansätze, Migrationshintergrund zu definieren. Ebenso sind die Gruppen von Kindern mit Migrationshintergrund nicht nur von diesem, sondern auch von weiteren potenziell bildungsbenachteiligenden Faktoren geprägt (Autorengruppe Bildungsberichterstattung, 2016). Daher ist zu prüfen, welche Indikatoren zur Erfassung des Migrationshintergrundes im Rahmen von TIMSS 2023 herangezogen werden können, welche weiteren Heterogenitätsmerkmale für eine differenzierte Betrachtung von Schüler:innen aus Familien mit Migrationserfahrung verfügbar sind und wie diese im Zusammenhang mit den Kompetenzen in Mathematik und den Naturwissenschaften stehen.

In TIMSS hat es sich seit 2007 (national) und seit 2015 (international) bewährt, zur Erfassung des Migrationshintergrundes nach dem Geburtsland der Eltern zu fragen. Hierbei wird auf Basis der Informationen über das Herkunftsland der Eltern zwischen Kindern unterschieden, bei denen (1) *beide Eltern in Deutschland geboren wurden*, bei denen (2) e*in Elternteil im Ausland geboren wurde* und schließlich bei denen (3) *beide Elternteile im Ausland geboren wurden*. Um einen Einblick in die Lebensbedingungen der Schüler:innen mit Migrationshintergrund geben und verschiedene Heterogenitätsmerkmale gemeinsam betrachten zu können, werden zusätzlich der *heimische Sprachgebrauch*, die *Armutsgefährdung* der Familie, der *Besuch eines Kindergartens* sowie der *Erwerbstätigkeitsstatus der Eltern* aufgegriffen.

Im Folgenden werden zunächst Befunde im internationalen Vergleich dargestellt, indem die Differenzen der mathematischen und naturwissenschaftlichen Kompetenzen der Schüler:innen mit und ohne Migrationshintergrund für die an

TIMSS beteiligten Staaten vergleichend gegenübergestellt werden. Da die Erfassung des Migrationshintergrundes nach dem Geburtsland der Eltern auf internationaler Ebene erst seit 2015 erfolgt, wird auf eine Betrachtung im internationalen Vergleich im Trend hinsichtlich der Zyklen 2007 und 2019 verzichtet (Wendt et al., 2020). Stattdessen erfolgt diese auf nationaler Ebene.

Vertiefend für Deutschland werden anschließend verschiedene Heterogenitätsmerkmale der Kinder gemeinsam in den Blick genommen sowie die erreichten Kompetenzstände der Kinder mit und ohne Migrationshintergrund in Mathematik und Naturwissenschaften im Trend seit 2007 betrachtet.

Im Anschluss wird dargestellt, wie sich das Selbstkonzept, die fachspezifischen Einstellungen und die soziale Eingebundenheit zwischen Kindern mit und ohne Migrationshintergrund unterscheiden. Abschließend werden die Befunde interpretierend zusammengefasst.

An dieser Stelle ist auf eine Besonderheit in der Darstellung der naturwissenschaftlichen Kompetenzen der Schüler:innen für dieses Kapitel hinzuweisen: National und international kommt es in seltenen Fällen vor, dass Schüler:innen zu wenige Testaufgaben erfolgreich bearbeitet haben, sodass eine zuverlässige Feststellung ihrer Kompetenzen mit großen Unsicherheiten verbunden ist. Der Anteil von Schüler:innen mit solchen nicht skalierbaren Leistungswerten ist für Deutschland insgesamt nicht auffällig (< 1 %). Jedoch zeigt sich bei differenzierter Betrachtung dieser Anteile im Zusammenhang mit dem Migrationshintergrund, dass in Deutschland für die Naturwissenschaften ein anteilig gesehen deutlich höherer Anteil zu beobachten ist (ca. 4 %). Bei den nationalen Analysen wurden diese Schüler:innen ausgeschlossen, da es sonst zu einer möglichen systematischen Unterschätzung der Kompetenzwerte für diese Gruppen kommen könnte. Somit weicht die Datengrundlage für die in TIMSS 2023 ermittelten naturwissenschaftlichen Kompetenzen in den nationalen Abbildungen und Tabellen leicht von der internationalen Datengrundlage ab.

8.4 Mathematische und naturwissenschaftliche Kompetenzen der Schüler:innen nach Migrationshintergrund im internationalen Vergleich

Wie bereits in den früheren Zyklen von TIMSS dargestellt, zeigt sich auch in 2023, dass der Migrationshintergrund im Zusammenhang mit den mathematischen und naturwissenschaftlichen Kompetenzen der Grundschulkinder steht.

Abbildung 8.1 zeigt die Testleistungen der Viertklässler:innen für Mathematik differenziert nach dem Migrationshintergrund der Eltern. Dem linken Tabellenteil sind für die teilnehmenden Staaten und Benchmark-Teilnehmer die Testleistungen der Schüler:innen zu entnehmen, bei denen kein Elternteil, ein Elternteil oder beide Elternteile im Ausland geboren wurden. Der rechte Tabellenteil bildet den Leistungsvorsprung der Kinder ab, deren Eltern beide im Inland geboren wurden. Verglichen wird dabei mit den Gruppen *ein Elternteil im Ausland geboren* bzw. *beide Elternteile im Ausland geboren*. In der Grafik sind die letztgenannten Differenzen visualisiert, wobei signifikante Differenzen grün eingefärbt sind. Balken nach rechts kennzeichnen höhere Kompetenzen für Schüler:innen, deren Eltern beide im Inland geboren wurden. Balken nach links zeigen entsprechend, dass Viertklässler:innen, deren Eltern beide im Ausland geboren wurden, höhere Kompetenzen aufweisen.

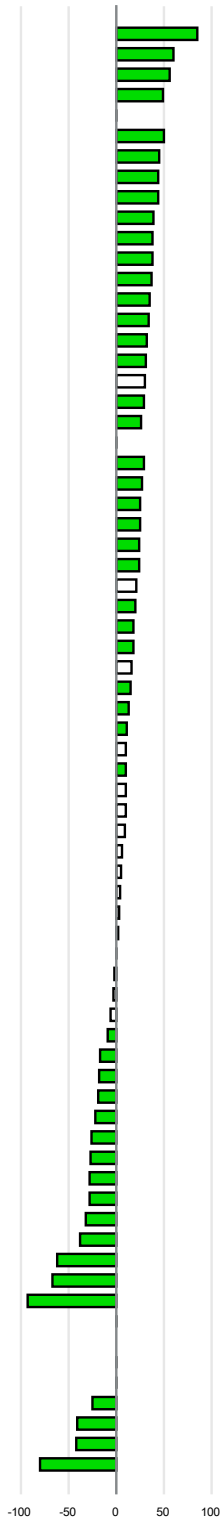
Abbildung 8.1 ist zu entnehmen, dass in Deutschland insgesamt rund 40 Prozent der Schüler:innen in Familien leben, in denen mindestens ein Elternteil im Ausland geboren wurde. Kinder, deren Eltern beide in Deutschland geboren wurden, weisen einen Kompetenzvorsprung von 11 Punkten (verglichen mit der Gruppe *ein Elternteil im Ausland geboren*) bzw. 38 Punkten (verglichen mit der Gruppe *beide Elternteile im Ausland geboren*) auf. Letzteres kann wie ein potenzieller Lernfortschritt von fast einem Jahr interpretiert werden (Wendt et al., 2017).

Im internationalen Vergleich zeigt sich, dass vier Staaten verglichen mit Deutschland eine signifikant größere Differenz in den Mathematikkompetenzen zwischen Kindern mit Eltern ohne Migrationshintergrund und Eltern, die beide im Ausland geboren wurden, aufweisen. Dies sind die Republik Korea (Südkorea), Brasilien, Finnland und Taiwan. Im Vergleich mit dem internationalen Mittelwert (10 Punkte Differenz) sowie der Vergleichsgruppe der OECD mit 20 Punkten weist Deutschland mit 38 Punkten eine signifikant größere Differenz in Mathematik auf. Kein statistisch signifikanter Unterschied kann zwischen Deutschland, der Vergleichsgruppe der EU (26 Punkte) und weiteren dreizehn teilnehmenden Staaten (mit Differenzen von 50 bis 29 Punkten) festgestellt werden. Wie bereits in früheren TIMSS-Untersuchungen wiesen auch in 2023 wieder einige Staaten (n = 14) höhere Kompetenzen von Grundschulkindern auf, deren Eltern beide im Ausland geboren wurden, verglichen mit Kindern, deren Eltern beide im Teilnehmerstaat geboren wurden. Die geringste signifikante Differenz für Mathematik weist die USA mit 9 Punkten auf, die größte Differenz mit 93 Punkten ist in Kuwait vorzufinden.

Nach der Betrachtung der zuwanderungsbedingten Disparitäten in den mathematischen Kompetenzen wird dies im Folgenden für den naturwissenschaftlichen Sachunterricht aufgegriffen. Mit Blick auf Abbildung 8.2 (die entsprechend wie Abbildung 8.1 aufgebaut ist) zeigt sich, wie auch bereits in TIMSS 2019, für die Naturwissenschaftskompetenz ein etwas anderes Bild als bei der Mathematikkompetenz. Hier ist der Kompetenzvorsprung der Gruppe *kein Elternteil im Ausland geboren* vor der Gruppe *beide Elternteile im Ausland geboren* in Deutschland mit 73 Punkten deutlich stärker ausgeprägt und entspricht damit einem Lernunterschied von fast zwei Schuljahren. Lediglich Finnland weist mit 96 Punkten eine signifikant größere Differenz auf. Disparitäten in den naturwissenschaftlichen Kompetenzen, die sich statistisch nicht signifikant von Deutschland unterscheiden, weisen drei Staaten (Republik Korea [Südkorea], Brasilien und Schweden) auf. Hingegen können für 53 Staaten signifikant geringere Kompetenzunterschiede in den Naturwissenschaften zwischen Kindern, deren Eltern beide im Teilnehmerstaat geboren wurden, und den Kindern, deren Eltern beide im Ausland geboren wurden, festgestellt werden. Von diesen Staaten weisen wiederum elf Staaten einen positiven Kompetenzvorsprung von Kindern auf, deren Eltern beide im Ausland geboren wurden. Hier liegen die signifikanten Differenzen zwischen 9 Punkten (Australien) und 101 Punkten (Kuwait).

Abbildung 8.1: Testleistungen der Schüler:innen in Mathematik nach Migrationshintergrund der Eltern im internationalen Vergleich.

Teilnehmer	Geburtsland im Ausland: kein Elternteil %	M_K	(SE)	ein Elternteil %	M_E	(SE)	beide Elternteile %	M_B	(SE)	Differenz A M_K-M_E	(SE)		Differenz A M_K-M_B	(SE)	
Republik Korea (Südkorea)	93.9	599	(2.4)	5.0	579	(8.9)	1.1	514	(29.1)	20	(9.1)	*	85	(28.8)	**
23 *Brasilien*	91.4	416	(2.1)	7.3	364	(5.2)	1.4	357	(9.6)	52	(5.0)	**	60	(8.9)	**
Finnland	84.1	539	(1.8)	7.8	531	(4.9)	8.1	483	(6.4)	9	(4.6)		56	(6.0)	**
Taiwan	89.0	619	(1.4)	7.3	593	(3.7)	3.8	570	(6.0)	26	(3.6)	**	49	(5.7)	**
123 *Türkei (5. Jgst.)*	94.1	559	(4.0)	4.7	523	(10.0)	1.2	509	(17.7)	36	(10.2)	**	50	(17.5)	**
2 *Litauen*	82.7	572	(2.4)	6.5	546	(6.9)	10.8	527	(4.4)	26	(6.9)	**	45	(4.7)	**
2 *Schweden*	69.1	544	(2.2)	12.2	536	(3.5)	18.7	500	(6.5)	8	(4.1)	*	44	(6.6)	**
2 *Bosnien und Herzegowina*	87.3	454	(3.2)	8.0	449	(7.0)	4.7	411	(9.3)	6	(7.1)		44	(8.7)	**
2 *Ungarn*	89.2	529	(3.2)	6.3	514	(9.9)	4.5	489	(8.6)	15	(9.8)		39	(8.7)	**
Marokko	93.5	402	(4.2)	4.9	374	(5.9)	1.6	364	(12.3)	29	(5.6)	**	38	(12.4)	**
Deutschland	60.0	540	(1.9)	14.3	529	(3.7)	25.6	502	(2.9)	11	(3.9)	**	38	(3.1)	**
2 *Armenien*	91.4	520	(2.6)	6.0	500	(6.4)	2.6	484	(8.8)	20	(6.2)	**	37	(8.3)	**
Lettland	82.9	544	(2.7)	9.3	522	(6.9)	7.7	509	(6.1)	22	(7.0)	**	35	(6.3)	**
23 *Dänemark*	76.3	533	(2.2)	12.1	520	(4.7)	11.6	499	(4.9)	13	(5.1)	*	34	(4.7)	**
23 *Belgien (Fläm. Gem.)*	68.4	537	(2.0)	12.6	518	(5.1)	19.0	505	(5.0)	19	(5.0)	**	32	(5.1)	**
2 *Frankreich*	69.6	498	(3.1)	15.7	481	(5.4)	14.7	467	(6.3)	17	(5.2)	**	31	(6.3)	**
23 *Rumänien*	91.6	556	(4.5)	7.1	514	(11.0)	1.3	525	(28.0)	41	(10.5)	**	30	(26.9)	
Iran	89.6	427	(3.8)	5.6	398	(8.0)	4.8	398	(6.8)	30	(8.0)	**	29	(6.7)	**
VG EU	**77.6**	**533**	**(2.6)**	**10.6**	**518**	**(6.5)**	**11.8**	**507**	**(8.9)**	**15**	**(6.3)**	*	**26**	**(8.8)**	**
12 *Norwegen (5. Jgst.)*	67.4	539	(2.2)	16.3	534	(3.5)	16.3	510	(3.7)	5	(3.6)		29	(4.0)	**
2 *Italien*	76.9	525	(2.9)	11.3	506	(4.6)	11.8	499	(5.4)	20	(4.4)	**	27	(5.4)	**
2 *Slowenien*	79.9	521	(1.9)	9.9	509	(4.1)	10.1	495	(4.5)	11	(3.9)	**	25	(4.3)	**
2 *Spanien*	73.2	508	(2.0)	10.6	498	(2.9)	16.1	483	(2.7)	9	(2.8)	**	25	(2.9)	**
Nordmazedonien	86.5	488	(2.9)	9.3	465	(7.0)	4.2	464	(6.9)	22	(6.3)	**	24	(6.9)	**
3 *Niederlande*	76.9	545	(1.9)	13.3	523	(6.0)	9.8	522	(5.3)	23	(5.7)	**	24	(5.6)	**
23 *Albanien*	89.8	521	(5.0)	7.5	488	(10.3)	2.7	500	(11.6)	33	(10.5)	**	21	(11.5)	
VG OECD	**76.6**	**533**	**(2.6)**	**10.6**	**521**	**(6.1)**	**12.8**	**513**	**(8.8)**	**12**	**(6.2)**		**20**	**(8.8)**	*
2 *Tschechien*	85.1	536	(1.9)	7.1	524	(5.1)	7.8	518	(5.1)	12	(4.6)	**	18	(4.7)	**
Jordanien	71.4	439	(4.3)	20.1	424	(5.9)	8.5	421	(6.4)	15	(5.7)	**	18	(7.2)	*
Slowakei	91.0	520	(3.5)	5.7	504	(7.4)	3.3	504	(9.7)	16	(8.1)	*	16	(9.5)	
2 *Belgien (Frz. Gem.)*	58.5	500	(2.4)	17.6	495	(3.9)	23.9	484	(3.5)	4	(3.8)		15	(3.6)	**
2 *Kasachstan*	88.1	492	(3.2)	7.3	490	(6.1)	4.6	478	(6.3)	2	(5.5)		13	(6.5)	*
2 *Georgien*	80.9	507	(3.3)	6.0	484	(7.2)	13.1	496	(4.7)	24	(7.6)	**	11	(5.2)	*
Internationaler Mittelwert	**77.2**	**510**	**(3.0)**	**11.1**	**496**	**(6.4)**	**11.6**	**500**	**(9.1)**	**15**	**(6.3)**	*	**10**	**(9.0)**	
Irland	68.2	551	(3.0)	12.9	556	(4.9)	18.9	541	(4.8)	-5	(4.8)		10	(4.8)	*
2 *Serbien*	84.4	527	(3.6)	12.5	533	(4.6)	3.1	516	(15.2)	-6	(5.2)		10	(15.6)	
Portugal	75.8	521	(2.8)	12.7	523	(4.8)	11.4	511	(5.9)	-2	(4.7)		10	(5.5)	
2 *Zypern*	58.6	520	(2.3)	21.1	522	(3.1)	20.2	511	(5.5)	-2	(3.2)		9	(5.4)	
Bulgarien	93.9	534	(3.5)	4.2	512	(12.6)	2.0	528	(15.6)	22	(11.3)	*	6	(16.0)	
1 *Südafrika (5. Jgst.)*	81.4	379	(3.8)	14.1	332	(6.7)	4.5	374	(9.2)	47	(6.0)	**	5	(8.1)	
23 *Chile*	84.2	452	(2.6)	8.2	430	(5.3)	7.6	448	(7.0)	22	(5.5)	**	4	(7.2)	
2 *Montenegro*	72.6	483	(1.9)	18.9	485	(3.2)	8.5	480	(5.1)	-2	(3.5)		3	(4.9)	
2 *Polen*	95.9	549	(1.9)	2.0	516	(9.3)	2.1	547	(10.6)	34	(9.2)	**	2	(10.3)	
Aserbaidschan	84.9	504	(3.0)	7.1	479	(7.4)	8.0	504	(5.3)	25	(7.2)	**	0	(5.3)	
2 *Usbekistan*	86.2	448	(3.2)	8.7	442	(5.3)	5.2	450	(5.8)	6	(5.9)		-2	(5.7)	
Japan	96.6	594	(2.4)	2.2	585	(9.0)	1.1	597	(13.6)	9	(8.7)		-3	(13.6)	
2 *Kosovo*	92.6	457	(2.9)	5.7	451	(7.2)	1.7	464	(9.0)	7	(7.1)		-6	(8.6)	
23 *USA*	72.7	529	(2.7)	13.4	512	(4.1)	13.9	538	(4.8)	16	(3.3)	**	-9	(4.4)	*
23 *Kanada*	58.6	506	(1.9)	12.3	511	(6.0)	29.2	523	(2.6)	-5	(6.3)		-17	(3.1)	**
123 *England*	64.1	556	(3.0)	15.7	566	(6.0)	20.3	574	(5.4)	-10	(6.2)		-18	(5.3)	**
3 *Hongkong*	68.7	600	(3.6)	16.9	597	(6.2)	14.4	620	(9.5)	4	(5.0)		-19	(8.3)	*
Macau	38.0	577	(1.9)	28.3	582	(1.7)	33.7	599	(1.8)	-6	(2.5)	*	-22	(2.9)	**
123 *Neuseeland*	57.8	494	(2.8)	17.8	512	(5.5)	24.4	520	(4.5)	-18	(5.8)	**	-26	(5.2)	**
2 *Saudi-Arabien*	76.6	430	(3.7)	16.3	412	(5.3)	7.1	457	(9.0)	18	(4.8)	**	-27	(8.2)	**
Bahrain	66.2	467	(3.9)	13.6	437	(4.5)	20.2	495	(7.2)	31	(4.2)	**	-28	(7.2)	**
Australien	58.6	528	(2.7)	17.2	529	(3.6)	24.2	556	(3.8)	-1	(4.1)		-28	(4.6)	**
Oman	79.4	433	(4.0)	12.5	402	(4.2)	8.0	465	(3.5)	31	(4.7)	**	-32	(4.4)	**
2 *Singapur*	56.2	610	(2.3)	20.8	619	(2.8)	23.0	647	(2.5)	-9	(2.6)	**	-38	(2.5)	**
Katar	58.4	458	(2.9)	11.6	430	(5.3)	30.0	520	(3.6)	28	(5.3)	**	-62	(4.0)	**
Vereinigte Arabische Emirate	54.6	487	(1.5)	9.6	460	(2.2)	35.8	554	(1.3)	27	(2.1)	**	-67	(1.8)	**
Kuwait	65.1	367	(4.5)	14.5	374	(11.1)	20.4	460	(6.5)	-7	(10.9)		-93	(7.7)	**
Benchmark-Teilnehmer															
23 *Québec, Kanada*	62.1	521	(2.8)	11.3	516	(3.8)	26.5	522	(3.5)	6	(4.2)		0	(4.5)	
2 *Ontario, Kanada*	54.9	503	(2.8)	12.2	505	(3.7)	33.0	528	(4.0)	-3	(4.4)		-25	(4.5)	**
Schardscha, VAE	52.6	499	(4.2)	8.4	477	(4.1)	39.0	540	(2.5)	22	(5.5)	**	-41	(4.0)	**
Dubai, VAE	42.3	542	(2.4)	6.8	536	(4.2)	50.9	584	(1.7)	6	(4.7)		-42	(2.5)	**
Abu Dhabi, VAE	59.7	452	(1.7)	12.3	419	(3.2)	28.0	532	(2.1)	33	(3.2)	**	-80	(2.6)	**

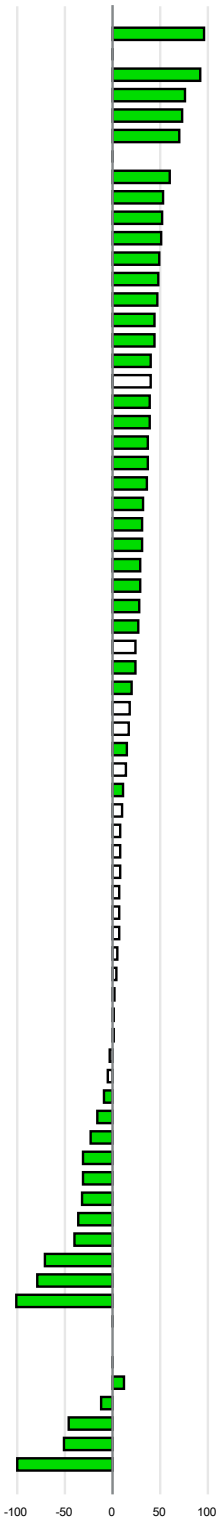


Keine statistisch signifikanten Unterschiede zwischen *kein Elternteil* und *beide Elternteile* ($p > .05$)
Statistisch signifikante Unterschiede zwischen *kein Elternteil* und *beide Elternteile* ($p \leq .05$)
▲ = Differenzwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
* = Statistisch signifikante Unterschiede ($p \leq .05$)
** = Statistisch signifikante Unterschiede ($p \leq .01$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

Abbildung 8.2: Testleistungen der Schüler:innen in Naturwissenschaften nach Migrationshintergrund der Eltern im internationalen Vergleich

Teilnehmer	Geburtsland im Ausland: kein Elternteil %	M_K	(SE)	ein Elternteil %	M_E	(SE)	beide Elternteile %	M_B	(SE)	Differenz M_K-M_E	(SE)		Differenz M_K-M_B	(SE)		A
Finnland	84.1	556	(2.1)	7.8	541	(4.7)	8.1	460	(6.4)	15	(4.0)	**	96	(5.7)	**	▲
Republik Korea (Südkorea)	93.9	588	(2.3)	5.0	563	(6.8)	1.1	496	(32.2)	25	(7.2)	**	92	(31.9)	**	■
23 *Brasilien*	91.4	446	(2.2)	7.3	377	(4.9)	1.4	370	(9.6)	68	(4.6)	**	76	(9.3)	**	■
Deutschland	60.0	543	(2.8)	14.3	520	(4.4)	25.6	470	(3.1)	23	(4.1)	**	73	(3.5)	**	■
2 *Schweden*	69.1	553	(2.3)	12.2	539	(4.0)	18.7	483	(6.5)	14	(4.1)	**	70	(6.4)	**	■
23 *Dänemark*	76.3	535	(2.3)	12.1	522	(5.1)	11.6	475	(4.7)	14	(4.8)	**	60	(4.6)	**	▼
12 *Norwegen (5. Jgst.)*	67.4	544	(2.2)	16.3	535	(4.0)	16.3	491	(5.0)	10	(3.7)	**	53	(4.5)	**	▼
2 *Slowenien*	79.9	536	(2.4)	9.9	521	(4.5)	10.1	484	(4.8)	14	(4.2)	**	52	(4.8)	**	▼
23 *Belgien (Fläm. Gem.)*	68.4	508	(2.4)	12.6	483	(6.0)	19.0	457	(5.0)	25	(5.7)	**	51	(5.2)	**	▼
Marokko	93.5	402	(4.7)	4.9	371	(10.0)	1.6	353	(11.4)	31	(9.3)	**	49	(11.1)	**	▼
2 *Bosnien und Herzegowina*	87.3	456	(3.6)	8.0	450	(7.6)	4.7	409	(12.3)	7	(7.9)		48	(12.0)	**	▼
Taiwan	89.0	586	(1.4)	7.3	555	(4.3)	3.8	539	(6.6)	30	(4.3)	**	47	(6.7)	**	▼
2 *Litauen*	82.7	548	(2.3)	6.5	525	(7.1)	10.8	504	(5.4)	23	(6.9)	**	44	(5.8)	**	▼
2 *Ungarn*	89.2	531	(3.3)	6.3	513	(9.0)	4.5	487	(8.0)	18	(9.8)		44	(8.6)	**	▼
123 *Türkei (5. Jgst.)*	94.1	574	(3.3)	4.7	546	(9.5)	1.2	534	(16.4)	28	(9.4)	**	40	(15.8)	*	▼
23 *Rumänien*	91.6	539	(4.5)	7.1	506	(8.2)	1.3	499	(21.3)	33	(8.0)	**	40	(20.7)		▼
3 *Niederlande*	76.9	527	(2.7)	13.3	506	(6.3)	9.8	487	(7.4)	20	(5.6)	**	39	(6.7)	**	▼
VG EU	**77.6**	**529**	**(2.9)**	**10.6**	**513**	**(6.7)**	**11.8**	**490**	**(8.0)**	**15**	**(6.4)**	*	**39**	**(7.9)**	**	▼
2 *Armenien*	91.4	464	(2.5)	6.0	446	(5.5)	2.6	428	(9.6)	19	(5.1)	**	37	(9.3)	**	▼
Iran	89.6	442	(4.1)	5.6	400	(8.6)	4.8	405	(7.5)	42	(8.8)	**	37	(8.2)	**	▼
2 *Frankreich*	69.6	503	(2.7)	15.7	484	(6.2)	14.7	467	(6.7)	19	(5.6)	**	36	(6.0)	**	▼
VG OECD	**76.6**	**535**	**(2.6)**	**10.6**	**523**	**(6.0)**	**12.8**	**504**	**(9.0)**	**13**	**(5.9)**	*	**32**	**(8.9)**	**	▼
Nordmazedonien	86.5	455	(3.4)	9.3	429	(6.2)	4.2	424	(8.1)	26	(6.8)	**	31	(6.9)	**	▼
2 *Tschechien*	85.1	533	(2.3)	7.1	520	(4.7)	7.8	502	(5.1)	12	(4.4)	**	31	(4.8)	**	▼
Lettland	82.9	536	(3.2)	9.3	513	(7.1)	7.7	506	(5.6)	23	(6.9)	**	29	(6.5)	**	▼
2 *Belgien (Frz. Gem.)*	58.5	495	(2.4)	17.6	486	(4.1)	23.9	466	(5.0)	9	(3.7)	*	29	(5.1)	**	▼
2 *Polen*	95.9	553	(2.1)	2.0	531	(8.9)	2.1	525	(9.1)	23	(8.8)	*	28	(8.7)	**	▼
2 *Italien*	76.9	521	(2.7)	11.3	509	(4.3)	11.8	494	(4.6)	12	(4.2)	**	27	(5.0)	**	▼
Bulgarien	93.9	534	(4.6)	4.2	504	(15.6)	2.0	509	(14.5)	30	(13.5)	*	24	(14.8)		▼
2 *Spanien*	73.2	513	(2.0)	10.6	506	(3.1)	16.1	489	(2.6)	7	(2.8)	*	24	(2.9)	**	▼
2 *Zypern*	58.6	494	(2.8)	21.1	495	(3.9)	20.2	474	(6.9)	-1	(4.0)		20	(6.7)	**	▼
Slowakei	91.0	527	(3.9)	5.7	508	(7.8)	3.3	508	(9.3)	19	(8.7)	*	18	(9.5)		▼
Internationaler Mittelwert	**77.2**	**503**	**(3.1)**	**11.1**	**486**	**(6.5)**	**11.6**	**486**	**(9.3)**	**16**	**(6.3)**	**	**17**	**(9.2)**		▼
Irland	68.2	537	(3.3)	12.9	547	(5.4)	18.9	522	(5.6)	-10	(5.0)	*	15	(5.9)	*	▼
23 *Albanien*	89.8	502	(4.5)	7.5	468	(8.8)	2.7	488	(13.9)	34	(8.9)	**	14	(14.0)		▼
2 *Georgien*	80.9	474	(3.7)	6.0	453	(6.8)	13.1	463	(5.3)	21	(7.1)	**	11	(5.7)	*	▼
2 *Kasachstan*	88.1	469	(3.2)	7.3	481	(6.5)	4.6	459	(6.7)	-12	(6.0)	*	10	(6.6)		▼
2 *Montenegro*	72.6	467	(2.2)	18.9	471	(3.0)	8.5	460	(4.6)	-3	(3.7)		8	(4.8)		▼
Jordanien	71.4	429	(4.4)	20.1	415	(6.5)	8.5	421	(9.3)	14	(6.2)	*	8	(9.6)		▼
2 *Serbien*	84.4	513	(3.5)	12.5	520	(4.5)	3.1	506	(14.6)	-6	(5.0)		8	(14.8)		▼
Portugal	75.8	514	(2.5)	12.7	514	(4.1)	11.4	507	(4.8)	0	(4.1)		7	(4.7)		▼
1 *Südafrika (5. Jgst.)*	81.4	329	(5.1)	14.1	264	(8.6)	4.5	322	(13.4)	65	(7.5)	**	7	(12.2)		▼
23 *USA*	72.7	547	(2.6)	13.4	527	(3.4)	13.9	540	(4.3)	20	(3.1)	**	7	(3.9)		▼
Japan	96.6	558	(2.5)	2.2	544	(9.9)	1.1	553	(12.3)	14	(9.1)		5	(12.1)		▼
123 *England*	64.1	563	(2.8)	15.7	572	(5.1)	20.3	559	(4.6)	-9	(4.9)		4	(4.7)		▼
Aserbaidschan	84.9	429	(2.9)	7.1	405	(5.7)	8.0	428	(6.2)	24	(5.8)	**	2	(5.3)		▼
23 *Chile*	84.2	489	(2.6)	8.2	464	(5.1)	7.6	488	(6.1)	26	(5.3)	**	1	(6.1)		▼
2 *Kosovo*	92.6	409	(3.4)	5.7	404	(5.8)	1.7	408	(9.3)	5	(5.5)		1	(10.5)		▼
2 *Usbekistan*	86.2	416	(3.2)	8.7	402	(6.7)	5.2	418	(7.9)	14	(6.4)	*	-3	(7.5)		▼
23 *Kanada*	58.6	526	(2.0)	12.3	531	(5.5)	29.2	531	(2.3)	-5	(5.4)		-5	(2.6)		▼
Australien	58.6	556	(2.3)	17.2	558	(2.8)	24.2	565	(3.4)	-2	(3.4)		-9	(3.6)	**	▼
123 *Neuseeland*	57.8	523	(2.8)	17.8	541	(4.7)	24.4	539	(4.4)	-19	(4.8)	**	-16	(4.4)	**	▼
Macau	38.0	532	(2.2)	28.3	533	(1.9)	33.7	555	(1.7)	-1	(2.4)		-23	(2.6)	**	▼
2 *Saudi-Arabien*	76.6	443	(3.4)	16.3	414	(5.7)	7.1	474	(9.6)	29	(5.1)	**	-31	(8.9)	**	▼
3 *Hongkong*	68.7	550	(3.3)	16.9	552	(6.1)	14.4	581	(11.3)	-3	(5.0)		-31	(10.4)	**	▼
2 *Singapur*	56.2	604	(2.3)	20.8	612	(2.7)	23.0	636	(2.7)	-8	(2.4)	**	-32	(2.5)	**	▼
Bahrain	66.2	482	(3.7)	13.6	448	(4.3)	20.2	518	(6.3)	34	(4.5)	**	-36	(6.5)	**	▼
Oman	79.4	446	(3.9)	12.5	407	(5.6)	8.0	486	(4.6)	39	(5.9)	**	-40	(5.0)	**	▼
Katar	58.4	466	(3.2)	11.6	432	(5.5)	30.0	538	(3.0)	34	(5.5)	**	-71	(3.4)	**	▼
Vereinigte Arabische Emirate	54.6	483	(2.1)	9.6	451	(2.7)	35.8	562	(1.6)	32	(2.2)	**	-79	(1.9)	**	▼
Kuwait	65.1	360	(5.7)	14.5	366	(12.7)	20.4	462	(9.3)	-6	(12.5)		-101	(10.7)	**	▼
Benchmark-Teilnehmer																
23 *Québec, Kanada*	62.1	518	(3.1)	11.3	508	(4.0)	26.5	506	(2.9)	11	(4.3)	*	12	(4.2)	**	▼
2 *Ontario, Kanada*	54.9	528	(3.1)	12.2	534	(4.6)	33.0	540	(3.3)	-6	(4.2)		-12	(3.8)	**	▼
Dubai, VAE	42.3	547	(2.2)	6.8	539	(4.6)	50.9	593	(2.2)	8	(5.0)		-46	(2.6)	**	▼
Schardscha, VAE	52.6	496	(4.9)	8.4	473	(5.4)	39.0	548	(2.4)	23	(5.2)	**	-51	(4.8)	**	▼
Abu Dhabi, VAE	59.7	438	(2.3)	12.3	398	(4.2)	28.0	538	(2.7)	39	(3.7)	**	-100	(2.9)	**	▼



Keine statistisch signifikanten Unterschiede zwischen *kein Elternteil* und *beide Elternteile* ($p > .05$)
Statistisch signifikante Unterschiede zwischen *kein Elternteil* und *beide Elternteile* ($p \leq .05$)
▲ = Differenzwert im Vergleich zu Deutschland statistisch signifikant höher ($p \leq .05$)
▼ = Differenzwert im Vergleich zu Deutschland statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zum Differenzwert von Deutschland nicht statistisch signifikant ($p > .05$)
* = Statistisch signifikante Unterschiede ($p \leq .05$)
** = Statistisch signifikante Unterschiede ($p \leq .01$)
Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

© TIMSS 2023

8.5 Lebensbedingungen von Schüler:innen mit und ohne Migrationshintergrund in Deutschland

Wie bereits in TIMSS 2019 dargelegt, unterliegt die Zusammensetzung der Schülerschaft einem ständigen Wandel, der durch unterschiedliche Aspekte wie z. B. demografischer Wandel, Einbindung in die EU und Migration bedingt ist (Kasper et al., 2020). Diese Veränderung lässt sich mit unterschiedlichen Akzentuierungen und unterschiedlichen Hintergrundmerkmalen auch in TIMSS 2023 beobachten (siehe Kapitel 9 in diesem Band). Für Schüler:innen mit und ohne Migrationshintergrund lässt sich anhand von Tabelle 8.1 feststellen, dass die Gruppe der Kinder, deren Eltern beide im Ausland geboren wurden, mit 25.6 Prozent um 3.7 Prozentpunkte größer ausfällt als in TIMSS 2019. Der Anteil von Kindern, deren Elternteile beide in Deutschland geboren wurden, beträgt wiederum 60.0 Prozent und ist damit 3.5 Prozentpunkte niedriger als in TIMSS 2019. Der Anteil von Kindern, bei denen ein Elternteil im Ausland geboren wurde, ist in TIMSS 2023 mit 14.3 Prozent (TIMSS 2019: 14.4 %) praktisch unverändert.

Für den heimischen Sprachgebrauch kann für TIMSS 2023 festgestellt werden, dass in Familien, in denen beide Elternteile im Ausland geboren wurden, rund ein Drittel der Eltern *immer* oder *fast immer* Deutsch spricht, in Familien mit einem im Ausland geborenen Elternteil etwa zwei Drittel. In Familien, bei denen beide Elternteile in Deutschland geboren wurden, sprechen 95.5 Prozent der Eltern *immer* oder *fast immer* Deutsch. Diese Anteile fallen für alle betrachteten Gruppen etwas niedriger aus als in TIMSS 2019 (1.5 – 6.9 Prozentpunkte).

Bei der Betrachtung des Kindergartenbesuchs zeigen sich nur marginale Veränderungen zu 2019. Insgesamt gaben in TIMSS 2023 64.1 Prozent der Eltern, die beide im Ausland geboren wurden, an, dass ihr Kind mindestens drei Jahre den Kindergarten besucht hat. Wie auch schon in TIMSS 2019 (63.6 %) ist dieser Anteil im Vergleich der drei Gruppen am niedrigsten. So gingen 75.6 Prozent der Kinder, bei denen ein Elternteil im Ausland geboren wurde, mindestens drei Jahre in einen Kindergarten (TIMSS 2019: 80.0 %); sind beide Elternteile in Deutschland geboren, liegt dieser Anteil bei 75.1 Prozent (TIMSS 2019: 73.7 %). Substanzielle Veränderungen lassen sich somit im Vergleich zu 2019 in TIMSS 2023 nicht feststellen.

Ein ähnliches Bild zeichnet sich für die Erwerbstätigkeit ab: Auch wenn eine direkte Zuordnung zu Vater und Mutter, wie sie bei TIMSS 2019 noch vorgenommen wurde, in TIMSS 2023 nicht mehr möglich ist, bleiben die Prozentanteile der Erwerbstätigkeit auf ähnlichem Niveau. Zusammengenommen zeichnet sich eine etwas niedrigere Quote bei Familien ab, bei denen beide Elternteile im Ausland geboren wurden (73.6 bzw. 80.6 %), etwas höher (82.5 bzw. 86.4 %) ist dieser Anteil, wenn ein Elternteil im Ausland geboren wurde. Den höchsten Anteil hinsichtlich der Erwerbstätigkeit weisen Familien auf, in denen beide Elternteile in Deutschland geboren wurden (90.3 bzw. 93.1 %).

Für die *Armutsgefährdung* können große Unterschiede zwischen Familien mit und ohne Migrationshintergrund verzeichnet werden, wobei sich im Vergleich von TIMSS 2019 und TIMSS 2023 nur geringfügige Veränderungen feststellen lassen. So zeigt sich, dass der höchste Anteil (42.2 %) in Familien, bei denen beide Elternteile im Ausland geboren wurden, vorliegt und der niedrigste (13.0 %), wenn beide Elternteile in Deutschland geboren wurden.

Tabelle 8.1: Angaben zur Verteilung von Viertklässler:innen im Hinblick auf den heimischen Sprachgebrauch, den Erwerbsstatus der Eltern sowie den Kindergartenbesuch und die *Armutsgefährdung* der Familien nach Migrationshintergrund (Anteile in Prozent)

im Ausland geboren	Viertklässler:innen	Gebrauch der deutschen Sprache zu Hause				mind. 3 Jahre Kindergarten-besuch	Erwerbstätigkeit[A]		Armuts-gefährdung der Familie
		immer	fast immer	manchmal	nie		Elternteil A	Elternteil B	
	%	%	%	%	%	%	%	%	%
beide Elternteile	25.6	10.3	20.4	55.9	13.4	64.1	73.6	80.6	42.2
ein Elternteil	14.3	38.5	27.7	31.7	2.1	75.6	86.4	82.5	21.7
kein Elternteil	60.0	79.8	15.7	4.2	0.3	75.1	93.1	90.3	13.0

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.
A = Voll- oder Teilzeit

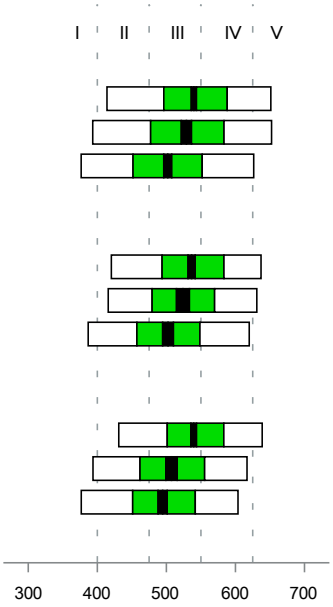
IEA: Trends in International Mathematics and Science Study © TIMSS 2023

8.6 Kompetenzunterschiede von Schüler:innen mit und ohne Migrationshintergrund in Deutschland im Trend

Die Kompetenzen in Mathematik, differenziert nach dem Migrationshintergrund von Schüler:innen, unterscheiden sich in ihren Abstufungen zwischen TIMSS 2007, 2019 und 2023 kaum (siehe Abbildung 8.3). Sowohl das generelle Leistungsniveau der Schülerschaft als auch die differenziert betrachteten Kompetenzen von Kindern mit und ohne Migrationshintergrund weisen ein sich kaum veränderndes Muster auf. Verglichen mit den Kompetenzen von Schüler:innen, deren Elternteile beide in Deutschland geboren wurden (TIMSS 2023: 540 Punkte), liegen die Schüler:innen, bei denen ein Elternteil im Ausland geboren wurde, mit einer Differenz von 11 Punkten signifikant darunter. Bei Schüler:innen, deren Eltern beide im Ausland geboren wurden, ist diese Differenz mit 38 Punkten noch ausgeprägter. Eine signifikante Veränderung der Mathematikkompetenzen über die TIMSS-Zyklen hinweg ergibt sich in TIMSS 2023 lediglich für Schüler:innen, bei denen ein Elternteil im Ausland geboren wurde. Hier lag der Nachteil in 2007 bei 32 Punkten, in 2019 bei 13 Punkten. In 2023 ist dieser Wert mit nur noch 11 Punkten signifikant niedriger.

Stellt man die Unterschiede in den Kompetenzständen der jeweiligen Gruppen gegenüber (Abbildung 8.4.), kann festgestellt werden, dass sich die Kompetenzstände der Grundschulkinder in den drei unterschiedlichen Gruppen zwischen 2019 und 2023 nicht signifikant unterscheiden. Dies deutet darauf hin, dass sich keine der hier berichteten Gruppen über die COVID-19-Pandemie hinweg substanziell in ihren Kompetenzen verändert hat.

Abbildung 8.3: Testleistungen der Schüler:innen in Mathematik in Deutschland nach Migrationshintergrund der Eltern – TIMSS 2007, 2019 und 2023 im Vergleich



im Ausland geboren	% (SE)	M (SE)	SD (SE)	M_d [B] (SE)
TIMSS 2023				
kein Elternteil	60.0 (1.1)	540 (1.9) ■	71 (1.2)	-RG-
ein Elternteil	14.3 (0.6)	529 (3.7) ▼	79 (2.8)	-11 (4.1)
beide Elternteile	25.6 (0.9)	502 (2.9) ▼	75 (1.9)	-38 (3.4)
TIMSS 2019 [A]				
kein Elternteil	63.6 (1.3)	537 (2.6) ■	66 (1.5)	-RG-
ein Elternteil	14.4 (0.9)	524 (4.7) ▼	67 (3.0)	-13 (5.3)
beide Elternteile	22.0 (1.0)	502 (3.9) ▼	71 (3.7)	-34 (4.7)
TIMSS 2007				
kein Elternteil	71.4 (1.4)	540 (2.0) ■	66 (1.7)	-RG-
ein Elternteil	11.4 (0.7)	507 (4.1) ▼	70 (3.9)	-32 (4.6)*
beide Elternteile	17.2 (1.0)	495 (3.3) ▼	69 (2.4)	-45 (3.9)

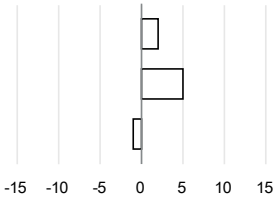
Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

* = Unterschied zu 2023 statistisch signifikant ($p \leq .05$)
▲ = Wert im Vergleich zu *kein Elternteil* statistisch signifikant höher ($p \leq .05$)
▼ = Wert im Vergleich zu *kein Elternteil* statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu *kein Elternteil* nicht statistisch signifikant ($p > .05$)
A = Abweichungen zu den Ergebnissen vorheriger Zyklen ergeben sich durch eine Vereinheitlichung des Rundungsverfahrens.
B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.
RG = Referenzgruppe

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Abbildung 8.4: Unterschiede in den mittleren Testleistungen in Mathematik im Trend in Deutschland nach Migrationshintergrund der Eltern – TIMSS 2007, 2019 und 2023 im Vergleich

im Ausland geboren	2007 M_{07} (SE)	2019 [A] M_{19} (SE)	2023 M_{23} (SE)	Veränderungen [B] M_{23}-M_{19} (SE)	M_{23}-M_{07} (SE)
kein Elternteil	540 (2.1)	537 (2.6)	540 (1.9)	2 (3.2)	-1 (2.8)
ein Elternteil	507 (4.3)	524 (4.7)	529 (3.7)	5 (6.0)	22 (5.6)
beide Elternteile	495 (3.5)	502 (3.9)	502 (2.9)	-1 (4.8)	6 (4.6)

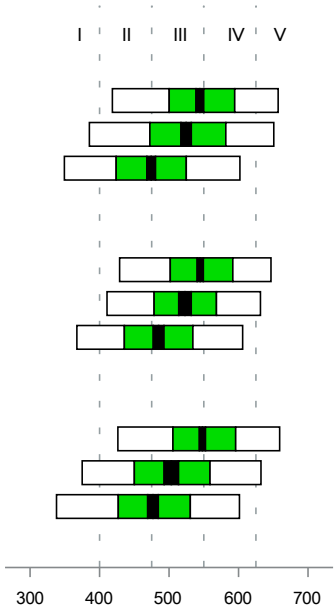


Keine statistisch signifikante Veränderung zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Veränderung zwischen 2019 und 2023 ($p \leq .05$)
A = Abweichungen zu den Ergebnissen vorheriger Zyklen ergeben sich durch eine Vereinheitlichung des Rundungsverfahrens.
B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Ein vergleichbares Bild zeigt sich auch für die naturwissenschaftlichen Kompetenzen (Abbildung 8.5). Kinder, deren Elternteile beide in Deutschland geboren wurden, zeigen die höchsten Kompetenzen (544 Punkte). Signifikant niedriger sind die Kompetenzen der Schüler:innen, bei denen ein Elternteil im Ausland geboren wurde (20 Punkte Differenz). Die größte Differenz in der mittleren Kompetenz mit 70 Punkten weisen die Schüler:innen auf, deren Elternteile beide im Ausland geboren wurden. Somit sind die relativen Differenzen zwischen den Gruppen in den naturwissenschaftlichen Kompetenzen – wie auch schon in TIMSS 2007 und 2019 – deutlich ausgeprägter als in den mathematischen Kompetenzen.

Abbildung 8.5: Testleistungen der Schüler:innen in den Naturwissenschaften in Deutschland nach Migrationshintergrund der Eltern – TIMSS 2007, 2019 und 2023 im Vergleich



im Ausland geboren	% (SE)	M (SE)	SD (SE)	M_d [C] (SE)
TIMSS 2023 [A]				
kein Elternteil	60.5 (1.1)	544 (2.8) ■	73 (1.7)	-RG-
ein Elternteil	14.3 (0.6)	525 (3.7) ▼	80 (2.3)	-20 (4.6)
beide Elternteile	25.2 (0.9)	474 (3.1) ▼	76 (2.2)	-70 (4.2)
TIMSS 2019 [B]				
kein Elternteil	64.3 (1.3)	545 (2.2) ■	66 (1.6)	-RG-
ein Elternteil	14.4 (0.9)	523 (4.5) ▼	68 (2.6)	-22 (5.0)
beide Elternteile	21.2 (1.0)	485 (3.8) ▼	72 (2.3)	-60 (4.4)
TIMSS 2007				
kein Elternteil	71.4 (1.4)	548 (2.2) ■	73 (1.7)	-RG-
ein Elternteil	11.4 (0.7)	503 (5.2) ▼	82 (4.8)	-45 (5.6)*
beide Elternteile	17.2 (1.0)	477 (3.8) ▼	79 (2.8)	-71 (4.4)

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

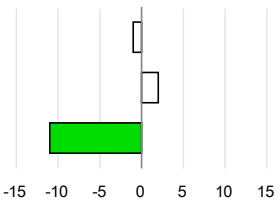
* = Unterschied zu 2023 statistisch signifikant ($p \leq .05$)
▲ = Wert im Vergleich zu *kein Elternteil* statistisch signifikant höher ($p \leq .05$)
▼ = Wert im Vergleich zu *kein Elternteil* statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zu *kein Elternteil* nicht statistisch signifikant ($p > .05$)
A = Abweichungen zur Abbildung 8.2 sind in einer minimal veränderten Stichprobenzusammensetzung begründet.
B = Abweichungen zu den Ergebnissen vorheriger Zyklen ergeben sich durch eine Vereinheitlichung des Rundungsverfahrens.
C = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.
RG = Referenzgruppe

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Rücken wiederum die Veränderungen der Kompetenzstände innerhalb der einzelnen Gruppen zwischen TIMSS 2007, 2019 und 2023 in den Fokus, ist zu beobachten, dass sich die Kompetenzen in den Naturwissenschaften nur bei Kindern, deren Eltern beide im Ausland geboren wurden, zwischen 2019 und 2023 unterscheiden (Abbildung 8.6). Hier schneiden die Kinder in TIMSS 2023 signifikant niedriger ab als noch in TIMSS 2019.

Abbildung 8.6: Unterschiede in den mittleren Testleistungen in den Naturwissenschaften im Trend in Deutschland nach Migrationshintergrund der Eltern – TIMSS 2007, 2019 und 2023 im Vergleich

im Ausland geboren	2007	2019	2023 [A]	Veränderungen [B]			
	M_{07} (SE)	M_{19} (SE)	M_{23} (SE)	M_{23}-M_{19}	(SE)	M_{23}-M_{07}	(SE)
kein Elternteil	548 (2.2)	545 (2.2)	544 (2.8)	-1	(3.6)	-4	(3.5)
ein Elternteil	503 (5.2)	523 (4.5)	525 (3.7)	2	(5.8)	21	(6.4)
beide Elternteile	477 (3.8)	485 (3.8)	474 (3.1)	-11	(4.9)	-3	(4.9)



Keine statistisch signifikante Veränderung zwischen 2019 und 2023 ($p > .05$)
Statistisch signifikante Veränderung zwischen 2019 und 2023 ($p \leq .05$)
A = Abweichungen zu Abbildung 8.2 sind in einer minimal veränderten Stichprobenzusammensetzung begründet.
B = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Für beide Domänen ist jedoch gleichermaßen festzustellen, dass die berichteten Differenzen der Gruppenmittelwerte jeweils mit deutlichen Kompetenzstreuungen innerhalb der Gruppen einhergehen. So kann sowohl für Mathematik- (Abbildung 8.3) als auch Naturwissenschaftskompetenzen (Abbildung 8.5) festgestellt werden, dass substanzielle Anteile von Kindern, bei denen ein Elternteil oder beide Elternteile im Ausland geboren wurden, nicht nur das mittlere Niveau der Kinder ohne Migrationshintergrund erreichen, sondern dieses auch übertreffen. Entsprechend gibt es in der Gruppe von Schüler:innen ohne Migrationshintergrund Anteile, deren Kompetenzen unter den mittleren Kompetenzen der Schüler:innen mit Migrationshintergrund liegen.

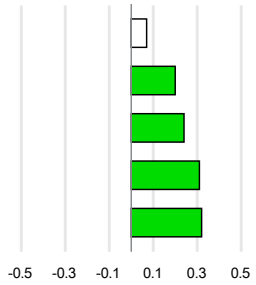
8.7 Fachbezogene Selbstkonzepte, Einstellung und soziale Eingebundenheit von Schüler:innen mit und ohne Migrationshintergrund

Schüler:innen in der Grundschule haben im wesentlichen positive Selbstkonzepte und auch positive Einstellungen zum Mathematik- und Sachunterricht (für Befunde zu TIMSS 2019 siehe Selter et al., 2020; Steffensky et al., 2020; für TIMSS 2023 siehe auch Kapitel 3 und 4 in diesem Band). Dies zeigt sich ebenfalls – wenn auch mit marginalen Veränderungen – in TIMSS 2023 (Abbildung 8.7). Sowohl für die Selbstkonzepte als auch für die Einstellungen liegen die Antworten bei einem Skalenbereich von 1 bis 4 durchgängig über einem theoretisch zu erwartendem Mittelwert von 2.5. Dabei zeigt sich, dass Kinder, deren Eltern beide in Deutschland geboren wurden, signifikant höhere mathematik- und sachunterrichtsbezogene Selbstkonzepte sowie eine positivere Einstellung zum Sachunterricht aufweisen als die Kinder, deren Eltern beide im Ausland geboren wurden. Bei der Einstellung zur Mathematik gibt es keine statistisch auffälligen Unterschiede zwischen den Gruppen.

Ein vergleichbares Muster zeigt sich auch in Bezug auf die soziale Eingebundenheit[2]. Auch hier geben Kinder, deren Eltern beide in Deutschland geboren wurden, an, sich signifikant besser einbezogen zu fühlen als Kinder, deren Eltern beide im Ausland geboren wurden.

Abbildung 8.7: Mittelwerte und Differenzen von mathematik- und naturwissenschaftsbezogenen Selbstkonzepten, Einstellungen zur Mathematik und zum Sachunterricht sowie sozialer Eingebundenheit nach Migrationshintergrund

	im Ausland geboren: kein Elternteil		im Ausland geboren: ein Elternteil		im Ausland geboren: beide Elternteile		Differenz [C]	
	M_K	(SE)	M_E	(SE)	M_B	(SE)	M_K-M_B	(SE)
Einstellung zur Mathematik	2.94	(0.02)	2.82	(0.04)	2.86	(0.04)	0.07	(0.04)
Mathematikbezogenes Selbstkonzept	3.13	(0.01)	3.05	(0.03)	2.93	(0.03)	0.20	(0.03)
Einstellung zum Sachunterricht [A]	3.23	(0.02)	3.21	(0.03)	2.99	(0.02)	0.24	(0.03)
Sachunterrichtsbezogenes Selbstkonzept [A]	3.24	(0.02)	3.17	(0.03)	2.93	(0.02)	0.31	(0.02)
Soziale Eingebundenheit [B]	9.47	(0.04)	9.41	(0.08)	9.15	(0.06)	0.32	(0.07)



Unterschied nicht statistisch signifikant ($p > .05$)
Unterschied statistisch signifikant ($p \leq .05$)
A = Abweichende Datengrundlage (weiterführende Informationen dazu finden sich im Text)
B = Der abweichende Wertebereich ist in der internationalen Skalierung begründet.
C = Inkonsistenzen in den berichteten Differenzen sind im Rundungsverfahren begründet.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

8.8 Zusammenfassung

Werden die Kompetenzen von Grundschüler:innen betrachtet, deren Eltern über unterschiedliche Migrationserfahrungen verfügen, zeigt sich im Rahmen von TIMSS 2023 ein recht einheitliches Bild in Deutschland. Viertklässler:innen, deren Eltern beide in Deutschland geboren wurden, weisen höhere Kompetenzstände in Mathematik und den Naturwissenschaften auf als Schüler:innen, bei denen ein Elternteil oder beide Elternteile im Ausland geboren wurden. Die jeweiligen Niveaus dieser Gruppen verändern sich in Deutschland über die Zeit nur gering, auch im internationalen Vergleich zeigen sich seit 2007 keine substanziellen Veränderungen.

Bei den Mathematikkompetenzen lassen sich für die drei Gruppen im Vergleich zu 2019 keine signifikanten Veränderungen feststellen. In den naturwissenschaftlichen Kompetenzen liegt das Niveau von Kindern, deren Eltern beide im Ausland geboren wurden, in 2023 um 10 Punkte niedriger und damit signifikant unter dem in 2019 festgestellten Niveau. Trotz der wiederkehrenden Befundlagen (siehe auch Tarelli et al., 2012 für TIMSS 2011 und Wendt et al., 2016 für TIMSS 2015) ist bei den Betrachtungen zu berücksichtigen, dass die berichteten Befunde sogenannte Mittelwertsaussagen sind. Dies bedeutet, dass hier über Gruppen von Schüler:innen gesprochen wird, die ein besonderes Merkmal teilen und somit auch als Gruppen unterschieden werden können. Das ist im Sinne einer vereinfachenden Modellbildung und zur übersichtlichen Situationsbeschrei-

2 Der Wertebereich der Skala *soziale Eingebundenheit* basiert auf der Zusammenfassung von fünf Einzelitems und einer anschließenden Skalierung durch die internationale Studienleitung. Hierbei werden international Werte kleiner als 7.2 als geringe Eingebundenheit und Werte über 9.6 als starke Eingebundenheit interpretiert (Yin & Fishbein, 2020).

bung sinnvoll und bewährt. Im gleichen Zuge ist jedoch festzustellen, dass sich die Individuen in diesen Gruppen in Bezug auf weitere persönliche Merkmale durchaus deutlich unterscheiden können und auch tatsächlich unterscheiden. So erreicht auch in TIMSS 2023 ein erheblicher Anteil von Kindern aus Familien mit Migrationserfahrung Kompetenzen, die zum Teil auch deutlich über den mittleren Kompetenzständen von Kindern aus Familien ohne Migrationserfahrung liegen (siehe Abschnitt 8.4). Dieser exemplarische Befund verdeutlicht, dass bei der Auseinandersetzung mit unterschiedlichen Migrationshintergründen eine Stereotypisierung von Personen, die bei diesen Betrachtungen oftmals nur das übergeordnete Merkmal ‚Migrationshintergrund' teilen, vermieden werden muss.

Vor diesem Hintergrund ist also zu berücksichtigen, dass eine pauschale Zuweisung von Gruppencharakteristika für die Beschreibung der Schülerkompetenzen zu kurz greift. Wird dieser Aspekt für Deutschland im Trend betrachtet, ist festzustellen, dass nicht von ‚den' Migrationsgruppen gesprochen werden kann und eine wachsende Heterogenität unter migrantischen Gruppen festzustellen ist. Dieses Phänomen wird differenziert unter dem Konzept der (Super-)Diversität betrachtet, analysiert und interpretiert (Nieswand, 2009; Schneider, 2020).

Inwieweit der Migrationshintergrund von Schüler:innen mit den erreichten Kompetenzständen zusammenhängt, sollte daher unter Einbezug weiterer Hintergrundmerkmale der Schüler:innen bzw. deren Eltern differenzierter betrachtet werden und wird Gegenstand der Analysen in Kapitel 9 dieses Bandes sein. Bei den betrachteten Befunden zu Kompetenzunterschieden von Schüler:innen ist in TIMSS 2023 festzustellen, dass neben bildungsrelevanten Merkmalen der Haushalte auch der Migrationshintergrund eine zentrale Rolle spielt.

Wie eingangs beschrieben gehen mit dem Migrationshintergrund weitere Heterogenitätsmerkmale einher, die zu berücksichtigen sind. So konnte im Rahmen von IGLU 2021 gezeigt werden, dass ein wesentliches, wenn auch nicht für alle Kinder mit Migrationshintergrund gleichermaßen relevantes, Merkmal einerseits der heimische Sprachgebrauch und andererseits auch das in der Schule wichtige Leseverständnis in Deutsch (Stubbe et al., 2023) ist.

Aktuell und auch zukünftig rückt im Umgang mit sprachlichen Defiziten von Kindern mit Migrationshintergrund der sprachsensible Fachunterricht bzw. ein herkunftssprachen-berücksichtigender Unterricht in den Fokus. Gerade auch für den Erwerb mathematischer und naturwissenschaftlicher Kompetenzen in der Grundschule stellt in der aktuellen Schul- und Unterrichtspraxis eine kompetente Beherrschung der deutschen Sprache eine wichtige Voraussetzung für gelingende Lernprozesse dar. Daher sollte an dieser Stelle bildungspolitisch angesetzt und zugleich die indirekte Förderung von Kindern über die Sprachförderung ihrer Eltern verstärkt in den Blick genommen werden. Damit geht einher, dass geplante bzw. auch realisierte Fördermaßnahmen sich nicht an den Stereotypen der Gruppen orientieren dürfen, sondern in jedem Fall eine individualisierte Analyse von Förderbedarfen vorauszugehen hat. Nicht alle Schüler:innen aus Familien mit Migrationserfahrungen haben einen sprachlichen und/oder fachlichen Förderbedarf, sodass dies im Rahmen der Förderung zu differenzieren ist.

Auch in TIMSS 2023 konnte darüber hinaus ähnlich wie in TIMSS 2019 bestätigt werden, dass sich Kinder mit Migrationshintergrund weniger sozial eingebunden fühlen und ein niedrigeres mathematikbezogenes sowie sachunterrichtsbezogenes Selbstkonzept haben als Kinder ohne Migrationshintergrund. Auch diese Aspekte sind für erfolgreiche Lernprozesse relevant und hinsichtlich des Abbaus von migrationsbedingten Leistungsdisparitäten zu berücksichtigen.

Literatur

Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Schümer, G., Stanat, P., Tillmann, K.-J., & Weiß, M. (Hrsg.). (2001). *PISA 2000. Zusammenfassung zentraler Befunde*. https://www.mpib-berlin.mpg.de/Pisa/ergebnisse.pdf

Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. wbv. https://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2016/pdf-bildungsbericht-2016/bildungsbericht-2016

Autor:innengruppe Bildungsberichterstattung. (2024). *Bildung in Deutschland 2024. Ein indikatorengestützter Bericht mit einer Analyse zu beruflicher Bildung*. wbv. https://www.bildungsbericht.de/de/bildungsberichte-seit-2006/bildungsbericht-2024/pdf-dateien-2024/bildungsbericht-2024.pdf

Bonsen, M., Kummer, N., & Bos, W. (2008). Schülerinnen und Schüler mit Migrationshintergrund. In W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther (Hrsg.), *TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 157–175). Waxmann.

Bourdieu, P. (1983). Ökonomisches Kapital, kulturelles Kapital, soziales Kapital. In R. Kreckel (Hrsg.), *Soziale Ungleichheiten* (S. 183–198). Schwartz.

Bourdieu, P., & Passerson, J. C. (1971). *Die Illusion der Chancengleichheit: Untersuchungen zur Soziologie des Bildungswesens am Beispiel Frankreichs*. Klett.

de Boer, H., & Merklinger, D. (Hrsg.). (2021). *Grundschule im Kontext von Flucht und Migration*. Kohlhammer.

Deutsche UNESCO-Kommission. (2017). *Bildungsagenda 2030. Aktionsrahmen für die Umsetzung von Sustainable Development Goal 4. Inklusive, chancengerechte und hochwertige Bildung sowie lebenslanges Lernen für alle. Kurzfassung der Deutschen UNESCO-Kommission*. https://www.unesco.de/sites/default/files/2018-01/Bildungsagenda%202030_Aktionsrahmen_Kurzfassung_DeutscheVersion_FINAL.pdf

Ditton, H., & Krüsken, J. (2006). Sozialer Kontext und schulische Leistungen – Zur Bildungsrelevanz segregierter Armut. *Zeitschrift für Soziologie der Erziehung und Sozialisation*, *26*(2), 135–157. https://doi.org/10.25656/01:5645

Esser, H. (2006). *Migration, Sprache und Integration.* Arbeitsstelle Interkulturelle Konflikte und gesellschaftliche Integration (AKI)-Forschungsbilanz 4. https://bibliothek.wzb.eu/pdf/2006/iv06-akibilanz4a.pdf

Gogolin, I. (1994). *Der monolinguale Habitus der multilingualen Schule*. Waxmann.

Henschel, S., Heppt, B., Camilla Rjosk, C., & Weirich, S. (2022). Zuwanderungsbezogene Disparitäten. In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (S. 181–219). Waxmann. https://doi.org/10.31244/9783830996064

Kasper, D., Wendt, H., Schwippert, K., & Köller, O. (2020). Trends in der Schülerzusammensetzung und mathematischen und naturwissenschaftlichen Kompetenzen. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 331–349). Waxmann. https://doi.org/10.31244/9783830993193

Klafki, W. (2007). *Neue Studien zur Bildungstheorie und Didaktik. Zeitgemäße Allgemeinbildung und kritisch-konstruktive Didaktik*. Beltz.

KMK (Ständige Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland). (2002). *Bericht ‚Zuwanderung'*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2002/2002_05_24-Zuwanderung.pdf

Nieswand, B. (2009). *Diversität und Gesellschaft*. https://www.mpg.de/359276/forschungs-Schwerpunkt

Schneider, J. (2020). *Was heißt Integration in superdiversen Nachbarschaften?* https://www.bpb.de/themen/migration-integration/kurzdossiers/integrationspolitik/308203/was-heisst-integration-in-superdiversen-nachbarschaften/

Selter, C., Walter, D., Heinze, A., Brandt, J., & Jentsch, A. (2020). Mathematische Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In K. Schwip-

pert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 56–113). Waxmann. https://doi.org/10.31244/9783830993193

Steffensky, M., Scholz, L. A., Kasper, D., & Köller, O. (2020). Naturwissenschaftliche Kompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 115–168). Waxmann. https://doi.org/10.31244/9783830993193

Stubbe, T. C., Kleinkorres, R. Krieg, M., Schaufelberger, R., & Schlitter, T. (2023). Soziale und migrationsbedingte Disparitäten in der Lesekompetenz von Viertklässlerinnen und Viertklässlern. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher & T. C. Stubbe (Hrsg.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (S. 151–177). Waxmann. https://doi.org/10.31244/9783830997009

Stubbe, T. C., Krieg, M., Beese, C., & Jusufi, D. (2020). Soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässlerinnen und Viertklässlern. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 263–289). Waxmann. https://doi.org/10.31244/9783830993193

Tarelli, I., Schwippert, K., & Stubbe, T. C. (2012). Mathematische und naturwissenschaftliche Kompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In W. Bos, H. Wendt, O. Köller & C. Selter (Hrsg.), *TIMSS 2011. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 247–267). Waxmann. https://elibrary.utb.de/doi/epdf/10.31244/9783830978145

UNESCO (United Nations Educational, Scientific and Cultural Organization). (2018). *Global education monitoring report 2019. Migration, displacement and education – Building bridges, not walls*. https://doi.org/10.54676/XDZD4287

Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality. Civic voluntarism in American politics*. Harvard University Press. https://doi.org/10.2307/j.ctv1pnc1k7

Walther, L., Fuchs, L. M., Schupp, J., & von Scheve, C. (2020). Living conditions and the mental health and well-being of refugees: Evidence from a large-scale German survey. *Journal of Immigrant Minority Health*, *22*, 903–913. https://doi.org/10.1007/s10903-019-00968-5

Wendt, H., Kasper, D., Bos, W., Vennemann, M., & Goy, M. (2017). Wie viele Punkte auf der TIMSS-Metrik entsprechen einem Lernjahr? Leistungszuwächse in Mathematik und Naturwissenschaften am Ende der Grundschulzeit. In T. Eckert & B. Gniewosz (Hrsg.), *Bildungsgerechtigkeit* (S. 121–153). Springer. https://doi.org/10.1007/978-3-658-15003-7_8

Wendt, H., Schwippert, K., & Stubbe, T. C. (2016). Mathematische und naturwissenschaftliche Kompetenzen von Schülerinnen und Schülern mit Migrationshintergrund. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 317–331). Waxmann. https://doi.org/10.25656/01:14022

Wendt, H., Schwippert, K., Stubbe, T. C., & Jusufi, D. (2020). Mathematische und naturwissenschaftliche Kompetenzen von Schülerinnen und Schülern mit und ohne Migrationshintergrund. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 291–313). Waxmann. https://doi.org/10.31244/9783830993193

Yin, L., & Fishbein, B. (2020). Creating and interpreting the TIMSS 2019 context questionnaire scales. In M. O. Martin, M. von Davier, & I. V. S. Mullis (Hrsg.), *Methods and Procedures: TIMSS 2019 Technical Report* (S. 16.1–16.331). TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2019/methods/chapter-16.html

Kapitel 9
Trends in der Schülerzusammensetzung und den mathematischen und naturwissenschaftlichen Kompetenzen am Ende der Grundschulzeit in Deutschland

Daniel Kasper, Jakob Sierau, Knut Schwippert und Olaf Köller

9.1 Einleitung

Durch die wiederholte Erhebung der mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässler:innen in der *Trends in International Mathematics and Science Study* (TIMSS) ist es möglich, Entwicklungen der Leistungsstände über die Jahre 2007, 2011, 2015, 2019 und 2023 nachzuzeichnen und diese in international vergleichender Perspektive einzuordnen. Wie in den vorangegangenen Kapiteln erläutert liegt der Fokus in dieser Berichtslegung auf der detaillierten Darstellung der Befunde für 2023. Darüber hinaus werden sowohl der langfristige Trend, d. h. die Veränderung in der Kompetenzverteilung von 2007 zu 2023, als auch der kurzfristige Trend, d. h. die Unterschiede in der Kompetenzverteilung von 2019 zu 2023, betrachtet.

Für Deutschland lassen sich aus der Perspektive von TIMSS 2023 im Vergleich zu Befunden der vorangegangenen Studienzyklen folgende Ergebnisse feststellen:

- *Mathematik*: Mit einem Leistungsmittelwert von 524 Punkten sind die Ergebnisse sowohl gegenüber 2007 (525 Punkte) als auch gegenüber 2019 (521 Punkte) in Deutschland stabil. Der Anteil der Schüler:innen, die lediglich die unteren beiden Kompetenzstufen erreichen, ist in TIMSS 2023 im langfristigen Trend hingegen leicht gestiegen und es lässt sich eine leichte Zunahme der Streuung der Leistungen beobachten (siehe Kapitel 3 in diesem Band).
- *Naturwissenschaften*: Mit einem Leistungsmittelwert von 515 Punkten liegen die Ergebnisse in 2023 3 Punkte unter dem Mittelwert von 2019 (518 Punkte) und 13 Punkte unter dem Mittelwert von 2007 (528 Punkte). Die Veränderung im langfristigen Trend ist statistisch signifikant. Auch lässt sich eine leichte Zunahme der Leistungsstreuung sowie ein gestiegener Anteil an Schüler:innen, die lediglich die unteren Kompetenzstufen erreichen, sowohl im kurzfristigen als auch im langfristigen Trend feststellen (siehe Kapitel 4 in diesem Band).

Bemerkenswert sind die 2023 im langfristigen Trend auf 25.1 Prozent (Mathematik) bzw. auf 29.8 Prozent (Naturwissenschaften) gestiegenen Anteile an Schüler:innen, die am Ende der Grundschulzeit unter Kompetenzstufe III liegen (siehe Kapitel 3 und 4 in diesem Band). Es ist davon auszugehen, dass diese in der Sekundarstufe I mit erheblichen Schwierigkeiten konfrontiert sein werden, dem weiteren Unterricht in Mathematik und den naturwissenschaftlichen Fächern zu folgen. Im Hinblick auf die Förderung dieser Kinder besteht somit dringender Handlungsbedarf.

Wie in den Kapiteln 5 bis 8 in diesem Band beschrieben, fallen die für Deutschland beobachteten Leistungsentwicklungen nicht für alle Subgruppen der Viertklässler:innen in gleicher Weise aus. So sind beispielsweise Jungen von der langfristigen negativen Trendentwicklung in Naturwissenschaften (20 Punkte Differenz zwischen 2007 und 2023) deutlich stärker betroffen als Mädchen (5 Punkte Differenz zwischen 2007 und 2023) und Schüler:innen mit Migrationshintergrund (ein Elternteil im Ausland geboren) weisen sogar eine positive langfristige Trendentwicklung in den Naturwissenschaften auf (siehe Kapitel 6 und 8 in diesem Band). Die für Deutschland beobachteten Trends müssen daher im Kontext von gesellschaftlichen und demografischen Entwicklungen sowie strukturellen und inhaltlichen Reformen im Bildungswesen diskutiert werden.

Im Rahmen der Berichtslegung von TIMSS 2015 wurden bereits zentrale Entwicklungen im Grundschulbereich seit der Jahrtausendwende skizziert (Kasper et al., 2016). In der Zusammenschau von Bildungsreformen der letzten 16 Jahre, die auf das Primarschulwesen in Deutschland gerichtet waren, offenbaren sich vielfältige und vielschichtige Veränderungsprozesse, die auf unterschiedlichen Ebenen des Bildungssystems umgesetzt wurden. Inwiefern diesen Maßnahmen auch Steuerungsfunktionen zuzuschreiben sind und ob diesen Reformen auch tatsächliche Veränderungen des mathematischen bzw. naturwissenschaftlichen Lehrens und Lernens in der Grundschule gefolgt sind, lässt sich nur schwer trennen. Offensichtlich ist hingegen, dass sich in den letzten Jahren die Zusammensetzung der Schülerschaft verändert hat (siehe Tabelle 9.1, Kapitel 7 und Kapitel 8 in diesem Band sowie die Ergebnisse von TIMSS 2015 und TIMSS 2019: Kasper et al., 2016; Kasper et al., 2020 und der *Internationalen Grundschul-Lese-Untersuchung* [IGLU] 2016: Bremerich-Vos, Wendt & Bos, 2017). In TIMSS 2019 konnten bei Modellierungen der Leistungstrends unter Berücksichtigung von Veränderungen der Struktur der Schülerschaft positive Leistungstrends für 2015 im Vergleich zu 2007 in Form eines Zuwachses von 15 Punkten (Naturwissenschaften) beobachtet werden (Kasper et al., 2020). In IGLU 2021 wurden, unter Berücksichtigung der Veränderung relevanter Merkmale der Schülerschaft über die IGLU-Erhebungszyklen hinweg, negative Trendentwicklungen für die Lesekompetenz beobachtet (Frey et al., 2023).

Neben Veränderungen in der Zusammensetzung der Schülerschaft könnte die kurzfristige Trendentwicklung allerdings auch durch die COVID-19-Pandemie bzw. die damit zusammenhängenden Schulschließungen beeinflusst sein. Auch wenn sich für Deutschland insgesamt im kurzfristigen Trend in Mathematik und Naturwissenschaften keine Leistungseinbußen durch die COVID-19-Pandemie zeigen, ist nicht auszuschließen, dass für Subgruppen der Schüler:innen durchaus negative Effekte zu verzeichnen sind. So konnte im Rahmen des Bildungstrends des *Instituts zur Qualitätsentwicklung im Bildungswesen* (IQB) 2021 etwa nachgewiesen werden, dass die erreichte Kompetenz in Mathematik von Viertklässler:innen in Deutschland u. a. mit der Qualität des Fernunterrichts (Einschätzung durch die Lehrkräfte) oder auch mit dem Anteil des Präsenzunterrichts (Ein-

schätzung durch die Lehrkräfte) positiv zusammenhing (Schneider et al., 2022). Für die Lesekompetenz von Viertklässler:innen in Deutschland im Rahmen von IGLU 2021 weisen die Autor:innen einen Pandemieeffekt von etwa 10 Punkten nach, d. h. unter sonst gleichen Bedingungen wäre davon auszugehen, dass die Schüler:innen in Deutschland in der Lesekompetenz 2021 rund 10 Punkte besser abgeschnitten hätten, wenn keine Pandemie bzw. damit einhergehende Veränderungen in der Beschulung der Schülerschaft vorhanden gewesen wären (Frey et al., 2023). Differenzielle Entwicklungsverläufe werden von den Autor:innen bei dem Ausweisen dieses Effekts allerdings nicht berücksichtigt. Neben der Qualität des Fernunterrichts und dem Anteil des Präsenzunterrichts kommen hier auch verschiedene andere differenzielle Faktoren in Betracht, die während der COVID-19-Pandemie relevant für die Schulleistung gewesen sein könnten. So wurden etwa Faktoren der häuslichen Lernumgebung (z. B. technische Ausstattung mit Endgeräten, ein ausreichender Internetzugang und ein eigener Raum zum Lernen), aber auch strukturelle Bedingungen auf Seiten der Lehrkräfte (z. B. Ausstattung der Schulen mit digitalen Medien und Fortbildungsangebote zum Umgang mit digitalen Medien) diskutiert (Eickelmann et al., 2019; Hasler-Waters et al., 2014; Huber & Helm, 2020; Tomasik et al., 2020; Wößmann et al., 2020). Um solche differenziellen Effekte in der kurzfristigen Trendentwicklung der mathematischen und naturwissenschaftlichen Kompetenzen in Deutschland analysieren zu können, wurden im Rahmen von TIMSS 2023 die auch im IQB-Bildungstrend 2021 eingesetzten Skalen zur COVID-19-Pandemie zu den Hintergrundbedingungen der Viertklässler:innen, der Erziehungsberechtigten, der Lehrkräfte sowie der Schulleitungen während der COVID-19-Pandemie miterhoben.[1]

In diesem Beitrag wird zunächst der Frage nachgegangen, inwieweit sich die Zusammensetzung der untersuchten Schülerpopulationen von TIMSS 2007, 2019 und 2023 verändert hat. Aufbauend darauf wird dargestellt, inwiefern differenzielle Effekte der COVID-19-Pandemie auf die Schülerleistungen in TIMSS 2023 zu verzeichnen sind. Im Anschluss daran wird auf die Trends in mathematischen und naturwissenschaftlichen Leistungen unter Kontrolle von Veränderungen zentraler Merkmale der Schülerpopulation eingegangen. Es wird untersucht, inwieweit sich die Veränderungen der mittleren Leistungen in Deutschland von TIMSS 2007 bzw. 2019 zu 2023 auf veränderte Schülerschaften zurückführen lassen. Weiterhin wird dargestellt, inwieweit sich eine etwas stärker von Vielfalt geprägte Schülerschaft auch in Klassenzusammensetzungen widerspiegelt.

1 Wir danken an dieser Stelle Herrn Dr. Stefan Schipolowski vom IQB für die Bereitstellung der entsprechenden Hintergrundfragen.

9.2 Charakteristika der Viertklässler:innen in TIMSS 2007, 2019 und 2023

Demografischer Wandel, Europäisierung, Migration sowie schulstrukturelle Veränderungen (z. B. Inklusion) bedingen eine veränderte Komposition der Schülerschaft an Grundschulen in Deutschland (Autorengruppe Bildungsberichterstattung, 2016, 2020; Autor:innengruppe Bildungsberichterstattung, 2024; Statistisches Bundesamt, 2017). In Tabelle 9.1 sind Entwicklungen für die Zyklen 2007, 2019 und 2023 dargestellt, die sich anhand zentraler Hintergrundmerkmale der Schüler:innen in TIMSS beobachten lassen. Der veränderte Anteil an Mädchen und Jungen in TIMSS 2023 ist dem Umstand geschuldet, dass erstmalig die Kategorie *Anderes* berücksichtigt wurde, die auf etwa 2 Prozent der Schüler:innen zutrifft. Der Anteil an Kindern aus armutsgefährdeten Familien ist 2023 (wie auch schon 2019) im langfristigen Trend statistisch signifikant gesunken. Im Vergleich zu den vorangegangenen Erhebungsrunden ist wiederum der Anteil von Kindern mit Migrationshintergrund (beide Eltern im Ausland geboren) in TIMSS 2023 mit 25.6 Prozent statistisch signifikant höher und spiegelt reale Veränderungen in der Komposition der Schülerschaft an Grundschulen in Deutschland wider (Autor:innengruppe Bildungsberichterstattung, 2024). Gleiches gilt für den ebenfalls im Vergleich zu 2007 statistisch signifikant höheren Anteil an Schüler:innen mit besonderen Unterstützungsbedarfen, dieser liegt in TIMSS 2023 bei 6 Prozent. Hierzu zählen in TIMSS Kinder, die nach Angabe der Schule einen, nach Verfahren des jeweiligen Landes der Bundesrepublik Deutschland diagnostizierten sonderpädagogischen Förderbedarf haben. Da für TIMSS 2007 keine spezifi-

Tabelle 9.1: Charakteristika von Viertklässler:innen in TIMSS 2007, 2019 und 2023

	TIMSS 2007		TIMSS 2019		TIMSS 2023	
	%	(*SE*)	%	(*SE*)	%	(*SE*)
Geschlecht						
Mädchen	49.0	(0.6)	49.6	(0.8)	48.4	(0.5)
Jungen	51.0	(0.6)	50.4	(0.8)	49.5	(0.6)
Anderes	—	—	—	—	2.0	(0.2)
Status						
Regelschüler:innen	96.6	(0.7)	94.1	(0.5)-	94.0	(0.3)-
Kinder mit bes. Unterstützungsbedarfen	3.4	(0.7)	5.9	(0.5)+	6.0	(0.3)+
Im Ausland geboren						
kein Elternteil	71.4	(1.4)	63.6	(1.3)-	60.0	(1.1)-▼
ein Elternteil	11.4	(0.7)	14.4	(0.9)+	14.3	(0.6)+
beide Elternteile	17.2	(1.0)	22.0	(1.0)+	25.6	(0.9)+▲
Kinder aus armutsgefährdeten Familien	33.9	(1.3)	21.5	(1.0)-	20.1	(0.8)-

Signifikanzniveau:

+/- = Unterschied zu TIMSS 2007 statistisch signifikant ($p \leq .05$)
▼/▲ = Unterschied zu TIMSS 2019 statistisch signifikant ($p \leq .05$)

© TIMSS 2023

schen Angaben zu sonderpädagogischen Förderbedarfen vorliegen, wurden hier Schüler:innen zu den Kindern mit besonderen Förderbedarfen gezählt, wenn es sich bei ihnen nach Angabe der Schule um Inklusionskinder handelte (Wendt, Bos et al., 2017). Die Ergebnisse weisen darauf hin, dass im Zuge der Inklusion zunehmend mehr Kinder mit sonderpädagogischen Förderbedarfen an Regelschulen unterrichtet werden.

9.3 Entwicklungen während der COVID-19-Pandemie

Um etwas über die Rahmenbedingungen des Lernens der Schüler:innen während der COVID-19-Pandemie zu erfahren, wurden in TIMSS 2023 sowohl die Schulleitungen als auch die Lehrkräfte und die Erziehungsberechtigten um Auskunft zu spezifischen Aspekten der Lernumgebung während der COVID-19-Pandemie gebeten. In den Schuljahren 2019/20 und 2020/21 waren die Grundschulen in Deutschland in den sogenannten Lockdowns laut OECD-Bericht 64 Tage vollständig geschlossen (OECD, 2021). Ausnahmen in Form eines Notbetriebs galten nur für Kinder, deren Erziehungsberechtigte in kritischer Infrastruktur arbeiteten oder bei Kindern von Alleinerziehenden. Über die 64 Tage hinaus kam es immer wieder zu lokalen bzw. regionalen Schulschließungen in Folge von Infektionsausbrüchen in Schulen.

In Tabelle 9.2 finden sich Angaben der Schulleitungen darüber, wie lange ihre Grundschule während der COVID-19-Pandemie für den Präsenzunterricht vollständig geschlossen war. Den Angaben der Schulleitungen zufolge war etwas mehr als die Hälfte (59 %) bzw. gut die Hälfte (49 %) der Schulen dabei vor allem in den Schuljahren 2019/20 und 2020/21 über ein bis drei Monate geschlossen. Zu bedenken ist hier, dass in beiden Schuljahren bundesweit alle Schulen bis auf den oben bereits erwähnten Notbetrieb zeitweise geschlossen waren. Die Auskünfte von 21.7 bzw. 28.4 Prozent der Schulleitungen, dass ihre Schulen gar nicht geschlossen waren, kann sich daher nur darauf beziehen, dass in den Lockdowns für wenige Kinder eine Notbetreuung angeboten wurde. Längere Phasen der Schulschließung in diesem Zeitraum berichten nur etwa 8 Prozent bzw. 11 Prozent der Schulen. In den Schuljahren 2021/22 und 2022/23 mussten nur noch sehr wenige Schulen (6 %) bzw. gar keine Schulen mehr als einen Monat vollständig schließen. Insbesondere im Schuljahr 2022/23 musste dann praktisch keine Schule mehr vollständig schließen (98 %). Die offizielle

Tabelle 9.2: Dauer der vollständigen Schulschließung für den Präsenzunterricht in den Schuljahren 2019/20, 2020/21, 2021/22 und 2022/23 nach Angabe der Schulleitungen

	keine vollständige Schulschließung		weniger als ein Monat		ein bis drei Monate		vier bis sechs Monate		mehr als sechs Monate	
	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
Schuljahr 2019/20	21.7	(2.7)	11.0	(1.8)	59.2	(3.5)	5.3	(1.5)	2.8	(1.7)
Schuljahr 2020/21	28.4	(3.0)	11.9	(3.0)	48.5	(3.7)	10.2	(1.9)	1.0	(1.0)
Schuljahr 2021/22	80.8	(3.2)	13.2	(2.8)	5.6	(1.6)	0.3	(0.0)	0.0	(0.0)
Schuljahr 2022/23	98.0	(1.7)	2.0	(1.7)	0.0	(0.0)	0.0	(0.0)	0.0	(0.0)

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

© TIMSS 2023

OECD-Statistik (OECD, 2021) sowie die Angaben der Schulleitungen in TIMSS 2023 lassen vermuten, dass die große Mehrheit der in TIMSS 2023 getesteten Viertklässler:innen während der COVID-19-Pandemie über einen längeren Zeitraum (rund drei Monate) keinen Präsenzunterricht erhalten hat. Betroffen waren dabei im Wesentlichen das erste und zweite Schuljahr der Kinder.

Bei den Angaben zu den Schulschließungen bleibt offen, was anstelle des Präsenzunterrichts von den Schulen angeboten wurde. Neben Wechselunterricht kommt hier beispielsweise Distanzunterricht, Distanzlernen (asynchron und eigenständig), Notbetreuung (mit Unterrichtsinhalten) und Notbetreuung (ohne Unterrichtsinhalte) in Frage. In Ergänzung zu der Dauer der vollständigen Schulschließungen für den Präsenzunterricht wurden die Schulleitungen deswegen gefragt, wie lange die Schüler:innen der TIMSS-Klassen während der gesamten Zeit der COVID-19-Pandemie in diesen Formaten unterrichtet wurden.

Aus Tabelle 9.3 wird ersichtlich, dass etwa 80 Prozent der Schulleitungen angaben, dass die Schüler:innen mindestens einen Monat im Wechselunterricht unterrichtet wurden. Gleichsam gaben etwa 60 Prozent der Schulleitungen an, dass die Schüler:innen mindestens einen Monat im Distanzunterricht gelernt haben. Distanzlernen (asynchron und eigenständig) und die Notbetreuung (mit Unterrichtsinhalten) waren nach Angaben der meisten Schulleitungen (64 % bzw. 84 %) Formate, in denen die Schüler:innen für mindestens einen Monat unterrichtet wurden. Eine Notbetreuung ohne Unterrichtsinhalte haben die Schüler:innen nach Angaben der meisten Schulleitungen nicht (73 %) oder zumindest nicht mehr als drei Monate (20 %) erfahren. Es wird demnach deutlich, dass der überwiegende Anteil der in TIMSS getesteten Viertklässler:innen zwar über einen längeren Zeitraum während der COVID-19-Pandemie keinen Präsenzunterricht erfahren hat, während dieser Zeit aber eine alternative Form des Unterrichts, die nicht aus einer reinen Notbetreuung ohne Unterrichtsinhalte bestand, nutzen konnte.

Tabelle 9.3: Dauer unterschiedlicher Formate, in denen die Schüler:innen der TIMSS-Klassen während der Zeit der COVID-19-Pandemie unterrichtet wurden (Angabe der Schulleitungen)

	gab es nicht		weniger als ein Monat		ein bis drei Monate		vier bis sechs Monate		mehr als sechs Monate	
	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
Wechselunterricht	9.4	(2.3)	10.2	(2.1)	50.7	(3.8)	20.6	(4.1)	9.1	(2.6)
Distanzunterricht (synchron)	35.8	(3.7)	6.7	(2.0)	39.3	(3.6)	15.6	(2.7)	2.7	(1.3)
Distanzlernen (asynchron, eigenständig)	28.2	(3.4)	7.7	(1.7)	42.0	(3.8)	16.7	(2.7)	5.5	(1.6)
Notbetreuung (mit Unterrichtsinhalten)	12.5	(2.7)	3.5	(1.0)	41.9	(3.8)	25.1	(2.8)	16.9	(3.2)
Notbetreuung (ohne Unterrichtsinhalte)	72.7	(3.1)	9.0	(2.0)	10.7	(1.9)	4.2	(1.8)	3.3	(0.9)

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Da ein großer Teil der in TIMSS getesteten Viertklässler:innen über einen längeren Zeitraum während der COVID-19-Pandemie zu Hause lernen musste, stellt sich zudem die Frage, wie diese Lernumgebung gestaltet war. Dies ist insofern relevant, als dass zum einen einige der von den Schulen angebotenen alternativen Unterrichtsformate spezifische Anforderungen an die Lernumgebung stellten (etwa erfordert online angebotener Distanzunterricht einen Internetzugang und

ein geeignetes Endgerät), zum anderen aber auch im IQB-Bildungstrend 2021 nachgewiesen werden konnte, dass einige Ausstattungsmerkmale der häuslichen Lernumgebung (wie etwa ein Drucker) während der COVID-19-Pandemie mit einem höheren Kompetenzwert von Viertklässler:innen in Mathematik einhergingen (Schneider et al., 2022).

Um ein Bild davon zu erhalten, wie die häusliche Lernumgebung der Viertklässler:innen während der Schulschließungen gestaltet war, wurden die Erziehungsberechtigten des Kindes befragt. In Tabelle 9.4 finden sich die entsprechenden Angaben. Mindestens 80 Prozent der Erziehungsberechtigten gaben an, dass es in der Lernumgebung des Kindes während der Schulschließung die ganze Zeit einen Raum zum ungestörten Lernen, einen eigenen Schreibtisch, einen ausreichenden Internetzugang und einen Drucker gab. Über ein eigenes für den Distanzunterricht geeignetes Endgerät verfügten laut Angaben der Erziehungsberechtigten nur etwa 64 Prozent der Kinder während der gesamten Zeit. 21 Prozent der Erziehungsberechtigten gaben hingegen an, dass das Kind nur einen Teil der Zeit über ein eigenes Endgerät verfügte. Worauf diese nur teilweise Verfügbarkeit des eigenen Endgerätes zurückzuführen ist (ob die Erziehungsberechtigten beispielsweise ein entsprechendes Endgerät erst kaufen mussten oder etwa zeitversetzt ein Endgerät durch die Schule bereitgestellt wurde), lässt sich den Daten nicht entnehmen.

Tabelle 9.4: Verfügbarkeit von Merkmalen der Lernumgebung des Kindes zur Zeit der Schulschließungen (Elternangaben)

	ja, die ganze Zeit		ja, einen Teil der Zeit		nein	
	%	(*SE*)	%	(*SE*)	%	(*SE*)
Raum zum ungestörten Lernen	84.9	(0.6)	10.8	(0.5)	4.3	(0.3)
Eigener Schreibtisch	89.1	(0.5)	4.8	(0.3)	6.1	(0.4)
Eigenes Endgerät	63.5	(0.9)	21.4	(0.6)	15.1	(0.9)
Ausreichender Internetzugang	86.8	(0.6)	9.4	(0.5)	3.8	(0.3)
Drucker	81.5	(0.7)	6.9	(0.4)	11.6	(0.6)
Weitere schulpflichtige Kinder	57.8	(0.7)	3.1	(0.2)	39.1	(0.7)
Kinder zwischen 0 und 5 Jahren	35.6	(0.8)	1.8	(0.2)	62.6	(0.8)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

In knapp 58 Prozent der Fälle lebten neben den in TIMSS getesteten Schüler:innen weitere schulpflichtige Kinder im Haushalt. Dies könnte zum einen dazu geführt haben, dass sich das Kind entsprechende Ressourcen der Lernumgebung mit diesen weiteren Kindern teilen musste, gleichsam könnte es aber auch für die Erziehungsberechtigten bei weiteren schulpflichtigen Kindern im Haushalt eine größere Herausforderung dargestellt haben, alle angemessen beim Lernen zu unterstützen. Insofern könnten weitere schulpflichtige Kinder in der Lernumgebung der Schüler:innen deren Lernerfolg geschmälert haben. Noch nicht schulpflichtige Kinder zwischen 0 und 5 Jahren im Haushalt konkurrieren zwar nicht oder weniger um die Ressourcen in der Lernumgebung der Viertklässler:innen, wohl aber um die Zeit, die die Erziehungsberechtigten zur Verfügung haben, um die schulpflichtigen Kinder beim Lernen angemessen zu unterstützen. Etwa 36 Prozent der Erziehungsberechtigten gaben an, dass sie Kinder zwischen 0 und 5 Jahren während der gesamten Zeit der Schulschließung im Haushalt betreuten.

Im IQB-Bildungstrend 2021 konnte gezeigt werden, dass die Mathematikkompetenz von Viertklässler:innen in Deutschland niedriger war, wenn während der COVID-19-Pandemie weitere schulpflichtige Kinder oder Kinder zwischen 0 und 5 Jahren im Haushalt lebten (Schneider et al., 2022).

Weiterhin wurden die Erziehungsberechtigten danach gefragt, wie gut sie ihre Kinder während des Distanzunterrichts beim Lernen zu Hause unterstützen konnten. Aus Tabelle 9.5 wird ersichtlich, dass die Erziehungsberechtigten nach eigenen Angaben ihre Kinder größtenteils (über 80 %) beim Lernen, sowohl während der Schulschließungen als auch im Schuljahr 2022/23, gut unterstützen konnten. Allerdings hatten während der Schulschließungen deutlich mehr Erziehungsberechtigte Probleme, ihre Kinder beim Lernen zu unterstützen als im Schuljahr 2022/23 (19 % bzw. 11 %). Dies mag auf die oben bereits beschriebene Mehrfachbelastung durch die notwendige simultane Unterstützung von mehreren Kindern im Haushalt zurückzuführen sein. Im IQB-Bildungstrend 2021 konnte festgestellt werden, dass die Unterstützung durch die Erziehungsberechtigten während der COVID-19-Pandemie einen positiven Effekt auf die Mathematikkompetenz von Viertklässler:innen in Deutschland hatte (Schneider et al., 2022).

Neben quantitativen Aspekten der Lernumgebung während der COVID-19-Pandemie ist auch nach der Qualität des Distanzunterrichts gefragt worden. So wurden sowohl die Erziehungsberechtigten als auch die Lehrkräfte danach gefragt, wie gut aus ihrer Perspektive der Distanzunterricht hinsichtlich verschiedener Aspekte funktioniert hat. Da die Fragen im gleichen Wortlaut gestellt wurden, können diese Einschätzungen in Abbildung 9.1 einander gegenübergestellt werden. An den Werten für die Gesamtskala ist zu erkennen, dass sowohl die Erziehungsberechtigten als auch die Lehrkräfte davon ausgehen, dass der Distanzunterricht eher gut funktioniert hat (dreistufige Antwortskala; 1 = sehr schlecht/eher schlecht, 2 = eher gut, 3 = sehr gut). Hinsichtlich der einzelnen Aspekte schätzen die Erziehungsberechtigten den Kontakt zwischen den Schüler:innen und den Lehrkräften, die Übermittlung von Arbeitsaufträgen und Lernmaterialien, die Rückmeldung zu bearbeiteten Arbeitsaufträgen durch die Lehrkräfte und die technische Umsetzung des Fernunterrichts auf Seiten der Schule bzw. der Lehrkräfte als etwas weniger gut gelungen ein als die Lehrkräfte. Umgekehrt schätzen die Lehrkräfte die technische Umsetzung des Fernunterrichts auf Seiten der Schüler:innen weniger gut ein als die Erziehungsberechtigten. Die Lehrkräfte unterscheiden sich vor allem hinsichtlich der Einschätzung zur Motivierung der Schüler:innen: Sachunterrichtslehrkräfte geben an, dass es ihnen deutlich besser gelungen sei, die Schüler:innen zu motivieren als dies Mathematiklehrkräfte für sich beanspruchten. Erziehungsberechtigte wiederum antworten im Mittel ähnlich wie die Mathematiklehrkräfte, wobei Erziehungsberechtigte im Gruppenvergleich

Tabelle 9.5: Angaben der Erziehungsberechtigten, wie gut sie ihre Kinder während des Distanzunterrichts beim Lernen zu Hause unterstützen konnten

			trifft nicht zu		trifft eher nicht zu		trifft eher zu		trifft genau zu	
	M	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
Während der Schulschließungen konnte ich mein Kind gut beim Lernen unterstützen.	3.2	(0.01)	4.3	(0.3)	14.9	(0.7)	40.0	(0.8)	40.7	(0.8)
Im aktuellen Schuljahr konnte ich mein Kind gut unterstützen.	3.3	(0.01)	2.6	(0.2)	8.1	(0.5)	46.4	(0.8)	43.0	(0.7)

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

 © TIMSS 2023

Abbildung 9.1: Einschätzungen der Erziehungsberechtigten und Lehrkräfte zur Qualität des Distanzunterrichts

Merkmal	einschätzende Person		*M*	(*SE*)
Kontakt zwischen Schüler:innen und Lehrkräften	Erziehungsberechtigte	31% / 35% / 34%	2.03	(0.02)
	Mathematiklehrkraft	8% / 56% / 36%	2.27	(0.04)
	Sachunterrichtslehrkraft	11% / 47% / 42%	2.31	(0.05)
Übermittlung von Arbeitsaufträgen und Lernmaterialien	Erziehungsberechtigte	17% / 34% / 50%	2.34	(0.02)
	Mathematiklehrkraft	2% / 46% / 52%	2.51	(0.04)
	Sachunterrichtslehrkraft	5% / 42% / 54%	2.49	(0.05)
Rückmeldung zu bearbeiteten Arbeitsaufträgen durch die Lehrkräfte	Erziehungsberechtigte	37% / 34% / 30%	1.93	(0.02)
	Mathematiklehrkraft	22% / 50% / 29%	2.07	(0.06)
	Sachunterrichtslehrkraft	23% / 49% / 28%	2.06	(0.05)
Motivierung der Schüler:innen	Erziehungsberechtigte	44% / 34% / 23%	1.78	(0.02)
	Mathematiklehrkraft	35% / 54% / 11%	1.76	(0.05)
	Sachunterrichtslehrkraft	28% / 56% / 16%	1.88	(0.06)
Technische Umsetzung des Fernunterrichts auf Seiten der Schule/Lehrkräfte	Erziehungsberechtigte	44% / 34% / 22%	1.78	(0.02)
	Mathematiklehrkraft	30% / 53% / 17%	1.87	(0.05)
	Sachunterrichtslehrkraft	32% / 47% / 22%	1.90	(0.05)
Technische Umsetzung des Fernunterrichts auf Seiten der Schüler:innen	Erziehungsberechtigte	15% / 39% / 47%	2.32	(0.02)
	Mathematiklehrkraft	40% / 48% / 12%	1.71	(0.05)
	Sachunterrichtslehrkraft	38% / 46% / 16%	1.79	(0.06)
Gesamtskala	Erziehungsberechtigte		2.03	(0.02)
	Mathematiklehrkraft		2.05	(0.03)
	Sachunterrichtslehrkraft		2.08	(0.04)

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.
Aufgrund sehr geringer Fallzahlen für die Antwortkategorie *hat sehr schlecht funktioniert* in den einzelnen Personengruppen wurde diese für die Darstellung im Balkendiagramm mit der Kategorie *hat eher schlecht funktioniert* zur Kategorie *hat nicht gut funktioniert* zusammengefasst.

hat nicht gut funktioniert
hat eher gut funktioniert
hat sehr gut funktioniert

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

mit 23 Prozent am häufigsten eine sehr gute Einschätzung, aber mit 44 Prozent ebenfalls am häufigsten eine nicht gute Einschätzung abgaben.

Die Ergebnisse zeigen somit, dass sich die Lernbedingungen während der COVID-19-Pandemie für die in TIMSS getesteten Viertklässler:innen recht unterschiedlich gestalteten. Daraus resultiert die Frage, ob die Qualität der alternativen Lernumgebungen einen differenziellen Effekt auf die Kompetenzstände der Viertklässler:innen in Mathematik und Naturwissenschaften hatte. Zur Beantwortung dieser Frage wurden in Mehrebenenmodellen die Kompetenzen der Schüler:innen in Mathematik und Naturwissenschaften durch die spezifischen Aspekte der durch die COVID-19-Pandemie bedingten Lernumgebung vorhergesagt. In Tabelle 9.6 finden sich die Ergebnisse dieser Analysen.

In Modell 1 wurden zunächst nur die spezifischen Aspekte der Lernumgebung als Prädiktoren für die Kompetenz in Mathematik und Naturwissenschaften berücksichtigt, in Modell 2 wurde darüber hinaus um Hintergrundmerkmale der Schüler:innen kontrolliert und in Modell 3 schließlich um entsprechende Kompositionsmerkmale der Klassen. Es zeigt sich, dass ohne Kontrolle von Hintergrundmerkmalen der Viertklässler:innen sowohl die Unterstützung durch die Erziehungsberechtigten (während der Schulschließungen und im aktuellen Schuljahr) als auch der Grad, wie der Fernunterricht funktioniert hat (Einschätzung

Tabelle 9.6: Effekte von Maßen der häuslichen Lernumgebung in der COVID-19-Pandemie auf die Kompetenzen der Schüler:innen in Mathematik und Naturwissenschaften (Mehrebenenmodell)

	Mathematik						Naturwissenschaften					
	Modell 1		Modell 2		Modell 3		Modell 1		Modell 2		Modell 3	
	b	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)
Konstante	415.0	(24.5)**	-1221.3	(68.5)**	-1579.8	(58.1)**	420.9	(29.0)**	-1324.6	(86.7)**	-1741.0	(93.8)**
Schülerebene												
Raum zum Lernen[A]	6.4	(7.8)^{ns}	6.6	(6.6)^{ns}	7.2	(6.6)^{ns}	10.8	(8.4)^{ns}	6.2	(6.7)^{ns}	6.8	(6.6)^{ns}
Eigener Schreibtisch[B]	-2.3	(7.6)^{ns}	-10.9	(6.8)^{ns}	-10.4	(6.9)^{ns}	0.1	(9.0)^{ns}	-6.9	(8.0)^{ns}	-6.2	(7.7)^{ns}
Eigenes Gerät[C]	-9.1	(4.2)*	-2.0	(3.6)^{ns}	-2.1	(3.6)^{ns}	-10.0	(5.4)^{ns}	-2.5	(4.9)^{ns}	-2.6	(4.8)^{ns}
Ausreichend Internet[D]	-1.3	(6.1)^{ns}	-10.7	(5.8)^{ns}	-11.3	(5.8)*	-3.8	(7.0)^{ns}	-13.6	(6.9)^{ns}	-14.3	(6.9)*
Drucker[E]	4.3	(5.4)^{ns}	-3.2	(5.2)^{ns}	-2.9	(5.1)^{ns}	9.5	(5.8)^{ns}	-1.6	(5.0)^{ns}	-1.0	(4.9)^{ns}
Weitere schulpflichtige Kinder[F]	0.3	(4.4)^{ns}	-1.1	(3.7)^{ns}	-0.8	(3.7)^{ns}	-4.9	(4.7)^{ns}	-6.0	(4.2)^{ns}	-5.8	(4.2)^{ns}
Kinder zwischen 0 und 5 Jahren[G]	7.9	(4.3)^{ns}	1.0	(3.8)^{ns}	1.4	(3.8)^{ns}	6.3	(4.7)^{ns}	0.1	(3.8)^{ns}	0.5	(3.8)^{ns}
Unterstützung durch Erziehungsberechtigte während der Schulschließung	9.2	(3.6)**	5.3	(3.4)^{ns}	5.6	(3.5)^{ns}	9.6	(3.3)**	4.3	(3.1)^{ns}	4.6	(3.1)^{ns}
Unterstützung durch Erziehungsberechtigte im aktuellen Schuljahr	9.0	(3.7)*	4.0	(3.5)^{ns}	4.2	(3.5)^{ns}	8.3	(4.1)*	2.1	(4.0)^{ns}	2.4	(4.0)^{ns}
Funktionieren des Fernunterrichts (Elterneinschätzung)	5.2	(2.5)*	2.0	(2.3)^{ns}	1.6	(2.3)^{ns}	-0.7	(3.1)^{ns}	-1.4	(2.6)^{ns}	-1.8	(2.6)^{ns}
MGH (ein Elternteil)[H]			3.9	(7.1)^{ns}	1.8	(7.3)^{ns}			25.3	(8.7)**	22.6	(8.8)**
MGH (kein Elternteil)[I]			6.6	(5.9)^{ns}	5.4	(6.9)^{ns}			40.9	(6.7)**	38.7	(7.6)**
Familien mit hohem SES[J]			6.5	(0.6)**	5.8	(0.7)**			7.2	(0.7)**	6.5	(0.7)**
Kognitive Fähigkeiten			10.6	(0.6)**	10.5	(0.6)**			10.8	(0.7)**	10.7	(0.7)**
Schulebene												
Anteil des Präsenzunterrichts (Lehrkräfteeinschätzung)	12.0	(29.4)^{ns}	-1.3	(20.5)^{ns}	-3.4	(20.1)^{ns}	7.4	(35.8)^{ns}	-12.6	(21.8)^{ns}	-15.5	(20.8)^{ns}
Funktionieren des Fernunterrichts (Lehrkräfteeinschätzung)	15.8	(5.9)**	11.3	(3.9)**	9.8	(3.7)**	16.8	(7.0)*	10.3	(4.3)*	8.2	(4.2)*
Mittelwert des SES					4.4	(1.7)**					5.0	(2.1)*
Anteil der Schüler:innen mit MGH (ein Elternteil)					2.3	(14.1)^{ns}					5.9	(19.1)^{ns}
Anteil der Schüler:innen mit MGH (kein Elternteil)					3.1	(8.3)^{ns}					8.5	(9.9)^{ns}
Mathematik												
Erklärte Varianzanteile												
Zwischen den Schulen (26.6%)	41.1		41.1		87.4							
Innerhalb der Schulen (73.3%)	4.2		27.9		27.9							
Gesamt	14.0		30.7		40.5							
Naturwissenschaften												
Erklärte Varianzanteile												
Zwischen den Schulen (25.7%)							32.4		32.4		88.6	
Innerhalb der Schulen (74.2%)							5.3		36.9		36.9	
Gesamt							11.9		35.8		49.5	

Anmerkungen: MGH = Migrationshintergrund; ICC = Interklassenkorrelation

b = Regressionsgewicht (unstandardisiert)

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

A = Raum zum Lernen (0 = Nein; 1 = Ja)

B = Eigener Schreibtisch (0 = Nein; 1 = Ja)

C = Eigenes Gerät (0 = Nein; 1 = Ja)

D = Ausreichend Internet (0 = Nein; 1 = Ja)

E = Drucker (0 = Nein; 1 = Ja)

F = Weitere schulpflichtige Kinder (0 = Nein; 1 = Ja)

G = Kinder zwischen 0 und 5 Jahren (0 = Nein; 1 = Ja)

H = Migrationshintergrund nach Geburtsland der Eltern (0 = beide Elternteile im Ausland geboren; 1 = ein Elternteil im Ausland geboren; 0 = kein Elternteil im Ausland geboren)

I = Migrationshintergrund nach Geburtsland der Eltern (0 = beide Elternteile im Ausland geboren; 0 = ein Elternteil im Ausland geboren; 1 = kein Elternteil im Ausland geboren)

J = Sozioökonomischer Status der Familie: Sozialindex (unter der Annahme der Messinvarianz seit TIMSS 2007 skaliert)

© TIMSS 2023

durch die Lehrkräfte), einen positiven Effekt auf die Kompetenz in Mathematik und Naturwissenschaften zeigt. In Mathematik sagen darüber hinaus die Verfügbarkeit eines geeigneten Geräts für den digitalen Unterricht und der Grad, wie der Fernunterricht funktioniert hat (Einschätzung durch die Erziehungsberechtigten), die Kompetenzunterschiede vorher. Nach Kontrolle von relevanten Schülermerkmalen kann nur noch der Grad, wie der Fernunterricht funktioniert hat (Einschätzung durch die Lehrkräfte), die Leistungsunterschiede in Mathematik und Naturwissenschaften vorhersagen. Hier zeigt sich, dass viele der pandemiespezifischen Merkmale der Lernumgebung stark mit dem sozialen Hintergrund der Schüler:innen korrelieren (etwa dem Grad, inwiefern Erziehungsberechtigte in der Lage sind, ihre Kinder beim Lernen zu unterstützen).

Der Effekt eines gut funktionierenden Fernunterrichts liegt bei 11 Punkten (Mathematik) bzw. 10 Punkten (Naturwissenschaften). Nach Kontrolle von relevanten Kompositionsmerkmalen der Klasse (Modell 3) reduziert sich dieser Effekt auf rund 10 Punkte (Mathematik) bzw. 8 Punkte (Naturwissenschaften). Das bedeutet, dass Klassen, in denen der Fernunterricht aus Sicht der Lehrkräfte sehr gut funktioniert hat (entspricht 4 Punkten auf der Gesamtskala), im Durchschnitt etwa 40 Punkte (Mathematik) bzw. 32 Punkte (Naturwissenschaften) mehr erreichten als Klassen, bei denen der Fernunterricht nicht gut funktioniert hat (entspricht 1 Punkt auf der Gesamtskala). In der TIMSS-Stichprobe konnte allerdings als minimaler Wert für den Grad, wie Fernunterricht funktioniert, 1.8 und als maximaler Wert 4.0 auf der Gesamtskala beobachtet werden, d. h. die in TIMSS zu beobachtenden Kompetenzunterschiede zulasten der COVID-19-Pandemie zwischen den Klassen lagen bei etwa 20 Punkten (Mathematik) bzw. 16 Punkten (Naturwissenschaften). Der Leistungsvorsprung (bzw. die Leistungseinbuße) von etwa 20 Punkten entspricht dabei etwa einem Drittel Schuljahr (Wendt, Kasper et al., 2017) und ist damit auch praktisch bedeutsam.

9.4 Trends in den Kompetenzen der Schüler:innen

In diesem Abschnitt werden die Trends in Mathematik und in den Naturwissenschaften unter Berücksichtigung einer veränderten Zusammensetzung der Schülerschaft betrachtet. Dabei ist zu berücksichtigen, dass die gerade vorgestellten spezifischen Aspekte der Lernumgebung während der COVID-19-Pandemie nicht direkter Bestandteil dieser Betrachtung sein können, da entsprechende Informationen nur im Rahmen von TIMSS 2023 erhoben wurden.

Im internationalen Vergleich zeigt sich, dass sich für mehr als die Hälfte aller Staaten, die – wie Deutschland – an TIMSS 2007, 2019 und 2023 teilgenommen haben, zwischen dem Jahr 2007 und dem Jahr 2023 in Mathematik und/oder Naturwissenschaften positive Veränderungen der mittleren Leistungswerte beobachten ließen (siehe auch Kapitel 3 und 4 in diesem Band). Für Deutschland ließen sich – wie auch für fünf weitere Teilnehmer (Neuseeland, Niederlande, Dänemark, Italien und Québec, Kanada) – in Mathematik hingegen keine signifikanten Veränderungen in den Leistungen zwischen TIMSS 2007 und 2023 feststellen. Für die Naturwissenschaften zeigt sich zwischen 2007 und 2023 ein statistisch signifikanter Leistungsrückgang von 13 Punkten auf der TIMSS-Gesamtskala. Nur bei vier weiteren Teilnehmerstaaten (Italien, Ungarn, Ontario, Kanada und Québec, Kanada) lässt sich ebenfalls ein statistisch signifikanter negativer Trend beobachten. Gleichzeitig lassen sich für den gleichen Zeitraum sowohl in der amtlichen Statistik (Autor:innengruppe Bildungsberichterstattung,

2024) als auch in den Schülerpopulationen Veränderungen in zentralen Merkmalen der Schüler:innen feststellen (siehe Abschnitt 9.2).

9.4.1 Trends in den mittleren Leistungen unter Berücksichtigung der Mehrebenenstruktur

Komplexe Trendanalysen bieten die Möglichkeit, Veränderungen der Leistungen zwischen TIMSS 2007, 2019 und 2023 zu modellieren, die sich ergeben hätten, wenn die Zusammensetzung der Schülerschaft hinsichtlich verschiedener soziodemografischer Merkmale zwischen den Zyklen konstant geblieben wäre. Entsprechende Analysen können Hinweise darauf liefern, inwieweit der bei TIMSS 2023 im Vergleich zu 2007 beobachtete gestiegene Anteil von Kindern mit Migrationshintergrund und sonderpädagogischem Förderbedarf sowie geringere Anteile von Kindern aus armutsgefährdeten Familien als mögliche Erklärung für die veränderten Leistungsmittelwerte herangezogen werden können.

Eine Möglichkeit, die Bedeutung dieser Veränderungen für die mathematischen und naturwissenschaftlichen Leistungen von Viertklässler:innen abzuschätzen und damit quasi einen ‚Nettoeffekt' der Trendentwicklung zu ermitteln, bietet die Anwendung von Trendmodellen. In diesen Modellen wird die Leistungsdifferenz zwischen den Studienzyklen ermittelt, die sich ergeben hätte, wenn eine, hinsichtlich der soziodemografischen Merkmale, gleiche Zusammensetzung der Schülerschaft über die Zeit angenommen wird. Als erklärende Faktoren der Leistung wurden der Messzeitpunkt, die soziodemografischen Merkmale und die Interaktionen zwischen den soziodemografischen Merkmalen und dem Messzeitpunkt berücksichtigt.

Für die Analysen kam ein lineares Mehrebenenmodell mit Zufallseffekten (McCulloch et al., 2008) zum Einsatz. Da die Schüler:innen in der TIMSS-Stichprobe im Klassenverbund getestet wurden (siehe auch Kapitel 2 in diesem Band), ist davon auszugehen, dass ihre Leistungen voneinander abhängig sind. Diese Abhängigkeiten werden im Mehrebenenmodell durch die Aufnahme eines klassenspezifischen *Intercept*-Wertes kontrolliert. Da die in TIMSS untersuchten Klassen außerdem eine Zufallsauswahl von allen zur definierten Zielpopulation gehörenden Klassen in Deutschland darstellen, wird der *Intercept*-Wert als zufällig angenommen, d. h. dass ein sogenanntes *Random Intercept Model* berechnet wird (Raudenbush & Bryk, 2002).

Die Tabellen 9.7 und 9.8 zeigen, wie sich die Leistungstrends von Viertklässler:innen in Deutschland in TIMSS unter Kontrolle von Schülercharakteristika darstellen. Als ein Ergebnis dieser Analyse ist zunächst festzuhalten, dass sich ein Viertel aller Unterschiede (Mathematik: 26.6 %; Naturwissenschaften: 25.7 %) in der beobachteten Leistung auf Unterschiede zwischen den Schulen bzw. Klassen zurückführen lässt und die verbleibenden etwa drei Viertel auf individuelle Unterschiede zwischen den einzelnen Schüler:innen innerhalb der Klassen zurückgehen. Diese Varianzanteile entsprechen Werten aus vergleichbaren Studien (Bremerich-Vos, Wendt & Bos, 2017; Kasper et al., 2016; Kasper et al., 2020).

In Modell 1 in den Tabellen 9.7 und 9.8 wurden jeweils die Leistungstrends unter Berücksichtigung der Mehrebenenstruktur der Daten als Ausgangspunkt der folgenden Betrachtungen geschätzt. Es zeigt sich, dass sich für Mathematik kein signifikanter Unterschied in den mittleren Schülerleistungen zwischen 2007 und 2023 nachweisen lässt. Für die Naturwissenschaften wiederum zeigt sich für den

Tabelle 9.7: Veränderungen der Leistungen in Mathematik zwischen TIMSS 2007, 2019 und 2023 unter Kontrolle von Merkmalen der Schüler:innen (Mehrebenenmodell)

	Modell 1		Modell 2		Modell 3		Modell 4		Modell 5		Modell 6	
	b	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)
Konstante	521.6	(3.6)**	515.7	(3.6)**	420.9	(7.6)**	-857.8	(30.4)**	-859.7	(30.7)**	-1316.3	(66.0)**
Testzeitpunkt												
TIMSS 2019[A]	-2.3	(4.6)[ns]	-1.7	(4.7)[ns]	2.6	(3.7)[ns]	-7.5	(2.7)**	-2.6	(4.3)[ns]	-4.6	(4.1)[ns]
TIMSS 2023[B]	0.4	(4.7)[ns]	-1.7	(5.0)[ns]	3.1	(3.8)[ns]	-4.9	(2.7)[ns]	-2.1	(3.8)[ns]	-4.7	(3.6)[ns]
Schülerebene												
Geschlecht[C]			11.4	(1.8)**	12.7	(2.0)**	19.3	(1.9)**	19.3	(1.9)**	19.2	(1.9)**
MGH (ein Elternteil)[D]					12.9	(2.6)**	4.5	(2.3)*	3.9	(3.3)[ns]	3.0	(3.3)[ns]
MGH (kein Elternteil)[E]					28.0	(1.9)**	10.5	(1.6)**	14.2	(2.6)**	11.7	(2.5)**
SPF[F]					80.8	(6.5)**	51.5	(4.8)**	52.1	(6.1)**	45.4	(5.8)**
Geschlecht × T19			-0.8	(3.0)[ns]	0.5	(3.1)[ns]	-1.1	(2.8)[ns]	-1.3	(10.8)[ns]	0.8	(10.2)[ns]
Geschlecht × T23			4.1	(3.3)[ns]	5.6	(3.5)[ns]	2.3	(3.1)[ns]	4.0	(9.3)[ns]	5.0	(9.6)[ns]
MGH (ein Elternteil) × T19									1.3	(5.5)[ns]	1.5	(5.4)[ns]
MGH (kein Elternteil) × T19									-7.7	(4.1)[ns]	-8.1	(4.0)*
MGH (ein Elternteil) × T23									0.7	(4.6)[ns]	0.1	(4.5)[ns]
MGH (kein Elternteil) × T23									-4.1	(3.7)[ns]	-4.2	(3.6)[ns]
Geschlecht × SPF × T19									0.4	(10.7)[ns]	-2.2	(10.1)[ns]
Geschlecht × SPF × T23									-1.8	(8.9)[ns]	-2.9	(9.2)[ns]
Familien mit hohem SES[G]							5.2	(0.2)**	5.2	(0.2)**	4.8	(0.3)**
Kognitive Fähigkeiten							8.1	(0.2)**	8.1	(0.2)**	8.0	(0.2)**
Schulebene												
Mittelwert des SES											4.9	(0.7)**
Anteil der Schüler:innen ohne SPF											22.1	(11.6)[ns]
Anteil der Schüler:innen mit MGH (ein Elternteil)											-0.7	(9.0)[ns]
Anteil der Schüler:innen ohne MGH (kein Elternteil)											14.7	(4.2)**
Erklärte Varianzanteile												
Zwischen den Schulen (26.6%)	0.1		0.1		0.1		0.1		0.1		76.7	
Innerhalb der Schulen (73.3%)	0.0		1.1		4.8		30.1		30.2		30.2	
Gesamt	0.0		0.8		3.6		22.1		22.2		42.6	

Anmerkungen: MGH = Migrationshintergrund; SPF = sonderpädagogischer Förderbedarf; T19 = TIMSS 2019; T23 = TIMSS 2023

b = Regressionsgewicht (unstandardisiert)

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

A = Testzeitpunkt (0 = TIMSS 2007; 1 = TIMSS 2019; 0 = TIMSS 2023)

B = Testzeitpunkt (0 = TIMSS 2007; 0 = TIMSS 2019; 1 = TIMSS 2023)

C = Geschlecht (0 = Mädchen; 1 = Jungen)

D = Migrationshintergrund nach Geburtsland der Eltern (0 = beide Elternteile im Ausland geboren; 1 = ein Elternteil im Ausland geboren; 0 = kein Elternteil im Ausland geboren)

E = Migrationshintergrund nach Geburtsland der Eltern (0 = beide Elternteile im Ausland geboren; 0 = ein Elternteil im Ausland geboren; 1 = kein Elternteil im Ausland geboren)

F = Sonderpädagogischer Förderbedarf (0 = sonderpädagogischer Förderbedarf; 1 = kein sonderpädagogischer Förderbedarf)

G = Sozioökonomischer Status der Familie: Sozialindex (unter der Annahme der Messinvarianz seit TIMSS 2007 skaliert)

© TIMSS 2023

gleichen Zeitraum ein negativer Trend in den mittleren Schülerleistungen. Dieser Befund entspricht der generellen Befundlage in den Kapiteln 3 und 4 in diesem Band. Wird die Zusammensetzung der Schülerschaft hinsichtlich des Geschlechtes kontrolliert (Modell 2), dann resultiert für Mathematik ein kleiner negativer Leistungstrend von 2007 zu 2023, der aber nicht statistisch signifikant ist. Für die Naturwissenschaften wiederum reduziert sich der beobachtete negative Leistungstrend zwischen 2007 und 2023 und ist nicht mehr signifikant. In Mathematik kann darüber hinaus im Modell 2 über alle Zyklen hinweg ein genereller Kompetenzvorsprung von 11 Punkten für Jungen beobachtet werden, während dieser Effekt in den Naturwissenschaften für die Zyklen unterschiedlich ausfällt: In 2007 beträgt der Vorteil der Jungen gegenüber den Mädchen etwa 14 Punkte. In den Zyklen 2019 (3 Punkte) und 2023 (0.2 Punkte) haben die Jungen praktisch keinen Kompetenzvorteil mehr gegenüber den Mädchen.

In Modell 3 und Modell 4 wurde zusätzlich um die Merkmale Migrationshintergrund, sonderpädagogischer Förderbedarf, sozioökonomischer Status und kognitive Fähigkeiten der Viertklässler:innen kontrolliert. Sowohl für Mathematik als auch Naturwissenschaften ist weiterhin kein signifikanter Leistungstrend zu beobachten. In beiden Modellen ist darüber hinaus ein Kompetenzvorsprung in Abhängigkeit des Migrationsstatus ersichtlich, wobei dieser sich unter Berücksichtigung des sozioökonomischen Status und der kognitiven Fähigkeiten der Schüler:innen mehr als (Mathematik) bzw. nahezu (Naturwissenschaften) halbiert. Mit knapp 30 Prozent (Mathematik) und 32 Prozent (Naturwissenschaften) erklärt das Modell 4 auch bereits substanzielle Varianzanteile der Schülerleistung innerhalb der Schulen.

In Modell 5 wurden in Ergänzung zu Modell 4 noch die Interaktionen zwischen dem Migrationshintergrund und dem Testzeitpunkt sowie die Interaktion zwischen dem Geschlecht, dem sonderpädagogischen Förderbedarf und dem Testzeitpunkt eingefügt. Diese Interaktionen werden allerdings nicht signifikant und auch der erklärte Varianzanteil innerhalb der Schulen ändert sich praktisch nicht. In Modell 6 wurden schließlich noch Kompositionsmerkmale der Klassen als erklärende Variablen mitaufgenommen. Der langfristige negative Trend in den Naturwissenschaften wird in diesem Modell wieder signifikant und beträgt rund 10 Punkte. Darüber hinaus ist ersichtlich, dass sowohl in Mathematik als auch in den Naturwissenschaften höhere Leistungen in Klassen erzielt werden, in denen die Schüler:innen im Durchschnitt einen höheren sozioökonomischen Status aufweisen. Die Kompositionsmerkmale erklären dabei knapp 77 Prozent (Mathematik) bzw. 80 Prozent (Naturwissenschaften) der Leistungsunterschiede zwischen den Klassen. Insgesamt kann das Modell 6 in Mathematik 43 Prozent der Leistungsvarianz erklären, in den Naturwissenschaften sind es 44 Prozent.

Um die Interpretation der Ergebnisse aus den Tabellen 9.7 und 9.8 für zentrale Schülergruppen zu erleichtern, werden im Folgenden die Ergebnisse noch einmal für die Schüler:innen nach Geschlecht, besonderem Unterstützungsbedarf und Migrationshintergrund genauer betrachtet. In Abschnitt 9.5 wird anschließend noch einmal auf die Ergebnisse zur Klassenzusammensetzung aus Modell 6 eingegangen.

Tabelle 9.8: Veränderungen der Leistungen in den Naturwissenschaften zwischen TIMSS 2007, 2019 und 2023 unter Kontrolle von Merkmalen der Schüler:innen (Mehrebenenmodell)

	Modell 1		Modell 2		Modell 3		Modell 4		Modell 5		Modell 6	
	b	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)	*b*	(*SE*)
Konstante	524.8	(3.8)**	517.6	(3.9)**	397.1	(7.9)**	-984.3	(36.5)**	-984.5	(37.9)**	-1371.0	(74.9)**
Testzeitpunkt												
TIMSS 2019[A]	-7.8	(4.8)[ns]	-2.4	(5.0)[ns]	4.7	(4.1)[ns]	-6.2	(3.2)[ns]	-5.9	(6.0)[ns]	-7.8	(5.8)[ns]
TIMSS 2023[B]	-11.1	(5.4)*	-4.0	(5.4)[ns]	5.1	(4.0)[ns]	-3.5	(2.8)[ns]	-7.9	(4.4)[ns]	-10.1	(4.2)**
Schülerebene												
Geschlecht[C]			13.9	(2.5)**	15.1	(2.6)**	22.4	(2.6)**	22.4	(2.6)**	22.3	(2.6)**
MGH (ein Elternteil)[D]					31.2	(2.8)**	21.7	(2.3)**	16.2	(4.3)**	15.1	(4.3)**
MGH (kein Elternteil)[E]					58.3	(2.4)**	39.0	(2.1)**	37.6	(3.6)**	34.6	(3.5)**
SPF[F]					82.6	(6.9)**	50.5	(5.0)**	55.2	(7.4)**	49.9	(6.9)**
Geschlecht × T19			-10.5	(4.0)**	-8.8	(4.2)*	-10.7	(3.9)**	-2.2	(12.0)[ns]	-1.0	(11.8)[ns]
Geschlecht × T23			-13.7	(3.8)**	-13.0	(4.1)**	-16.9	(3.6)**	-9.0	(11.8)[ns]	-9.2	(11.8)[ns]
MGH (ein Elternteil) × T19									7.2	(6.5)[ns]	7.2	(6.4)[ns]
MGH (kein Elternteil) × T19									-1.9	(5.5)[ns]	-2.5	(5.4)[ns]
MGH (ein Elternteil) × T23									8.8	(6.2)[ns]	8.4	(6.1)[ns]
MGH (kein Elternteil) × T23									5.4	(4.2)[ns]	5.2	(4.2)[ns]
Geschlecht × SPF × T19									-8.9	(11.7)[ns]	-10.5	(11.5)[ns]
Geschlecht × SPF × T23									-8.4	(12.4)[ns]	-8.2	(12.2)[ns]
Familien mit hohem SES[G]							6.3	(0.3)**	6.3	(0.3)**	5.8	(0.3)**
Kognitive Fähigkeiten							8.1	(0.2)**	8.1	(0.2)**	8.0	(0.2)**
Schulebene												
Mittelwert des SES											4.2	(0.8)**
Anteil der Schüler:innen ohne SPF											15.8	(12.2)[ns]
Anteil der Schüler:innen mit MGH (ein Elternteil)											4.6	(9.2)[ns]
Anteil der Schüler:innen ohne MGH (kein Elternteil)											18.4	(3.8)**
Erklärte Varianzanteile												
Zwischen den Schulen (25.7%)	1.5		1.5		1.5		1.5		1.5		79.5	
Innerhalb der Schulen (74.2%)	0.0		0.4		9.5		32.1		32.2		32.2	
Gesamt	0.4		0.7		7.5		24.2		24.3		44.3	

Anmerkungen: MGH = Migrationshintergrund; SPF = sonderpädagogischer Förderbedarf; T19 = TIMSS 2019; T23 = TIMSS 2023

b = Regressionsgewicht (unstandardisiert)

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

A = Testzeitpunkt (0 = TIMSS 2007; 1 = TIMSS 2019; 0 = TIMSS 2023)

B = Testzeitpunkt (0 = TIMSS 2007; 0 = TIMSS 2019; 1 = TIMSS 2023)

C = Geschlecht (0 = Mädchen; 1 = Jungen)

D = Migrationshintergrund nach Geburtsland der Eltern (0 = beide Elternteile im Ausland geboren; 1 = ein Elternteil im Ausland geboren; 0 = kein Elternteil im Ausland geboren)

E = Migrationshintergrund nach Geburtsland der Eltern (0 = beide Elternteile im Ausland geboren; 0 = ein Elternteil im Ausland geboren; 1 = kein Elternteil im Ausland geboren)

F = Sonderpädagogischer Förderbedarf (0 = sonderpädagogischer Förderbedarf; 1 = kein sonderpädagogischer Förderbedarf)

G = Sozioökonomischer Status der Familie: Sozialindex (unter der Annahme der Messinvarianz seit TIMSS 2007 skaliert)

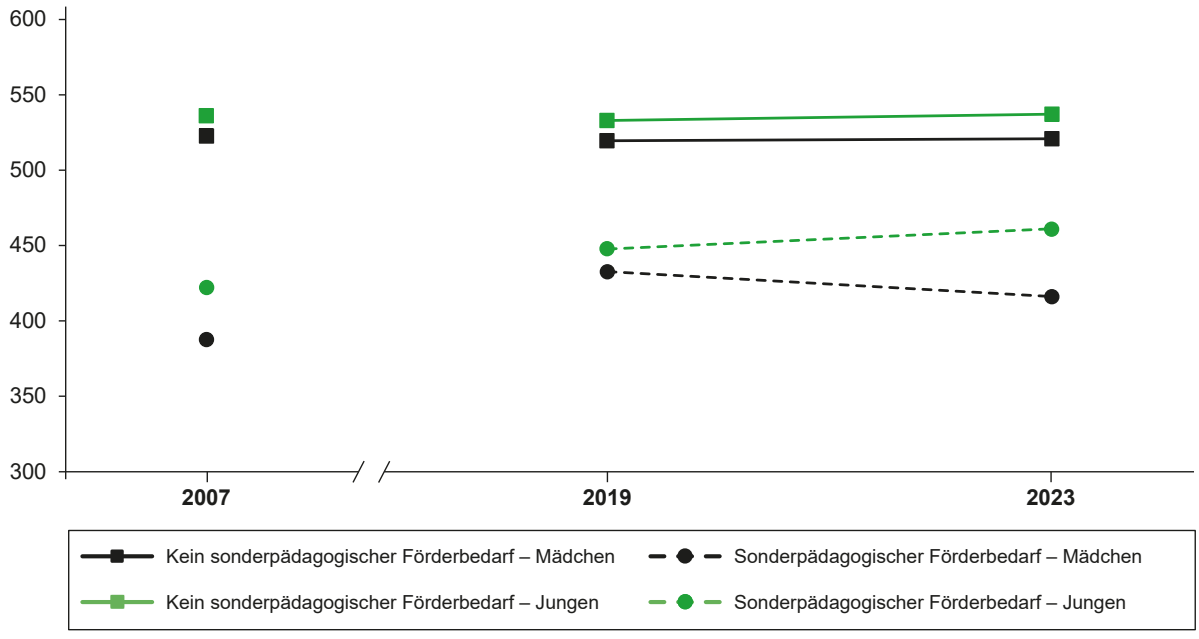
© TIMSS 2023

9.4.2 Trends in den Leistungen nach Geschlecht und Unterstützungsbedarfen

Wie in Kapitel 6 in diesem Band beschrieben, lässt sich für Deutschland für Mathematik keine Reduktion der Leistungsunterschiede zwischen Mädchen und Jungen von TIMSS 2007 zu 2023 beobachten. In den Naturwissenschaften haben sich die Geschlechterdifferenzen in diesem Zeitraum von 15 Punkten auf -2 Punkte statistisch signifikant reduziert. Für TIMSS 2015 konnten Wendt et al. (2016) zeigen, dass sich die von 2007 zu 2015 beobachtete Reduktion der Leistungsdisparitäten für Mathematik auf eine Veränderung der Zusammensetzung der Schülerschaft zurückführen ließ. Für die Naturwissenschaften blieb auch unter Kontrolle von Merkmalen der Schülerzusammensetzung ein signifikanter Leistungsvorsprung der Jungen vor den Mädchen bestehen. Die Ergebnisse wiesen dabei vor allem auf eine Interaktion zwischen dem Geschlecht und dem sonderpädagogischen Förderbedarf hin.

In Abbildung 9.2 sind die Leistungsentwicklungen für Regelschüler:innen sowie Kinder mit besonderen Unterstützungsbedarfen getrennt nach Geschlecht für die Mathematik dargestellt. Es zeigt sich folgendes Bild: Zunächst ist festzustellen, dass Kinder mit besonderen Unterstützungsbedarfen in allen Studienzyklen signifikant schlechtere Leistungen in Mathematik erzielen. Dem Modell 5 in Tabelle 9.7 ist zu entnehmen, dass diese im Durchschnitt über alle Studienzyklen hinweg einen Unterschied von 52.1 Punkten in Mathematik ausmachen. Darüber hinaus ist Abbildung 9.2 zu entnehmen, dass die geschlechterspezifischen Leistungsunterschiede zwischen Kindern mit besonderen Unterstützungsbedarfen insbesondere in 2007 und 2023 deutlich größer ausfallen als zwischen Regelschüler:innen. Wie in den Modellen 5 und 6 in Tabelle 9.7 zu erkennen ist, sind die entsprechenden Interaktionen zwischen Geschlecht, besonderen Unterstützungsbedarfen und Testzeitpunkt jedoch nicht signifikant, d. h., dieser diffe-

Abbildung 9.2: Veränderungen der Leistungen in Mathematik zwischen TIMSS 2007, 2019 und 2023 nach Geschlecht und Unterstützungsbedarfen

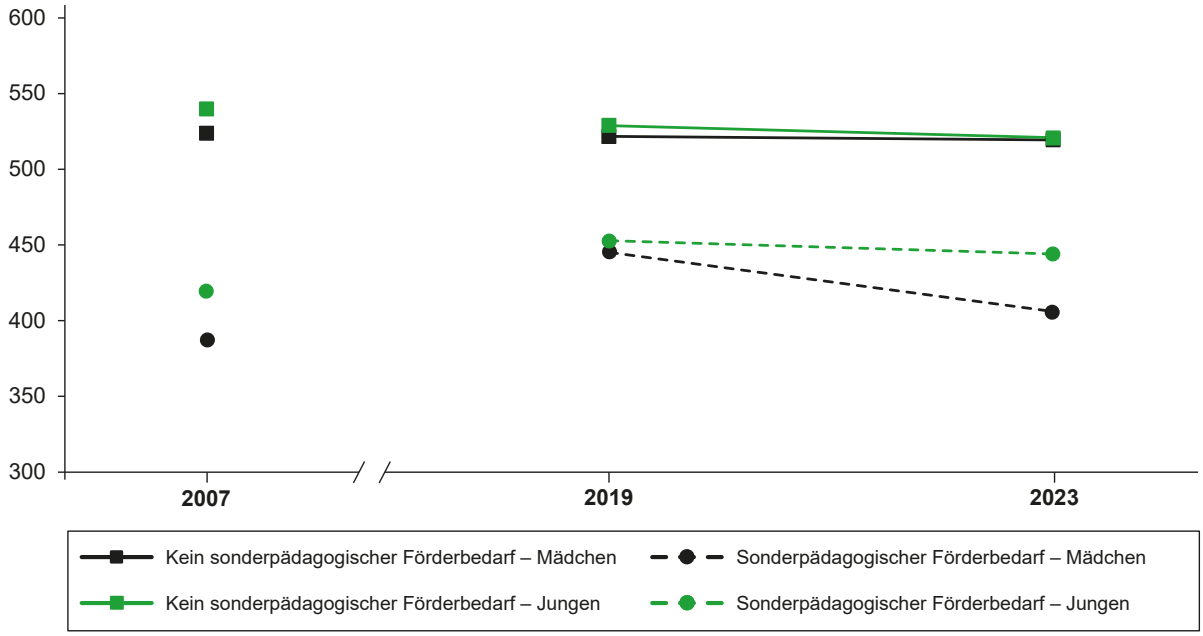


IEA: Trends in International Mathematics and Science Study © TIMSS 2023

renzielle Verlauf muss dem Zufall zugeschrieben werden. Für die Berechnung der geschlechterbezogenen Leistungsunterschiede sind hingegen substanzielle Leistungsunterschiede zugunsten der Jungen nachzuweisen. Diese betragen für Mathematik im Durchschnitt über alle Studienzyklen rund 19 Punkte. Ein Trend hinsichtlich der Reduktion der geschlechterspezifischen Leistungsdisparitäten von 2007 zu 2023 zeigt sich weder für die Regelschüler:innen noch die Kinder mit besonderen Unterstützungsbedarfen. Auch unter Berücksichtigung der Veränderungen in der Schülerzusammensetzung über die Zeit (Modelle 5 und 6) sind die geschlechterbezogenen Leistungsdisparitäten weiterhin nachzuweisen.

Die Ergebnisse zu den Geschlechterdifferenzen fallen für die Naturwissenschaften deutlich anders aus als für Mathematik. Zwar zeigt sich unter Berücksichtigung der Mehrebenenstruktur auch für die Naturwissenschaften ein signifikanter Vorteil der Jungen von rund 14 Punkten, dieser hängt aber, anders als bei Mathematik, vom Studienzyklus ab (Modell 2 in Tabelle 9.8). Die 14 Punkte zugunsten der Jungen zeigen sich in TIMSS 2007, reduzieren sich auf 4 Punkte in TIMSS 2019 und sind in TIMSS 2023 praktisch nicht mehr vorhanden. Die entsprechenden Interaktionen ‚Geschlecht x Zyklus' werden in Tabelle 9.8 jeweils signifikant. In Abbildung 9.3 wird ersichtlich, dass sich die Leistungen von Mädchen und Jungen mit sonderpädagogischem Förderbedarf und ohne sonderpädagogischen Förderbedarf in TIMSS 2007 deutlich voneinander unterscheiden. In TIMSS 2019 zeigen sich diese Unterschiede hingegen nicht und in TIMSS 2023 nur noch bei Schüler:innen mit sonderpädagogischem Förderbedarf. Allerdings ist diese Interaktion nicht signifikant (Modell 5 und 6, Tabelle 9.8), sodass dieser differenzielle Verlauf dem Zufall zugeschrieben werden muss.

Abbildung 9.3: Veränderungen der Leistungen in den Naturwissenschaften zwischen TIMSS 2007, 2019 und 2023 unter Kontrolle von Merkmalen der Schüler:innen nach Geschlecht und Unterstützungsbedarfen

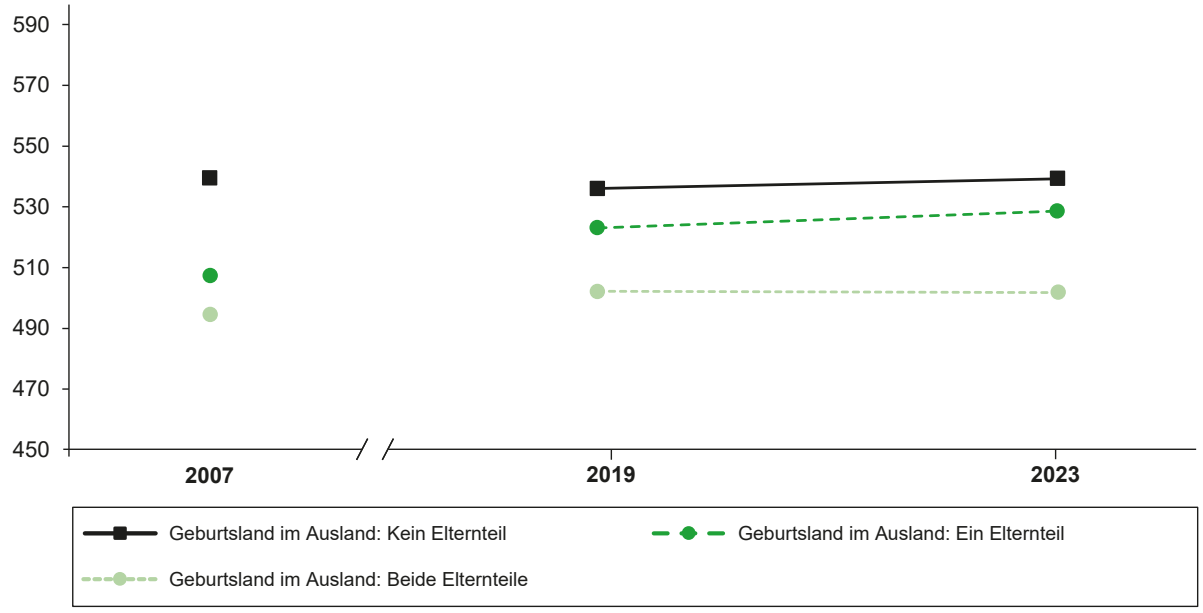


IEA: Trends in International Mathematics and Science Study © TIMSS 2023

9.4.3 Trends in den mittleren Leistungen nach Migrationshintergrund

Wie in Kapitel 8 beschrieben, lässt sich für Deutschland sowohl für Mathematik als auch für die Naturwissenschaften eine Reduktion der migrationsbezogenen Leistungsdisparitäten von TIMSS 2007 zu 2023 beobachten. In den Abbildungen 9.4 und 9.5 sind diese Ergebnisse grafisch dargestellt. In allen Studienzyklen lassen sich für Kinder mit Migrationshintergrund im Vergleich zu Kindern, deren Eltern beide in Deutschland geboren wurden, signifikant schlechtere Leistungen feststellen. Deutlich wird ein leichter Trend abnehmender migrationsspezifischer Leistungsunterschiede, der sich allerdings bei komplexer Modellierung (Modell 5 und Modell 6 in Tabelle 9.7 und Tabelle 9.8) für die Naturwissenschaften nicht und für Mathematik nur für TIMSS 2023 statistisch absichern lässt. Kinder mit einem in Deutschland und einem im Ausland geborenen Elternteil verzeichnen eine um 27 Punkte (Mathematik) bzw. 50.2 Punkte (Naturwissenschaften) höhere Leistung gegenüber Kindern, bei denen beide Elternteile im Ausland geboren wurden. Kinder ohne Migrationshintergrund wiederum verzeichnen sogar eine um 37.8 Punkte (Mathematik) bzw. 72.9 Punkte (Naturwissenschaften) höhere Leistung gegenüber Kindern, bei denen beide Elternteile im Ausland geboren wurden. Unter weiterer Kontrolle von sozialer Herkunft, kognitiven Grundfähigkeiten sowie Veränderungen in den Schülerzusammensetzungen (Modell 6) reduzieren sich diese Disparitäten. Für Kinder, bei denen ein Elternteil im Ausland geboren wurde, ist der Unterschied zu Kindern, bei denen beide Elternteile im Ausland geboren wurden, in Mathematik nicht mehr signifikant. Für die Naturwissenschaften bleiben sie weiterhin bestehen.

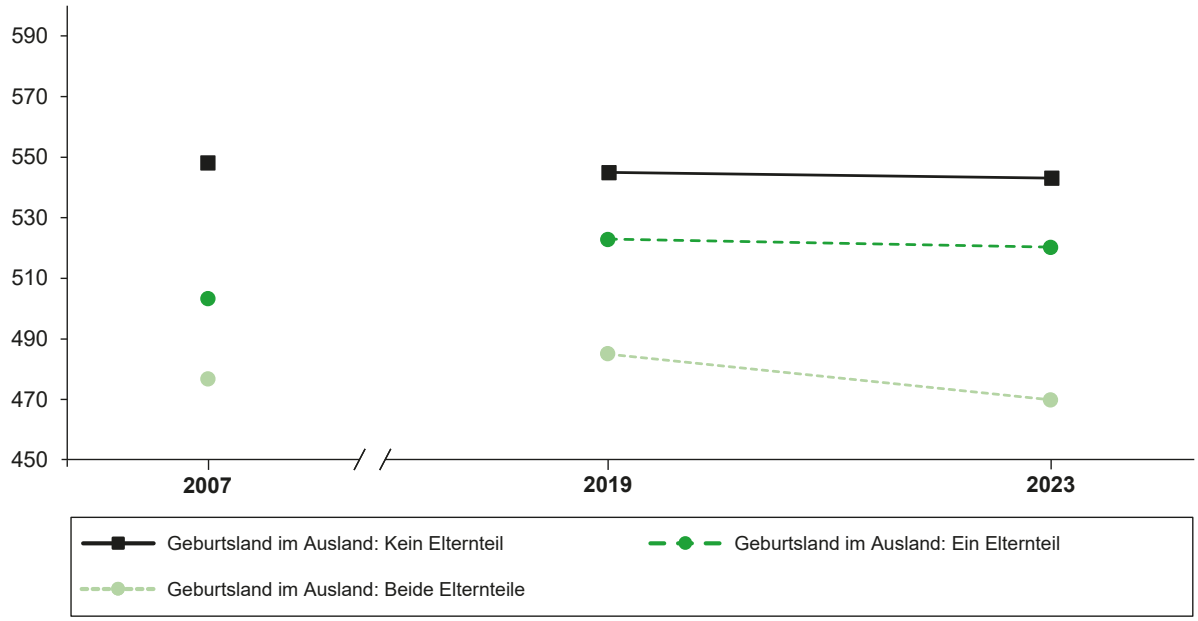
Abbildung 9.4: Veränderungen der Leistungen in Mathematik zwischen TIMSS 2007, 2019 und 2023 nach Migrationshintergrund



IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

Abbildung 9.5: Veränderungen der Leistungen in den Naturwissenschaften zwischen TIMSS 2007, 2019 und 2023 nach Migrationshintergrund



IEA: Trends in International Mathematics and Science Study © TIMSS 2023

9.5 Veränderungen in den Schülerzusammensetzungen

Wie in den vorangegangenen Abschnitten beschrieben, lassen sich für Grundschulen in Deutschland im Trend sowohl eine veränderte Schülerzusammensetzung mit gestiegenen Anteilen an Kindern mit Migrationshintergrund sowie höheren Anteilen von Kindern mit besonderen Unterstützungsbedarfen als auch eine Zunahme an Leistungsheterogenität beobachten. Anlass zur Sorge bieten die gestiegenen Anteile an Kindern, die lediglich die Kompetenzstufen I und II erreichen. Soll der gewachsenen Heterogenität Rechnung getragen werden, so gilt es, den Unterricht adaptiver bzw. individualisierter zu gestalten. Wer adaptiv lehrt, ist fähig, den „Unterricht so auf die individuellen Voraussetzungen der Schüler:innen auszurichten und während des Unterrichts laufend anzupassen, dass für möglichst viele Schüler:innen günstige Bedingungen für das Erreichen der Lernziele geschaffen werden“ (Brühwiler, 2014, S. 74). Der damit formulierte Anspruch ist hoch und ob bzw. in welchem Maß er zu erfüllen ist, hängt nicht zuletzt von der Zusammensetzung der jeweiligen Klasse ab. Zudem ist vor dem Hintergrund der eher geringen nominellen Veränderungen an Schüleranteilen zu fragen, an welchem Anteil an Schulen sich bedeutsame Veränderungen in den Schülerzusammensetzungen ergeben haben.

In diesem Zusammenhang ist es von besonderem Interesse, die Segregation von Schulen zu analysieren. Die Zuweisung von Schüler:innen zu Grundschulen erfolgt nicht zufällig, vielmehr ist nach wie vor die geografische Nähe der Schule zum Wohnort der Familie das entscheidende Kriterium dafür, auf welche Schule ein Kind geht. Insbesondere Großstädte sind heute stark sozial segregiert. Weiterhin verteilt sich die Bevölkerung mit Migrationshintergrund nicht zufällig über alle Stadtteile, sondern konzentriert sich regional (Bonsen et al., 2010). Als Folge solcher Segregationsprozesse und der Kopplung von Nachbarschaft und gewählter Schule ist zu erwarten, dass sich die Kompositionen von Klassen bzw.

Schulen erheblich unterscheiden und manche Schulen dementsprechend eine überwiegend benachteiligte Schülerschaft aufnehmen, andere eine privilegierte. Bremerich-Vos, Wendt und Hußmann (2017) konnten auf Basis der Daten von IGLU 2016 bereits zeigen, dass sich die Schülerzusammensetzungen zwischen Grundschulen erheblich unterscheiden. In den Abbildungen 9.6 und 9.7 werden die Klassenzusammensetzungen an Grundschulen in Deutschland nach verschiedenen Diversitätsmerkmalen im Vergleich von TIMSS 2007, 2019 und 2023 als Merkmale der Schulen betrachtet.

Abbildung 9.6: Klassenanteile von Schüler:innen nach unterschiedlichen Diversitätsmerkmalen in TIMSS 2007, 2019 und 2023 (Schulen nach Zusammensetzungsmerkmalen in Prozent) I

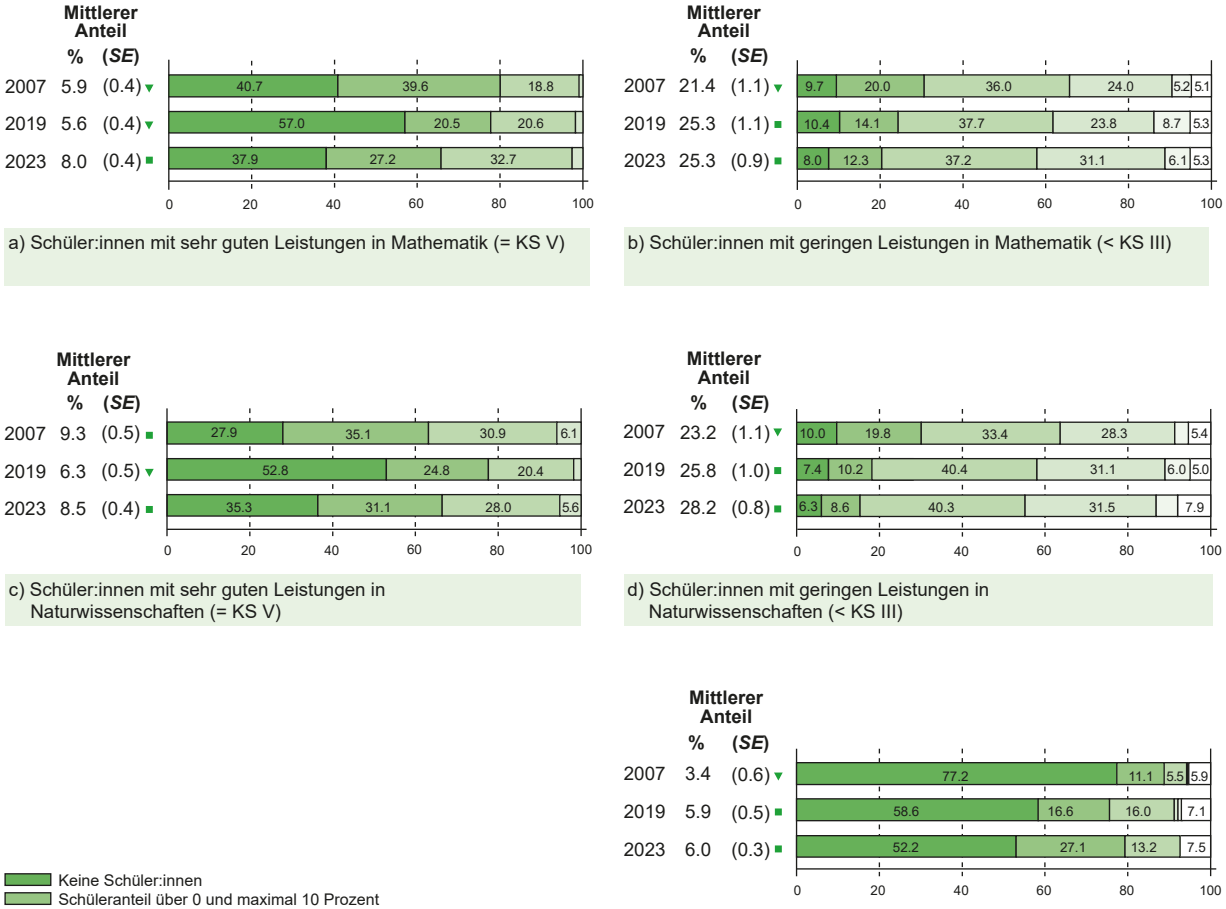
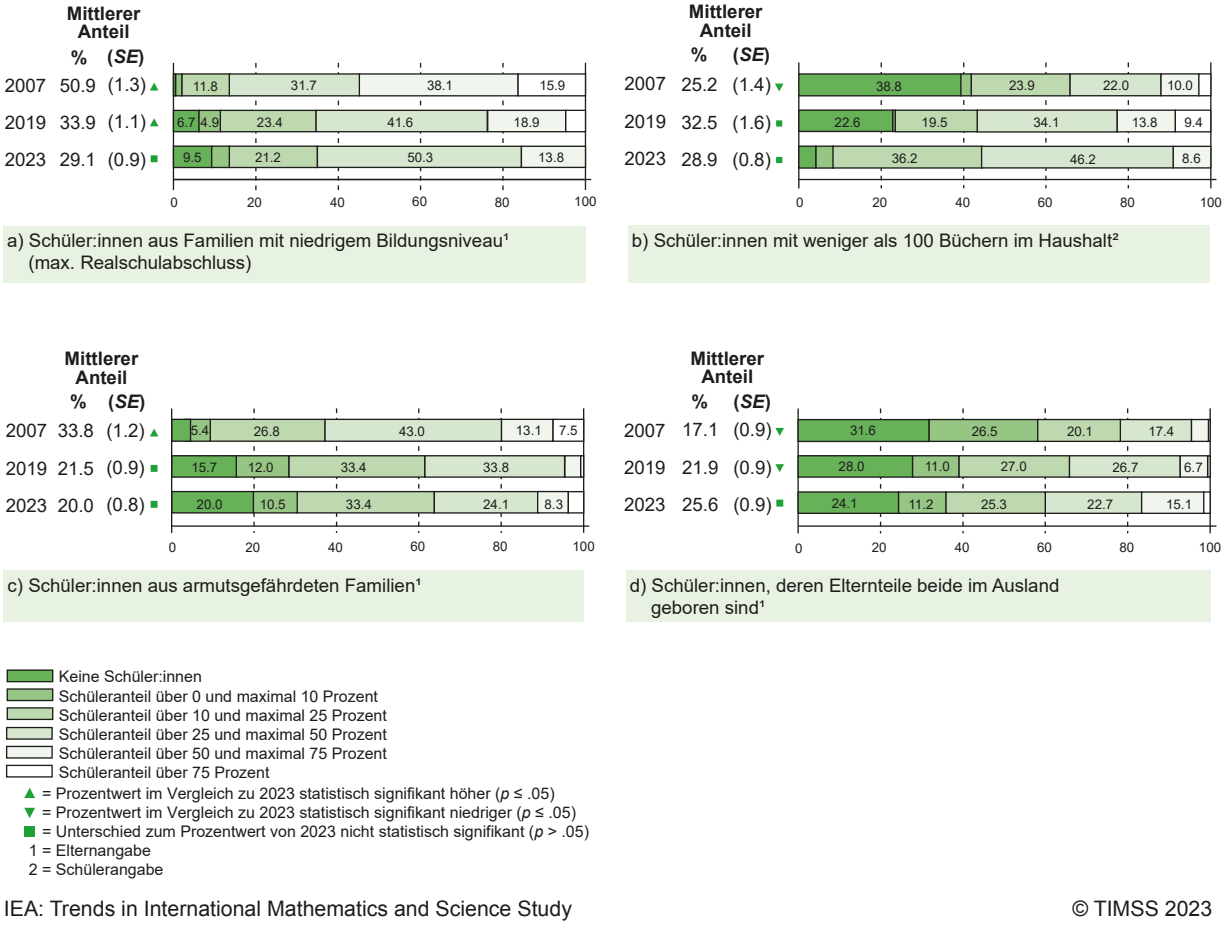


Abbildung 9.7: Klassenanteile von Schüler:innen nach unterschiedlichen Diversitätsmerkmalen in TIMSS 2007, 2019 und 2023 (Schulen nach Zusammensetzungsmerkmalen in Prozent) II



Leistungsstarke Schüler:innen: Abbildung 9.6 ist zu entnehmen, dass auf 38 Prozent der Grundschulen in 2023 keine Schüler:innen gehen, die eine Leistung auf dem höchsten Niveau (Kompetenzstufe V) in Mathematik erreichten. An weiteren 27 Prozent der Schulen liegt der Anteil dieser besonders leistungsstarken Kinder bei weniger als 10 Prozent und an etwa 33 Prozent der Schulen liegt der Anteil dieser Kinder zwischen 10 und 25 Prozent oder darüber. Im Jahr 2007 waren es rund 59 Prozent aller Grundschulen, an denen herausragende Leistungen in Mathematik festgestellt werden konnten, im Jahr 2019 waren es nur 43 Prozent der Schulen und im Jahr 2023 sind es 62 Prozent. Über die Studienzyklen hinweg fällt somit auf, dass es im Vergleich zu TIMSS 2019 vielen Schulen in 2023 wieder besser gelingt, leistungsstarke Kinder so zu fördern, dass sie herausragende Leistungen erzielen. Gleichzeitig ist Abbildung 9.6 auch zu entnehmen, dass der Anteil von Grundschulen, an denen mehr als 10 Prozent der Schüler:innen sehr gute Leistungen zeigen, leicht steigt. Für die Naturwissenschaften ist ebenfalls festzustellen, dass es in 2023 im Vergleich zu 2019 wieder mehr Grundschulen gibt, an denen Schüler:innen herausragende Leistungen auf der Kompetenzstufe V zeigen. Der Anteil an Schulen, bei denen mindestens 10 Prozent der Schüler:innen Leistungen auf der Kompetenzstufe V zeigen, steigt im selben Zeitraum von ca. 22 Prozent auf etwa 34 Prozent.

Leistungsschwache Schüler:innen: Für 2023 ist zunächst für Mathematik festzustellen, dass sich in 8 Prozent der Schulen keine Kinder mit Mathematikleis-

tungen unterhalb der Kompetenzstufe III finden. In weiteren 12.3 Prozent aller Grundschulen beträgt ihr Anteil maximal 10 Prozent, in 37 Prozent der Schulen liegt der Anteil dieser Kinder zwischen 10 und 25 Prozent und in weiteren 31 Prozent der Schulen liegt der Anteil der leistungsschwachen Schüler:innen zwischen 25 und 50 Prozent. In den verbleibenden 11 Prozent der Grundschulen kommt mindestens die Hälfte der Viertklässler:innen nicht über die beiden unteren Kompetenzstufen hinaus. Im Vergleich zu 2007 ist der Anteil von Schulen ohne oder mit sehr wenigen leistungsschwachen Kindern in Mathematik (weniger als 10 %) von knapp 30 Prozent aller Grundschulen auf knapp 25 Prozent (2019) bzw. rund 20 Prozent (2023) gesunken. Der Anteil an Klassen mit großen Anteilen leistungsschwacher Schüler:innen hat sich nur geringfügig verändert.

Für die naturwissenschaftlichen Leistungen zeigt sich, dass in TIMSS 2023 in 6 Prozent der Schulen kein Kind niedrige naturwissenschaftliche Kompetenzen unterhalb der Kompetenzstufe III zeigt. In weiteren 9 Prozent aller Grundschulen beträgt ihr Anteil maximal 10 Prozent, in 40 Prozent der Schulen zwischen 10 und 25 Prozent. In weiteren etwa 32 Prozent der Schulen liegt der Anteil der Kinder zwischen 25 und 50 Prozent. In etwas mehr als 13 Prozent aller Grundschulen kommt mindestens die Hälfte der Viertklässler:innen über die beiden unteren Kompetenzstufen nicht hinaus. Im Vergleich zu 2007 und 2019 hat sich die Zusammensetzung in 2023 insofern etwas weiter verschoben, als dass zunehmend mehr Schulen mehr Viertklässler:innen mit geringen Leistungen in den Naturwissenschaften aufweisen.

Kinder mit diagnostiziertem sonderpädagogischen Förderbedarf: In 2023 finden sich in mehr als der Hälfte (52 %) der Schulen keine Kinder mit diagnostiziertem sonderpädagogischen Förderbedarf. In 2007 lag dieser Anteil noch bei 77 Prozent, in 2019 bei 59 Prozent. Aufgrund der Stichprobenziehung (nur eine Klasse pro Grundschule) und besonderen Studienteilnahmebedingungen ist zu vermuten, dass der in Abbildung 9.6 berichtete Anteil etwas unterschätzt wird. An vielen Schulen werden zudem auch Kinder sonderpädagogisch gefördert, ohne dass für sie ein diagnostizierter Förderbedarf vorliegt. Für TIMSS 2015 konnte festgestellt werden, dass etwa 12 Prozent aller Grundschüler:innen sonderpädagogisch gefördert wurden, obwohl nur für knapp 6 Prozent ein diagnostizierter Förderbedarf vorlag (Wendt et al., 2016).

Differenziert nach Merkmalen der sozialen Herkunft der Schüler:innen (siehe Kapitel 7 in diesem Band) zeigt sich, dass Kinder aus Familien mit niedrigem Bildungsniveau in fast allen Klassen in Deutschland vertreten sind (siehe Abbildung 9.7). Im Vergleich zu 2007 hat sich der Anteil an Klassen, in denen kein Kind in einem Haushalt mit lediglich bis zu 100 Büchern lebt, von knapp 39 Prozent auf rund 10 Prozent reduziert. Gleichsam hat sich der Anteil an Klassen, in denen mindestens 50, aber maximal 75 Prozent der Kinder in einem Haushalt mit maximal 100 Büchern leben, von rund 10 Prozent auf 50 Prozent erhöht. Damit hat sich der schon 2019 beobachtete Trend verschärft: In immer mehr Klassen stammen immer mehr Schüler:innen aus Haushalten mit maximal 100 Büchern. Substanzielle Anteile von Kindern aus armutsgefährdeten Familien – d. h. Anteile im Umfang von mehr als 10 Prozent – gibt es in 70 Prozent aller Klassen. Hier zeigt sich im Vergleich zu 2007 eine günstigere Verteilung auf die Grundschulen. An 24 Prozent aller Grundschulen sind 2023 keine Kinder mit Migrationshintergrund (beide Elternteile im Ausland geboren). An weiteren 11 Prozent der Grundschulen macht ihr Anteil höchstens 10 Prozent aus. In 25 Prozent aller Grundschulen liegt ihr Anteil zwischen 10 und 25 Prozent der Schülerschaft. In rund 23 Prozent der Grundschulen in Deutschland hat etwa jedes zweite bis

vierte Kind Eltern, die nicht in Deutschland geboren wurden. In 15 Prozent der Grundschulen hat mehr als jedes zweite Kind einen Migrationshintergrund. Dieser Anteil lag 2019 noch bei 7 Prozent. Im Trend ist damit eine Zunahme an Schulen zu beobachten, bei denen mehr als 50 Prozent der Viertklässler:innen einen Migrationshintergrund haben.

Die Ergebnisse zeigen deutlich die große Varianz der Schülerzusammensetzungen zwischen Klassen bzw. Schulen in Deutschland. Berücksichtigt man die hier genannten Dimensionen von Heterogenität, so wird deutlich, dass die Herausforderungen im Hinblick auf adaptives Unterrichten je nach Zusammensetzung der Klasse stark variieren können. Forschungsergebnisse sprechen dafür, dass die Komposition der Klasse nicht nur die Leistungen der einzelnen Schüler:innen beeinflusst, sondern auch Auswirkungen auf das Selbstkonzept, die Motivation und die Lernfreude hat (Dumont et al., 2013; Lintorf et al., 2017).

Auch für den Grundschulbereich liegen mittlerweile einige Studien zu Zusammenhängen zwischen Merkmalen der Klassenzusammensetzung und den mathematischen und naturwissenschaftlichen Kompetenzen am Ende der vierten Jahrgangsstufe vor (für einen Überblick: Vennemann, 2018). Aufgrund des querschnittlichen Designs besitzen viele Studien allerdings nur begrenzte Aussagekraft, können aber bei mehrebenenanalytischer Modellierung bedeutsame Zusammenhänge zwischen unterschiedlichen Merkmalen der Komposition und Leistung feststellen. Vor diesem Hintergrund lassen sich die Ergebnisse der Modelle 6 aus den Tabellen 9.7 und 9.8 interpretieren. In den Modellen 6 wurden zusätzlich zu den Merkmalen aus den Modellen 5 auf Schulebene Merkmale der Schülerzusammensetzung mitaufgenommen. Es zeigt sich, dass der Anteil an Kindern mit besonderen Unterstützungsbedarfen sowie der Anteil an Kindern mit einem in Deutschland und einem im Ausland geborenen Elternteil keinen Zusammenhang mit den durchschnittlichen mathematischen oder naturwissenschaftlichen Kompetenzen in den Klassen aufweisen. Hingegen zeigen sich für die mittlere soziale und die migrationsspezifische Zusammensetzung signifikante Zusammenhänge.

Für die Interpretation ist es wichtig zu berücksichtigen, dass hier Klassen mit minimalen (also keinen Kindern mit Migrationshintergrund) und maximalen Anteilen (100 % der Kinder haben Eltern, die beide nicht in Deutschland geboren wurden) verglichen werden. In Klassen ohne Kinder mit Migrationshintergrund im Vergleich zu denen mit maximalen Anteilen zeigen Kinder im Durchschnitt – unabhängig von ihren Individualmerkmalen – etwas bessere Leistungen in der Größenordnung von 15 Punkten bzw. 18 Punkten. Längsschnittuntersuchungen, die für den Grundschulbereich vorliegen und zusätzlich für die Betrachtung von Kompositionseffekten auch Prozessmerkmale (z. B. Unterrichtsmerkmale) berücksichtigen (van Ophuysen & Wendt, 2009; Vennemann, 2018), zeigen, dass die Leistungsunterschiede zwischen Schüler:innen zu großen Anteilen durch Prozessmerkmale erklärt werden können. Die Befunde verdeutlichen damit, dass es vor allem auf einen angemessenen Umgang mit Heterogenität ankommt (Morris-Lange et al., 2013).

9.6 Zusammenfassung

Die Schülerschaft an Grundschulen in Deutschland ist im Jahr 2023 etwas mehr von Vielfalt geprägt, als dies noch im Jahr 2007 der Fall war. In Folge struktureller Reformbemühungen, inklusive Beschulungsangebote zu schaffen, ist ein Anstieg des Anteils von Kindern mit sonderpädagogischen Förderbedarfen an Regelschulen festzustellen. Auch hat sich die Schülerschaft hinsichtlich ihrer soziodemografischen Zusammensetzung, vor allem in Bezug auf die migrationsbezogenen Merkmale der Familien, verändert. In diesem Kapitel wurde anhand von Trendanalysen untersucht, welchen Zusammenhang diese Entwicklung mit den mittleren Leistungen der Schüler:innen in Deutschland von TIMSS 2007 zu 2019 und 2023 aufweist.

In Bezug auf die COVID-19-Pandemie zeigte sich, dass die große Mehrheit der in TIMSS getesteten Schüler:innen zeitweise eine vollständige Schulschließung erlebt hat. Als Alternativen zum Präsenzunterricht wurden Wechselunterricht, Distanzunterricht, Distanzlernen (asynchron und eigenständig) und Notbetreuung (mit Unterrichtsinhalten) von den Schulen angeboten. Über die notwendigen Ressourcen einer lernförderlichen häuslichen Lernumgebung (etwa Raum zum ungestörten Lernen oder ein ausreichender Internetzugang) scheinen die meisten Viertklässler:innen dabei die ganze Zeit verfügt zu haben. Ein eigenes Endgerät zum Durchführen des Distanzunterrichts stand hingegen nur in 64 Prozent der Fälle die ganze Zeit zur Verfügung und in 21 Prozent der Fälle nur einen Teil der Zeit. Aus den Daten lässt sich nicht entnehmen, worauf diese nur teilweise Verfügbarkeit des eigenen Endgerätes zurückzuführen ist.

Generell gaben die Erziehungsberechtigten an, dass sie ihre Kinder während der Schulschließungen gut beim Lernen unterstützen konnten. Allerdings berichteten auch immerhin 20 Prozent der Erziehungsberechtigten, dass sie ihre Kinder während der Schulschließungen nicht gut beim Lernen unterstützen konnten. Dieser Anteil liegt für die Zeit nach den Schulschließungen bei nur 10 Prozent, d. h. der hohe Anteil von 20 Prozent ist wahrscheinlich auch auf die Mehrfachbelastung durch die notwendige simultane Unterstützung von mehreren Kindern im Haushalt während der Schulschließungen zurückzuführen. Sowohl aus Sicht der Lehrkräfte als auch aus Sicht der Erziehungsberechtigten hat dabei die Umsetzung des Fernunterrichts eher gut funktioniert. Differenzielle Effekte gab es allerdings hinsichtlich der Einschätzung zum Kontakt zwischen Schüler:innen und Lehrkräften, der Übermittlung von Arbeitsaufträgen und Lernmaterialien, der Rückmeldung zu bearbeiteten Arbeitsaufträgen durch die Lehrkräfte sowie der technischen Umsetzung des Fernunterrichts auf Seiten der Schule bzw. der Lehrkräfte. Hier schätzen die Erziehungsberechtigten die Situation etwas pessimistischer ein als die Lehrkräfte, welche wiederum die technische Umsetzung des Fernunterrichts auf Seiten der Schüler:innen als weniger gut einschätzen als die Erziehungsberechtigten. In Bezug auf die Motivierung der Schüler:innen scheint diese nach eigenen Angaben der Lehrkräfte den Sachunterrichtslehrkräften etwas besser gelungen zu sein als den Mathematiklehrkräften.

Um den Effekt der durch die COVID-19-Pandemie bedingten veränderten Lernumgebung auf die Kompetenzstände der Viertklässler:innen in Mathematik und in den Naturwissenschaften zu analysieren, wurden in Mehrebenenmodellen verschiedene Aspekte der alternativen Lernumgebungen zur Vorhersage der Kompetenzstände genutzt. Dabei zeigt sich, dass nur die Einschätzung der Lehrkräfte darüber, wie gut der Fernunterricht funktioniert hat, die Kompetenzstände positiv vorhersagen konnte. Immerhin erzielten Schüler:innen in Klassen, in denen der

Fernunterricht nach Ansicht der Lehrkräfte gut funktioniert hat, rund 20 Punkte (Mathematik) bzw. 16 Punkte (Naturwissenschaften) mehr als Schüler:innen, bei denen der Fernunterricht nach Ansicht der Lehrkräfte nicht gut funktioniert hat, wobei ein Unterschied von rund 20 Punkten etwa dem Lernfortschritt eines Drittel Schuljahres entspricht (Wendt, Kasper et al., 2017).

Werden die Leistungsergebnisse im Trend und insbesondere im Vergleich der Resultate von TIMSS 2007 und 2023 betrachtet, so ergibt sich für Deutschland im Bereich Mathematik nach Kontrolle der veränderten Schülerschaft ein kleiner negativer Trend, der aber statistisch nicht signifikant ist. Für die Naturwissenschaften ließ sich in TIMSS 2019 erstmals ein um 10 Punkte signifikant niedrigerer Leistungsmittelwert bei Trendanalysen feststellen (Kasper et al., 2020). Für TIMSS 2023 betrug dieser negative Trend wiederum etwa 10 Punkte und konnte nicht mehr, wie noch in TIMSS 2019, durch komplexere statistische Modellierung, die die veränderte Schülerschaft berücksichtigt, erklärt werden. Die Analysen in diesem Kapitel zeigen, dass Veränderungen in der Schülerzusammensetzung an Grundschulen nicht als Erklärungsmuster für den Leistungsrückgang zwischen 2007 und 2023 herangezogenen werden können.

Die Analyse der Komposition der Schüler:innen an Grundschulen über die Zeit zeigt, dass sich diesbezüglich an jeder fünften Grundschule in Deutschland Veränderungen ergeben haben, die von praktischer Relevanz sind. Um zu verhindern, dass sich die aktuellen Ergebnisse tatsächlich zu noch stärkeren negativen Trends entwickeln, ist Handlungsbedarf geboten. Insbesondere jene Lehrkräfte an Grundschulen, denen es nicht gelingt, zumindest einer deutlichen Mehrheit ihrer Schüler:innen mathematische und naturwissenschaftliche Kompetenzen auf dem Niveau der Kompetenzstufe III zu vermitteln, sollten durch zusätzliche Ressourcen wie etwa Personal oder Reduzierung von Mehrbelastungen unterstützt werden. Dies betrifft in etwa jede vierte Grundschule in Deutschland. Auch wenn es geboten ist, Schüler:innen mit ungünstigen Voraussetzungen gezielt zu fördern, sind dabei die leistungsstarken Schüler:innen nicht zu vernachlässigen. Auffällig ist vor allem der Befund, dass es 2023 im Vergleich zu 2007 in immer weniger Schulen gelingt, gute Schüler:innen so zu fördern, dass sie auch herausragende Leistungen auf dem Niveau der Kompetenzstufe V in den Naturwissenschaften zeigen können. Waren Kinder mit sehr guten Leistungen in den Naturwissenschaften in 2007 noch an 70 Prozent aller Grundschulen zu finden, sind es 2023 nur noch knapp 60 Prozent der Grundschulen. Ein kompetenter Umgang mit Heterogenität ist leicht zu postulieren, aber nur schwer in der Praxis umzusetzen. Die in diesem Kapitel berichteten Ergebnisse zeigen deutlich, dass Herausforderungen im Hinblick auf adaptives Unterrichten je nach Zusammensetzung der Klasse stark variieren.

Einschränkend gilt, dass in diesem Kapitel lediglich einige zentrale Veränderungen in der Zusammensetzung der Schülerschaft betrachtet wurden. Damit bleibt unklar, inwiefern die beobachteten Trendverläufe auch für Gruppen gelten, die hier nicht abgebildet oder differenziert erfasst wurden. Beachtenswert ist, dass sich bedeutsame Leistungstrends vor allem für jene Gruppen beobachten lassen, deren Charakteristika immer schon sehr heterogen waren und sich im Zuge gesellschaftlicher Wandlungsprozesse deutlich verändert haben (wie beispielsweise Schüler:innen mit einem in Deutschland und einem im Ausland geborenen Elternteil). Die Frage, ob die Nutzung von übergeordneten kategorialen Unterscheidungen von Kindern in bestimmte Teilgruppen (beispielsweise mit und ohne Migrationshintergrund) auch im Trend sinnvoll ist, wäre in vertiefenden Analysen zu untersuchen. Auch muss bedacht werden, dass in den vorgestellten Modellen

die Veränderung in der Zusammensetzung der Schülerschaft in der Grundgesamtheit über die Zeit untersucht wurde. Denkbar ist allerdings auch, dass sich diese Veränderung unterschiedlich stark in der Schülerschaft einzelner Schulen dargestellt hat, Schulen also in diesem Sinne differenzielle Wandlungsprozesse erfahren haben (z. B. unterschiedliche Entwicklungen in städtischen oder ländlichen Gebieten). Ob diese Wandlungsprozesse auch zu heterogenen Leistungsgradienten beigetragen haben, wäre ebenfalls in vertiefenden Analysen zu untersuchen. Es ist ferner zu bedenken, dass es sich bei dem ausgewiesenen COVID-19-Effekt von etwa 20 bzw. 16 Punkten um durchschnittliche Klasseneffekte handelt, d. h., nicht alle Schüler:innen der betroffenen Klassen wiesen Leistungseinbußen in dieser Größenordnung auf.

Literatur

Autorengruppe Bildungsberichterstattung. (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*. wbv. https://doi.org/10.3278/6001820ew

Autorengruppe Bildungsberichterstattung. (2020). *Bildung in Deutschland 2020. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt*. wbv. https://doi.org/10.3278/6001820gw

Autor:innengruppe Bildungsberichterstattung. (2024). *Bildung in Deutschland 2024. Ein indikatorengestützter Bericht mit einer Analyse zu beruflicher Bildung*. wbv. https://dx.doi.org/10.3278/6001820iw

Bonsen, M., Bos, W., Gröhlich, C., Harney, B., Imhäuser, K., Makles, A., Schräpler, J., Terpoorten, T., Weishaupt, H., & Wendt, H. (Hrsg.). (2010). *Zur Konstruktion von Sozialindizes. Ein Beitrag zur Analyse sozialräumlicher Benachteiligung von Schulen als Voraussetzung für qualitative Schulentwicklung*. BMBF.

Bremerich-Vos, A., Wendt, H., & Bos, W. (2017). Lesekompetenzen im internationalen Vergleich: Testkonzeption und Ergebnisse. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 79–143). Waxmann.

Bremerich-Vos, A., Wendt, H., & Hußmann, A. (2017). Bausteine adaptiven Leseunterrichts angesichts gewachsener Heterogenität. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 317–330). Waxmann.

Brühwiler, C. (2014). *Adaptive Lehrkompetenz und schulisches Lernen*. Waxmann.

Dumont, H., Neumann, N., Maaz, K., & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen: Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, *60*(3), 163–183. https://doi.org/10.2378/peu2013.art14d

Eickelmann, B., Bos, W., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., Senkbeil, M., & Vahrenhold, J. (Hrsg.). (2019). *ICILS 2018 #Deutschland. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking*. Waxmann. https://doi.org/10.31244/9783830993278

Frey, A., Ludewig, U., König, C., Krampen, D., Lorenz, R., & Bos, W. (2023). Lesekompetenz von Viertklässlerinnen und Viertklässlern im internationalen Vergleich: 20-Jahre Trend. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher & T. C. Stubbe (Hrsg.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (S. 111–130). Waxmann. https://doi.org/10.31244/9783830997009

Hasler-Waters, L., Menchaca, M. P., & Borup, J. (2014). Parental involvement in K-12 online and blended learning. In R. E. Ferdig & K. Kennedy (Hrsg.), *Handbook of research on K-12 online and blended learning* (S. 303–324). ETC Press.

Huber, S. G., & Helm, C. (2020). Lernen in Zeiten der Corona-Pandemie. Die Rolle familiärer Merkmale für das Lernen von Schüler*innen. Befunde vom Schul-Barometer in Deutschland, Österreich und der Schweiz. In D. Fickermann & B. Edelstein (Hrsg.), *„Langsam vermisse ich die Schule …" Schule während und nach der Corona-Pandemie* (S. 37–60). Waxmann. https://doi.org/10.31244/9783830992318.01

Kasper, D., Wendt, H., Bos, W., & Köller, O. (2016). Trends in mathematischen und naturwissenschaftlichen Kompetenzen am Ende der Grundschulzeit in Deutschland. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 367–382). Waxmann. https://doi.org/10.25656/01:14022

Kasper, D., Wendt, H., Schwippert, K., & Köller, O. (2020). Trends in der Schülerzusammensetzung und mathematischen und naturwissenschaftlichen Kompetenzen. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 331–349). Waxmann. https://doi.org/10.31244/9783830993193

Lintorf, K., Guill, K., & Wendt, H. (2017). ‚Was dein Kind kann, kann meins schon lange!'. Effekte der sozialen Komposition von Grundschulklassen auf den Übergang. *Zeitschrift für Erziehungswissenschaft*, *20*(4), 711–727. https://doi.org/10.1007/s11618-017-0762-2

McCulloch, C. E., Searle, S. R., & Neuhaus, J. M. (2008). *Generalized, linear, and mixed models*. Wiley.

Morris-Lange, S., Wendt, H., & Wohlfarth, C. (2013). *Segregation an deutschen Schulen. Ausmaß, Folgen und Handlungsempfehlungen für bessere Bildungschancen*. SVR GmbH.

OECD (Organisation for Economic Cooperation and Development). (2021). The state of global education: 18 months into the pandemic. https://www.oecd-ilibrary.org/deliver/1a23bb23-en.pdf?itemId=%2Fcontent%2Fpublication%2F1a23bb23-en&mimeType=pdf

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods*. Sage Publication.

Schneider, R., Enke, F., Sachse, K. A., & Schipolowski, S. (2022). Lernbedingungen im Fern- und Wechselunterricht während der Coronavirus-Pandemie. In P. Stanat, S. Schipolowski, R. Schneider, K. A. Sachse, S. Weirich & S. Henschel (Hrsg.), *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich* (S. 115–126). Waxmann. https://doi.org/10.31244/9783830996064

Statistisches Bundesamt. (2017). *Bildung und Kultur. Allgemeinbildende Schulen.* https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00033325/2110100177004_korr08122017.pdf

Tomasik, M. J., Helbling, L. A., & Moser, U. (2020). Educational gains of in-person vs. distance learning in primary and secondary schools: A natural experiment during the COVID-19 pandemic school closures in Switzerland. *International Journal of Psychology*, *56*(4), 566–576. https://doi.org/10.1002/ijop.12728

van Ophuysen, S., & Wendt, H. (2009). Zur Veränderung der Mathematikleistung von Klasse 4 bis 6. Welchen Einfluss haben Kompositions- und Unterrichtsmerkmale? *Zeitschrift für Erziehungswissenschaft*, *12*, 302–327.

Vennemann, M. (2018). *Individual- und Kompositionseffekte und der Kompetenzzuwachs in Mathematik und Naturwissenschaft am Ende der Grundschule*. Springer. https://doi.org/10.1007/978-3-658-23506-2

Wendt, H., Steinmayr, R., & Kasper, D. (2016). Geschlechterunterschiede in mathematischen und naturwissenschaftlichen Kompetenzen. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 257–297). Waxmann.

Wendt, H., Bos, W., Goy, M., & Jusufi, D. (2017). *TIMSS 2015 – Skalenhandbuch zur Dokumentation der Erhebungsinstrumente und Arbeit mit den Datensätzen.* Waxmann.

Wendt, H., Kasper, D., Bos, W., Vennemann, M., & Goy, M. (2017). Wie viele Punkte auf der TIMSS-Metrik entsprechen einem Lernjahr? Leistungszuwächse in Mathematik und Naturwissenschaften am Ende der Grundschulzeit. In T. Eckert & B. Gniewosz (Hrsg.), *Bildungsgerechtigkeit* (S. 121–152). Springer. https://doi.org/10.1007/978-3-658-15003-7_8

Wößmann, L., Freundl, V., Grewenig, E., Lergetporer, P., Werner, K., & Zierow, L. (2020). Bildung in der Coronakrise: Wie haben die Schulkinder die Zeit der Schulschließungen verbracht, und welche Bildungsmaßnahmen befürworten die Deutschen? *ifo Schnelldienst, 73*(9), 25–39. https://www.ifo.de/DocDL/sd-2020-09-woessmann-etal-bildungsbarometer-corona.pdf

Kapitel 10
Unterrichten in der Grundschule: Unterrichtsqualität und Lehrkräftefortbildung

Luise A. Scholz, Christin Beese, Armin Jentsch, Jakob Sierau und Mirjam Steffensky

10.1 Einleitung

Unterricht zählt zu den einflussreichsten Faktoren für die leistungsbezogene sowie motivationale Entwicklung von Schüler:innen in der Schule (Hattie, 2009; Helmke & Weinert, 2022; Kunter, 2016; Seidel & Shavelson, 2007; Wang et al., 1993). Insbesondere die Qualität von Unterricht steht dabei seit Langem im Fokus der nationalen wie internationalen Forschung, die wiederholt deren Bedeutung untermauern konnte. So zeigten auch die letzten beiden Zyklen der *Trends in International Mathematics and Science Study* (TIMSS) aus den Jahren 2015 und 2019, dass verschiedene Facetten der Unterrichtsqualität bedeutsam mit den Leistungen in Mathematik und Naturwissenschaften zusammenhängen (Rieser et al., 2016; Stang et al., 2020). Wie Unterricht in Deutschland gestaltet ist, hängt neben curricularen Vorgaben und Schulkonzepten hauptsächlich von den unterrichtenden Lehrkräften ab. Inwiefern Lehrkräfte bereit sind und Gelegenheit bekommen, ihren Unterricht zu reflektieren, weiterzuentwickeln und sich über neue Konzepte oder Merkmale guten Unterrichts beispielsweise im Rahmen von Fortbildungen zu informieren, trägt zur Qualität und Aktualität von Unterricht bei (Richter, 2023). In TIMSS 2019 zeigte sich, dass die Teilnahme von Lehrkräften an Fortbildungen in Deutschland zwar regelmäßig ist, die Häufigkeit der Teilnahme jedoch unter dem europäischen Durchschnitt liegt. Gründe für die geringere Teilnahme konnten dabei wegen fehlender Datengrundlage weder in TIMSS 2015 noch in TIMSS 2019 ausgemacht werden. Um Hinweise auf mögliche Ursachen zu gewinnen, wurde dieser Aspekt in TIMSS 2023 bei der Konzeption der Fragebögen berücksichtigt, sodass in diesem Kapitel erstmals mögliche Gründe untersucht werden können.

Im Folgenden wird zunächst die Unterrichtsqualität sowohl im internationalen Vergleich als auch in Deutschland betrachtet. Anschließend wird sich mit dem Thema der Lehrkräftefortbildung mit dem Fokus auf internationale Vergleiche

hinsichtlich der Teilnahme befasst. Darüber hinaus wird für Deutschland berichtet, welche Fortbildungsangebote Lehrkräften bekannt sind und welche zukünftigen Bedarfe sie angeben.

10.2 Unterrichtsqualität

Unter Unterrichtsqualität wird die Güte von unterrichtlichen Lehr- und Lernumgebungen, die das Erreichen von Unterrichtszielen unterstützen, verstanden (Ditton, 2002; Senden et al., 2022; Weinert et al., 1989).[1] Die verschiedenen Merkmale, die zur Beschreibung von Unterricht genutzt werden können, lassen sich in sogenannte Sicht- und Tiefenstrukturen unterscheiden. Während Sichtstrukturen oberflächliche Merkmale wie Unterrichtsorganisation oder Methodenauswahl betreffen, werden tiefergehende Facetten wie Verstehensprozesse oder Interaktionen als Tiefenstrukturen beschrieben. Diese haben sich als besonders bedeutsam für die Erreichung von Unterrichtszielen herausgestellt (Kunter & Trautwein, 2013). In den letzten Jahrzehnten wurden international bereits verschiedene Modelle der Unterrichtsqualität, wie beispielsweise das Modell der *Quality of Instruction*, *Appropriateness*, *Incentives* und *Time* (QuAIT; Slavin, 2000) oder das *Classroom Assessment Scoring System* (CLASS; Pianta & Hamre, 2009), entwickelt. In der deutschsprachigen Forschung hat sich vor allem das Modell der Basisdimensionen der Unterrichtsqualität nach Klieme et al. (2001) durchgesetzt, das drei zentrale Tiefenstrukturen von Unterricht identifiziert. Das Modell wurde auf Basis der TIMSS-Videostudie von 1995 entwickelt, bei der in Deutschland und anderen Staaten Mathematikunterricht videografiert wurde, und seitdem in verschiedenen Studien im deutschsprachigen Raum genutzt. Als generisches Modell wurde es mittlerweile auch auf andere Fächer übertragen (Fauth et al., 2014; Kleickmann et al., 2020; Praetorius et al., 2018).

10.2.1 Basisdimensionen der Unterrichtsqualität

Die Basisdimensionen dieses Modells sind *Klassenführung*, *kognitive Aktivierung* und *konstruktive Unterstützung*. *Klassenführung* (im Englischen *classroom management;* Kounin, 1976) beschreibt Handlungen der Lehrkraft, die das Geschehen im Klassenzimmer strukturieren und ordnen sowie sicherstellen, dass die Lernzeit im Unterricht effektiv und störungsfrei genutzt werden kann. Dies kann beispielsweise durch klare Regeln und Strukturen, das Vorbeugen von Disziplinproblemen oder die Implementation von Routinen erreicht werden. *Klassenführung* ist somit eine wichtige Grundvoraussetzung für erfolgreichen Unterricht und eine intensive Auseinandersetzung mit den fachlichen Inhalten. Sie trägt also dazu bei, dass die anderen beiden Basisdimensionen im Unterricht umgesetzt werden können (Helmke & Schrader, 2008; Kunter & Trautwein, 2013). Die zweite Basisdimension *kognitive Aktivierung* beschreibt das Potenzial von Unterricht, Schüler:innen zum vertieften Nachdenken und zu einer elaborierten Auseinandersetzung mit dem Unterrichtsgegenstand anzuregen. Merkmale von kognitiv aktivierendem Unterricht sind beispielsweise das Formulieren von komplexen Aufgaben, das Anknüpfen an das Vorwissen der Schüler:innen oder das

1 Ausführlichere Darstellungen sowie multivariate Analysen zur Unterrichtsqualität finden sich im Berichtsband von TIMSS 2019, da Unterrichtsqualität 2019 als Schwerpunktthema in einem eigenen Kapitel (Stang et al., 2020) behandelt wurde.

Reflektieren von Lernprozessen und Lösungsvorschlägen durch die Schüler:innen (Lipowsky, 2020; Lipowsky et al., 2009; Steffensky & Neuhaus, 2018). Die dritte Basisdimension *konstruktive Unterstützung* umfasst Maßnahmen, die das Lernen schülerorientiert unterstützen und begleiten sollen. Die Unterstützung zielt zum einen auf eine schülergerechte Strukturierung von Inhalten (z. B. das Unterteilen von Aufgaben in einzelne, dem Wissensstand der Schüler:innen entsprechende Schritte) sowie Feedback und zum anderen auf Formen der emotionalen Unterstützung ab, wie eine positive Beziehung zwischen den Lehrkräften und Schüler:innen oder eine positive Fehlerkultur (Sliwka et al., 2019; Steffensky & Neuhaus, 2018).

Der Zusammenhang der Basisdimensionen von Unterrichtsqualität mit den Leistungen von Schüler:innen und ihren motivationalen Merkmalen konnte bereits in mehreren Untersuchungen und für verschiedene Fächer und Altersklassen belegt werden (im Überblick Praetorius et al., 2018). Die Befunde sind dabei zwar teils inkonsistent, es deutet sich jedoch an, dass die *Klassenführung* sowohl für die Leistungen als auch die motivationalen Merkmale der Schüler:innen, die *kognitive Aktivierung* vor allem für die Leistungen und die *konstruktive Unterstützung* vor allem für die motivationalen Merkmale bedeutsam zu sein scheinen (Henschel et al., 2023; Kunter et al., 2011; Lipowsky et al., 2009; Schiepe-Tiska et al., 2016). In TIMSS 2019 fanden Stang et al. (2020) mithilfe von Mehrebenenanalysen signifikante Effekte für *kognitive Aktivierung* und *konstruktive Unterstützung* auf die Leistungen in Mathematik und den Naturwissenschaften auf Schülerebene, wobei unter Kontrolle von verschiedenen Hintergrundmerkmalen nur der signifikante Zusammenhang zwischen *konstruktiver Unterstützung* und der Mathematikleistung bestehen blieb. Aus methodischen Gründen war die Dimension *Klassenführung* nicht Teil der Analysen (für weitere Informationen dazu siehe Stang et al., 2020). In TIMSS 2015 konnten Rieser et al. (2016) darüber hinaus zeigen, dass sich die Wahrnehmung der Basisdimensionen zwischen Schüler:innen verschiedener Leistungsgruppen unterscheidet.

10.2.2 Erfassung der Basisdimensionen in TIMSS 2023

In TIMSS 2023 wurden die drei Basisdimensionen über drei Skalen im Schülerfragebogen erfasst. Sie spiegeln somit die individuelle Wahrnehmung der Unterrichtsqualitätsdimensionen durch die Schüler:innen wider. Die Skalen *Wahrgenommene Störungen im Mathematikunterricht* bzw. *Wahrgenommene Störungen im Sachunterricht*[2] zur Erfassung einer Facette von *Klassenführung* werden dabei als Teil des internationalen Fragebogens erhoben. Da für den Sachunterricht zuvor ausschließlich eine abweichend formulierte nationale Skala existierte, sind die vorliegenden Ergebnisse zur *Klassenführung* im Sachunterricht nur eingeschränkt mit den Ergebnissen vorheriger TIMSS-Zyklen vergleichbar. Die internationale Skala zur *Klassenführung* umfasst für beide Domänen sechs Items (z. B. *Die Schülerinnen und Schüler halten sich nicht an die Klassenregeln.*), die auf einer vierstufigen Likertskala (1 = *in jeder oder fast jeder Stunde* bis 4 = *nie*) zu beantworten sind. Für die in diesem Kapitel berichteten nationalen Ergebnisse werden aufgrund geringerer inhaltlicher Passung nur fünf der sechs Items genutzt.

2 In diesem Kapitel wird aus Gründen der Einheitlichkeit von *Sachunterricht* und *Sachunterrichtslehrkräften* gesprochen, auch wenn es im internationalen Kontext keinen Sachunterricht, sondern naturwissenschaftlichen Unterricht gibt.

Für eine vollständige Erfassung des Modells der Basisdimensionen werden seit TIMSS 2015 zusätzlich die *kognitive Aktivierung* und die *konstruktive Unterstützung* im nationalen Teil des Schülerfragebogens über die Skalen *Wahrgenommene Unterstützung* und *Kognitive Aktivierung* erhoben. Die *kognitive Aktivierung* wird seitdem in Mathematik über sieben Items erfasst (z. B. *Unsere Mathematiklehrerin/unser Mathematiklehrer stellt Fragen, über die ich sehr gründlich nachdenken muss*.), in den Naturwissenschaften wurden in TIMSS 2023 zwei zusätzliche Items ergänzt, sodass die Skala neun Items enthält. Die *konstruktive Unterstützung* wird für beide Domänen über fünf Items erhoben (z. B. *Unsere Mathematiklehrkraft ist auch dann nett zu mir, wenn ich einen Fehler mache.*), wobei der Fokus vor allem auf der emotionalen Unterstützung der Schüler:innen liegt. Beide Basisdimensionen werden im Fragebogen ebenfalls auf einer vierstufigen Likertskala (1 = *stimme völlig zu* bis 4 = *stimme überhaupt nicht zu*) abgefragt.

10.2.3 Wahrgenommene Klassenführung im internationalen Vergleich

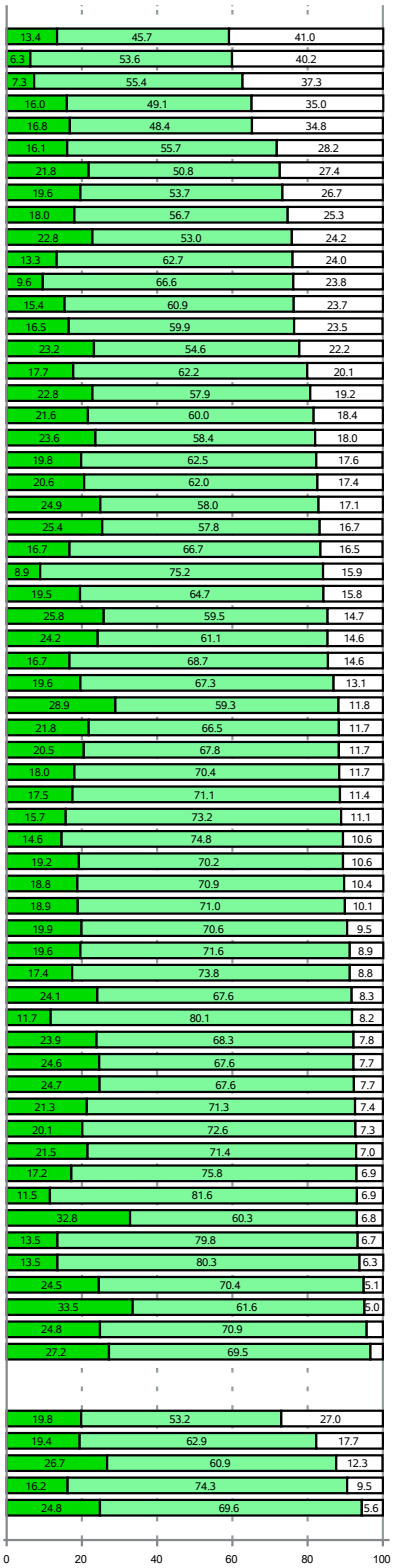
In diesem Abschnitt werden die Ergebnisse zu den von Schüler:innen wahrgenommenen unterrichtlichen Störungen im internationalen Vergleich berichtet. Dazu wurden die Antworten der Schüler:innen in den drei Kategorien *gering*, *mittel* und *hoch* zusammengefasst (für nähere Informationen zur internationalen Skalenbildung siehe von Davier et al., 2024). Dabei bedeutet ein als *gering* oder selten wahrgenommenes Auftreten unterrichtlicher Störungen eine *hoch* ausgeprägte *Klassenführung*. Analog weisen häufige Störungen auf eine weniger effiziente *Klassenführung* hin.

Abbildung 10.1 zeigt die von Schüler:innen in ihrem Mathematikunterricht wahrgenommene *Klassenführung* anhand der unterrichtlichen Störungen im internationalen Vergleich. Etwa drei Viertel der Viertklässler:innen in Deutschland empfinden die unterrichtlichen Störungen als moderat ausgeprägt. Der Anteil der Schüler:innen, die selten bzw. häufig unterrichtliche Störungen wahrnehmen, fällt entsprechend gering aus: etwa 7 Prozent der Schüler:innen empfinden die Störungen im Mathematikunterricht als niedrig ausgeprägt, ca. 17 Prozent nehmen dagegen in nahezu jeder Unterrichtsstunde Störungen wahr. Insgesamt deuten die Befunde also darauf hin, dass Viertklässler:innen in Deutschland die *Klassenführung* in überwiegender Mehrheit als moderat ausgeprägt einschätzen. Ähnliche Angaben zum Mathematikunterricht wie die Viertklässler:innen in Deutschland machen Schüler:innen in der französischen Gemeinschaft in Belgien, in Frankreich, Italien und Québec: Hier geben ebenfalls knapp 20 Prozent häufige oder sehr häufige Störungen an, während ungefähr drei Viertel der Schüler:innen moderate Störungen wahrnehmen. In der Tendenz ist diese Häufigkeitsverteilung international in vielen Staaten zu beobachten. Die Angaben von Schüler:innen in den EU- und OECD-Staaten liegen durchschnittlich bei etwa 18 bzw. 19 Prozent (häufige oder sehr häufige Störungen) und 71 Prozent (moderate Störungen).

Im Unterschied dazu zeigt sich, dass mindestens 35 Prozent der Schüler:innen in Albanien, Armenien, Bulgarien, Japan und Usbekistan ihren Mathematikunterricht als nahezu störungsfrei erleben. Zudem geben weniger als 10 Prozent der Schüler:innen in Bulgarien, Japan und Finnland häufige Störungen im Mathematikunterricht an. Umgekehrt weisen die Angaben von Viertklässler:innen in Brasilien und Chile auf eine weniger effiziente *Klassenführung* hin: So nimmt ein Drittel der Schüler:innen in diesen Staaten in fast jeder Unterrichtsstunde Störungen wahr.

Abbildung 10.1: Wahrgenommene Störungen im Mathematikunterricht nach *hoher*, *mittlerer* und *geringer* Ausprägung (Anteile der Schüler:innen in Prozent)

Teilnehmer [A]	gering %	(SE)	mittel %	(SE)	hoch %	(SE)
[23] *Albanien*	13.4	(1.3)	45.7	(1.7)	41.0	(2.2)
Bulgarien	6.3	(0.8)	53.6	(2.1)	40.2	(2.4)
Japan	7.3	(0.7)	55.4	(1.6)	37.3	(1.8)
[2] *Armenien*	16.0	(0.9)	49.1	(1.3)	35.0	(1.6)
[2] *Usbekistan*	16.8	(0.8)	48.4	(1.3)	34.8	(1.4)
[2] *Kasachstan*	16.1	(0.7)	55.7	(1.0)	28.2	(1.1)
Jordanien	21.8	(1.2)	50.8	(1.1)	27.4	(1.7)
Nordmazedonien	19.6	(1.0)	53.7	(1.0)	26.7	(1.1)
[2] *Kosovo*	18.0	(1.0)	56.7	(1.2)	25.3	(1.1)
[2] *Saudi-Arabien*	22.8	(0.7)	53.0	(0.9)	24.2	(0.9)
Marokko	13.3	(0.7)	62.7	(1.7)	24.0	(1.8)
Republik Korea (Südkorea)	9.6	(0.9)	66.6	(1.4)	23.8	(1.6)
[3] *Hongkong*	15.4	(1.4)	60.9	(1.2)	23.7	(1.3)
Aserbaidschan	16.5	(0.8)	59.9	(1.0)	23.5	(1.2)
Bahrain	23.2	(1.0)	54.6	(0.9)	22.2	(1.1)
[2] *Georgien*	17.7	(0.9)	62.2	(1.2)	20.1	(1.3)
Vereinigte Arabische Emirate (VAE)	22.8	(0.4)	57.9	(0.4)	19.2	(0.4)
[23] *Rumänien*	21.6	(1.6)	60.0	(1.5)	18.4	(1.4)
Katar	23.6	(0.6)	58.4	(0.8)	18.0	(0.7)
Taiwan	19.8	(1.0)	62.5	(1.0)	17.6	(0.9)
[2] *Serbien*	20.6	(1.1)	62.0	(1.3)	17.4	(1.4)
[2] *Montenegro*	24.9	(0.7)	58.0	(0.6)	17.1	(0.6)
Oman	25.4	(0.8)	57.8	(0.8)	16.7	(0.7)
Macau	16.7	(0.3)	66.7	(0.5)	16.5	(0.4)
Finnland	8.9	(0.6)	75.2	(0.7)	15.9	(0.8)
Internationaler Mittelwert	**19.5**	**(1.0)**	**64.7**	**(1.1)**	**15.8**	**(1.1)**
Iran	25.8	(0.9)	59.5	(0.9)	14.7	(0.8)
Kuwait	24.2	(1.2)	61.1	(1.1)	14.6	(1.0)
[2] *Ungarn*	16.7	(1.0)	68.7	(1.1)	14.6	(1.0)
Slowakei	19.6	(1.2)	67.3	(1.5)	13.1	(1.5)
[2] *Bosnien und Herzegowina*	28.9	(1.6)	59.3	(1.3)	11.8	(1.3)
[2] *Zypern*	21.8	(1.0)	66.5	(1.1)	11.7	(1.0)
[123] *Türkei (5. Jgst.)*	20.5	(1.0)	67.8	(0.9)	11.7	(0.8)
[2] *Litauen*	18.0	(1.1)	70.4	(1.0)	11.7	(0.8)
VG EU	**17.5**	**(1.0)**	**71.1**	**(1.1)**	**11.4**	**(1.0)**
Irland	15.7	(1.5)	73.2	(1.4)	11.1	(1.1)
[2] *Schweden*	14.6	(1.2)	74.8	(1.0)	10.6	(0.8)
[2] *Polen*	19.2	(1.2)	70.2	(1.1)	10.6	(0.9)
VG OECD	**18.8**	**(1.0)**	**70.9**	**(1.0)**	**10.4**	**(0.9)**
[2] *Slowenien*	18.9	(0.9)	71.0	(1.0)	10.1	(0.8)
Portugal	19.9	(1.2)	70.6	(1.0)	9.5	(0.7)
[2] *Frankreich*	19.6	(1.0)	71.6	(1.1)	8.9	(0.8)
[2] *Italien*	17.4	(1.0)	73.8	(0.9)	8.8	(0.8)
[1] *Südafrika (5. Jgst.)*	24.1	(0.9)	67.6	(0.8)	8.3	(0.4)
[3] *Niederlande*	11.7	(1.0)	80.1	(0.9)	8.2	(0.8)
[23] *USA*	23.9	(0.7)	68.3	(0.6)	7.8	(0.4)
Lettland	24.6	(1.2)	67.6	(1.2)	7.7	(0.9)
[2] *Tschechien*	24.7	(1.1)	67.6	(1.0)	7.7	(0.6)
[23] *Kanada*	21.3	(0.8)	71.3	(0.7)	7.4	(0.3)
[2] *Belgien (Frz. Gem.)*	20.1	(0.8)	72.6	(0.8)	7.3	(0.5)
[123] *England*	21.5	(1.4)	71.4	(1.2)	7.0	(0.6)
Deutschland	17.2	(0.6)	75.8	(0.7)	6.9	(0.5)
[23] *Dänemark*	11.5	(0.7)	81.6	(0.6)	6.9	(0.6)
[23] *Chile*	32.8	(1.1)	60.3	(1.0)	6.8	(0.6)
[12] *Norwegen (5. Jgst.)*	13.5	(0.7)	79.8	(0.8)	6.7	(0.7)
[23] *Belgien (Fläm. Gem.)*	13.5	(0.9)	80.3	(0.9)	6.3	(0.5)
Australien	24.5	(0.9)	70.4	(0.7)	5.1	(0.4)
[23] *Brasilien*	33.5	(0.7)	61.6	(0.7)	5.0	(0.5)
[2] *Spanien*	24.8	(0.8)	70.9	(0.8)	4.3	(0.3)
[123] *Neuseeland*	27.2	(1.2)	69.5	(1.1)	3.3	(0.3)
Benchmark-Teilnehmer						
Schardscha, VAE	19.8	(0.8)	53.2	(1.0)	27.0	(1.5)
Dubai, VAE	19.4	(0.8)	62.9	(0.9)	17.7	(0.7)
Abu Dhabi, VAE	26.7	(0.6)	60.9	(0.5)	12.3	(0.4)
[23] *Québec, Kanada*	16.2	(1.1)	74.3	(0.8)	9.5	(0.7)
[2] *Ontario, Kanada*	24.8	(1.5)	69.6	(1.4)	5.6	(0.5)



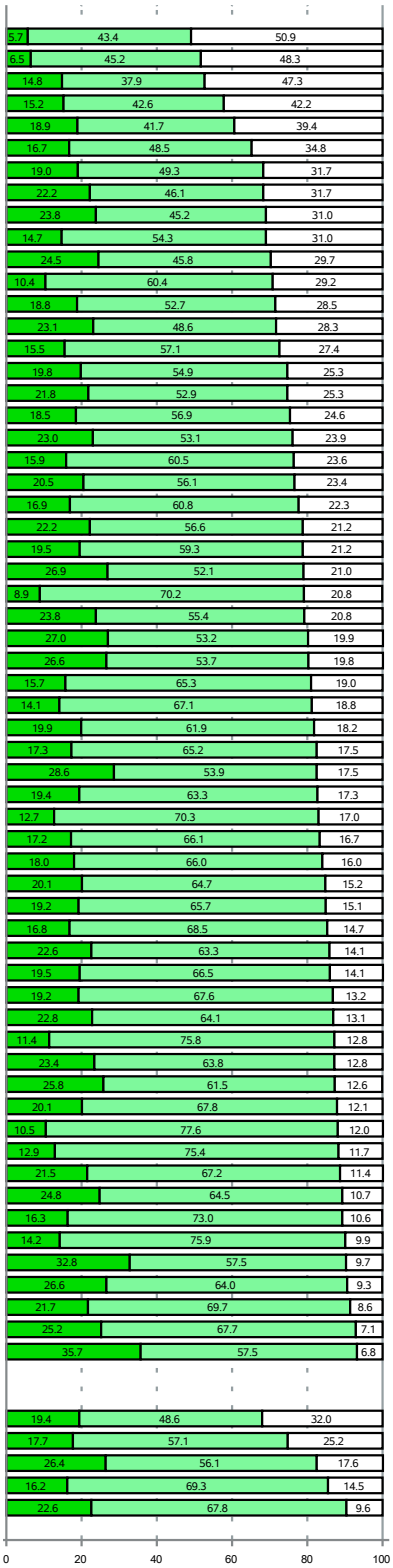
Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen, die die Klassenführung als gering einschätzen
% der Schüler:innen, die die Klassenführung als mittel einschätzen
% der Schüler:innen, die die Klassenführung als hoch einschätzen

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = In Singapur wurde die Klassenführung nicht erhoben. Somit ist Singapur nicht Teil dieser Darstellung und des berichteten internationalen Mittelwerts.

Abbildung 10.2: Wahrgenommene Störungen im naturwissenschaftsbezogenen Sachunterricht nach *hoher*, *mittlerer* und *geringer* Ausprägung (Anteile der Schüler:innen in Prozent)

Teilnehmer [A]	gering %	(SE)	mittel %	(SE)	hoch %	(SE)
Japan	5.7	(0.5)	43.4	(1.5)	50.9	(1.7)
Bulgarien	6.5	(0.8)	45.2	(2.0)	48.3	(2.5)
23 *Albanien*	14.8	(1.5)	37.9	(1.8)	47.3	(2.4)
2 *Armenien*	15.2	(0.9)	42.6	(1.3)	42.2	(1.8)
2 *Usbekistan*	18.9	(0.8)	41.7	(1.1)	39.4	(1.3)
2 *Kasachstan*	16.7	(0.7)	48.5	(1.0)	34.8	(1.1)
2 *Kosovo*	19.0	(0.8)	49.3	(1.3)	31.7	(1.4)
Nordmazedonien	22.2	(1.0)	46.1	(1.1)	31.7	(1.4)
Jordanien	23.8	(1.2)	45.2	(1.4)	31.0	(2.1)
3 *Hongkong*	14.7	(1.2)	54.3	(1.0)	31.0	(1.3)
2 *Saudi-Arabien*	24.5	(0.8)	45.8	(0.8)	29.7	(1.0)
Republik Korea (Südkorea)	10.4	(1.0)	60.4	(1.4)	29.2	(1.7)
Aserbaidschan	18.8	(0.8)	52.7	(1.0)	28.5	(1.2)
Bahrain	23.1	(1.0)	48.6	(1.0)	28.3	(1.2)
Marokko	15.5	(0.8)	57.1	(1.4)	27.4	(1.7)
23 *Rumänien*	19.8	(1.4)	54.9	(1.5)	25.3	(1.7)
Vereinigte Arabische Emirate (VAE)	21.8	(0.4)	52.9	(0.4)	25.3	(0.4)
2 *Georgien*	18.5	(1.0)	56.9	(1.4)	24.6	(1.5)
Katar	23.0	(0.7)	53.1	(0.6)	23.9	(0.6)
Macau	15.9	(0.3)	60.5	(0.5)	23.6	(0.4)
2 *Serbien*	20.5	(1.3)	56.1	(1.3)	23.4	(1.6)
2 *Ungarn*	16.9	(1.0)	60.8	(1.1)	22.3	(1.2)
Taiwan	22.2	(1.0)	56.6	(0.9)	21.2	(1.0)
Internationaler Mittelwert	**19.5**	**(1.0)**	**59.3**	**(1.1)**	**21.2**	**(1.2)**
2 *Montenegro*	26.9	(0.7)	52.1	(0.7)	21.0	(0.7)
Finnland	8.9	(0.5)	70.2	(0.9)	20.8	(1.0)
Kuwait	23.8	(1.2)	55.4	(1.3)	20.8	(1.5)
Oman	27.0	(0.8)	53.2	(0.7)	19.9	(0.7)
Iran	26.6	(1.1)	53.7	(0.9)	19.8	(0.9)
2 *Polen*	15.7	(0.9)	65.3	(1.1)	19.0	(1.2)
Irland	14.1	(1.3)	67.1	(1.3)	18.8	(1.3)
123 *Türkei (5. Jgst.)*	19.9	(1.1)	61.9	(0.9)	18.2	(1.1)
2 *Litauen*	17.3	(1.0)	65.2	(1.0)	17.5	(1.1)
2 *Bosnien und Herzegowina*	28.6	(1.9)	53.9	(1.4)	17.5	(1.8)
Slowakei	19.4	(1.2)	63.3	(1.5)	17.3	(1.7)
2 *Schweden*	12.7	(1.3)	70.3	(1.2)	17.0	(1.2)
VG EU	**17.2**	**(1.0)**	**66.1**	**(1.2)**	**16.7**	**(1.2)**
VG OECD	**18.0**	**(1.0)**	**66.0**	**(1.0)**	**16.0**	**(1.0)**
2 *Frankreich*	20.1	(1.1)	64.7	(1.2)	15.2	(1.0)
Portugal	19.2	(1.1)	65.7	(1.1)	15.1	(1.1)
2 *Italien*	16.8	(1.0)	68.5	(1.1)	14.7	(1.0)
23 *USA*	22.6	(0.7)	63.3	(0.6)	14.1	(0.6)
2 *Slowenien*	19.5	(0.9)	66.5	(1.0)	14.1	(1.0)
123 *England*	19.2	(1.2)	67.6	(1.1)	13.2	(1.1)
Lettland	22.8	(1.1)	64.1	(1.0)	13.1	(1.2)
3 *Niederlande*	11.4	(1.1)	75.8	(1.2)	12.8	(1.1)
2 *Tschechien*	23.4	(1.1)	63.8	(1.0)	12.8	(0.7)
2 *Zypern*	25.8	(1.2)	61.5	(1.3)	12.6	(0.9)
23 *Kanada*	20.1	(0.7)	67.8	(0.5)	12.1	(0.4)
23 *Dänemark*	10.5	(0.6)	77.6	(0.8)	12.0	(0.9)
12 *Norwegen (5. Jgst.)*	12.9	(0.8)	75.4	(1.0)	11.7	(1.0)
2 *Belgien (Frz. Gem.)*	21.5	(0.8)	67.2	(0.8)	11.4	(0.7)
1 *Südafrika (5. Jgst.)*	24.8	(0.7)	64.5	(0.8)	10.7	(0.6)
Deutschland	16.3	(0.7)	73.0	(0.7)	10.6	(0.6)
23 *Belgien (Fläm. Gem.)*	14.2	(0.9)	75.9	(0.9)	9.9	(0.7)
23 *Chile*	32.8	(1.0)	57.5	(0.8)	9.7	(0.7)
123 *Neuseeland*	26.6	(1.1)	64.0	(1.0)	9.3	(0.5)
Australien	21.7	(0.8)	69.7	(0.8)	8.6	(0.5)
2 *Spanien*	25.2	(0.7)	67.7	(0.6)	7.1	(0.5)
23 *Brasilien*	35.7	(0.9)	57.5	(0.8)	6.8	(0.5)
Benchmark-Teilnehmer						
Schardscha, VAE	19.4	(1.2)	48.6	(0.8)	32.0	(1.6)
Dubai, VAE	17.7	(0.7)	57.1	(0.8)	25.2	(0.9)
Abu Dhabi, VAE	26.4	(0.6)	56.1	(0.6)	17.6	(0.6)
23 *Québec, Kanada*	16.2	(0.9)	69.3	(0.7)	14.5	(0.8)
2 *Ontario, Kanada*	22.6	(1.2)	67.8	(0.9)	9.6	(0.6)



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen, die die Klassenführung als gering einschätzen
% der Schüler:innen, die die Klassenführung als mittel einschätzen
% der Schüler:innen, die die Klassenführung als hoch einschätzen

Kursiv gesetzt sind die Teilnehmer, für die von einer eingeschränkten Vergleichbarkeit der Ergebnisse ausgegangen werden muss.
1 = Die nationale Zielpopulation entspricht nicht oder nicht ausschließlich der vierten Jahrgangsstufe.
2 = Der Ausschöpfungsgrad und/oder die Ausschlüsse von der nationalen Zielpopulation erfüllen nicht die internationalen Vorgaben.
3 = Die Teilnahmequoten auf Schul- und/oder Schülerebene erreichen nicht die internationalen Vorgaben.
A = In Singapur wurde die Klassenführung nicht erhoben. Somit ist Singapur nicht Teil dieser Darstellung und des berichteten internationalen Mittelwerts.

 © TIMSS 2023

Die Befunde für den naturwissenschaftsbezogenen Sachunterricht fallen sehr ähnlich aus (siehe Abbildung 10.2): Etwa drei Viertel der Viertklässler:innen in Deutschland (73 %) empfinden die unterrichtlichen Störungen im Sachunterricht als moderat, 11 Prozent als niedrig ausgeprägt und 16 Prozent nehmen in nahezu jeder Unterrichtsstunde Störungen wahr. Sehr ähnliche Angaben zum Sachunterricht machen Schüler:innen in der flämischen Gemeinschaft in Belgien. In den EU- und OECD-Staaten liegen die Durchschnittswerte bei 17 bzw. 18 Prozent (häufige oder sehr häufige Störungen) und 66 Prozent (moderate Störungen).

In Albanien, Armenien, Bulgarien und Japan erleben mehr als 40 Prozent der Viertklässler:innen ihren Sachunterricht als weitestgehend störungsfrei. Weniger als 10 Prozent der Schüler:innen in Bulgarien und Japan geben an, dass Störungen im Sachunterricht häufig vorkommen. Wiederum weisen die Angaben von Viertklässler:innen in Brasilien und Chile auf eine weniger effiziente *Klassenführung* hin: Mehr als 30 Prozent der Schüler:innen in diesen Staaten empfinden ihren Sachunterricht als häufig gestört.

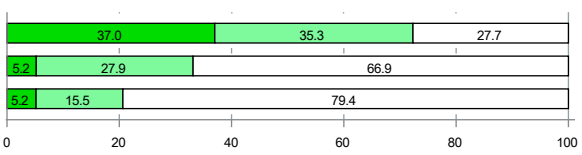
10.2.4 Wahrgenommene Unterrichtsqualität in Deutschland

In Abbildung 10.3 ist die mittlere von Schüler:innen in Deutschland wahrgenommene Qualität ihres Mathematikunterrichts dargestellt. Demnach wird die *Klassenführung* von den Schüler:innen auf der vierstufigen Antwortskala als eher unterdurchschnittlich eingeschätzt. Etwas mehr als ein Drittel stuft das Vorkommen von Unterrichtsstörungen dabei als häufig ein. Ungefähr der gleiche Anteil an Viertklässler:innen nimmt die Häufigkeit von Störungen im Mathematikunterricht als moderat ausgeprägt wahr und etwas mehr als ein Viertel der Schüler:innen schätzt sie als gering ein. Die Ergebnisse für die beiden anderen Basisdimensionen, *kognitive Aktivierung* und *konstruktive Unterstützung*, fallen deutlich positiver aus. Hier attestieren nur ca. 5 Prozent der Schüler:innen ihrem Mathematikunterricht eine geringe Unterrichtsqualität, während der Großteil diese als hoch einschätzt: Etwas mehr als zwei Drittel der Schüler:innen geben an, in ihrem Mathematikunterricht in hohem Maße kognitiv aktiviert zu sein. Knapp 80 Prozent fühlen sich in hohem Maße von ihrer Mathematiklehrkraft unterstützt.

Die nationalen Ergebnisse zur wahrgenommenen Qualität des naturwissenschaftsbezogenen Sachunterrichts sind in Abbildung 10.4 dargestellt. Es zeigt sich insgesamt ein nahezu identisches Bild zu den für den Mathematikunterricht berichteten Befunden: Auch in dieser Domäne wird die *Klassenführung*

Abbildung 10.3: Wahrgenommene Unterrichtsqualität im Mathematikunterricht (Anteile der Schüler:innen in Prozent)

Unterrichtsqualität im Mathematikunterricht	Mittlerer Skalenwert		Unterrichtsqualität					
			gering		mittel		hoch	
	M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
Klassenführung	2.37	(0.02)	37.0	(1.0)	35.3	(0.8)	27.7	(0.9)
Kognitive Aktivierung	3.16	(0.01)	5.2	(0.3)	27.9	(0.7)	66.9	(0.7)
Konstruktive Unterstützung	3.33	(0.01)	5.2	(0.4)	15.5	(0.5)	79.4	(0.7)



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen, die die Unterrichtsqualität als gering einschätzen
% der Schüler:innen, die die Unterrichtsqualität als mittel einschätzen
% der Schüler:innen, die die Unterrichtsqualität als hoch einschätzen

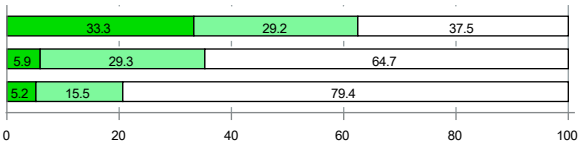
© TIMSS 2023

von einem Drittel der Schüler:innen niedrig, also wenig effizient eingeschätzt. Gemäß der in TIMSS vorgenommenen Operationalisierung deutet das auf häufige Störungen im (naturwissenschaftsbezogenen) Sachunterricht hin. Allerdings fallen auch hier – analog zu den Ergebnissen für den Mathematikunterricht – die wahrgenommene *konstruktive Unterstützung* und *kognitive Aktivierung* durch die Lehrkräfte sehr viel positiver aus: Wiederum berichten fast zwei Drittel der Schüler:innen von einer *hohen kognitiven Aktivierung* und knapp 80 Prozent von einer *hohen konstruktiven Unterstützung*. Lediglich etwa 5 Prozent der Viertklässler:innen geben an, dass sie sich durch ihre Lehrkraft nicht oder nur in geringem Ausmaß unterstützt fühlen. Ein ähnlich großer Anteil der Schüler:innen (ca. 6 %) empfindet den naturwissenschaftsbezogenen Sachunterricht zudem als wenig kognitiv aktivierend.

Im Vergleich zu den Befunden von TIMSS 2019 kann festgehalten werden, dass die Einschätzungen der Schüler:innen zum Mathematikunterricht in Deutschland in den Basisdimensionen *konstruktive Unterstützung* und *kognitive Aktivierung* in TIMSS 2023 weitestgehend ähnlich ausfallen. Die *Klassenführung* wurde im letzten TIMSS-Zyklus allerdings tendenziell etwas positiver wahrgenommen (für die Ergebnisse aus TIMSS 2019 siehe Stang et al., 2020). Für den naturwissenschaftsbezogenen Sachunterricht ergibt sich ein ähnliches Bild. Während die aktuellen Ergebnisse eine – im Sinne des Skalenmittelwerts – durchschnittliche *Klassenführung* zeigen, schätzten die Schüler:innen in TIMSS 2019 diese insgesamt etwas positiver ein. Die aktuellen Befunde zur *konstruktiven Unterstützung* und *kognitiven Aktivierung* fallen wiederum ähnlich aus wie diejenigen in TIMSS 2019.

Abbildung 10.4: Wahrgenommene Unterrichtsqualität im naturwissenschaftsbezogenen Sachunterrichts (Anteile der Schüler:innen in Prozent)

Unterrichtsqualität im Sachunterricht	Mittlerer Skalenwert		Unterrichtsqualität gering		mittel		hoch	
	M	*(SE)*	%	*(SE)*	%	*(SE)*	%	*(SE)*
Klassenführung	2.47	(0.02)	33.3	(1.0)	29.2	(0.7)	37.5	(1.1)
Kognitive Aktivierung	3.09	(0.01)	5.9	(0.4)	29.3	(0.7)	64.7	(0.8)
Konstruktive Unterstützung	3.26	(0.01)	5.2	(0.4)	15.5	(0.5)	79.4	(0.7)



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

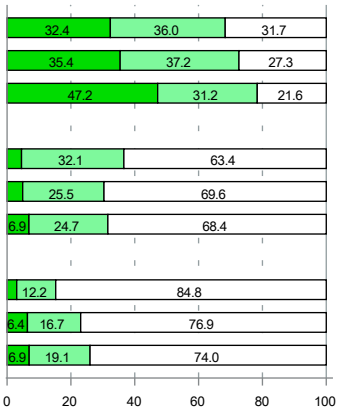
% der Schüler:innen, die die Unterrichtsqualität als gering einschätzen
% der Schüler:innen, die die Unterrichtsqualität als mittel einschätzen
% der Schüler:innen, die die Unterrichtsqualität als hoch einschätzen

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Abschließend werden die Einschätzungen der Schüler:innen zur Qualität ihres Mathematik- und Sachunterrichts in Abhängigkeit von ihrer mathematischen bzw. naturwissenschaftlichen Leistung (Kompetenzstufen) in TIMSS 2023 berichtet. Die Abbildungen 10.5 und 10.6 zeigen – ähnlich wie die zuvor erläuterten Abbildungen – Mittelwerte und Standardfehler für die drei Basisdimensionen der Unterrichtsqualität sowie die prozentualen Anteile der Schüler:innen, die diese als gering, mittel oder hoch ausgeprägt wahrnehmen. Für die *Klassenführung* im Mathematikunterricht ergeben sich demnach signifikante Unterschiede zwischen Schüler:innen der Kompetenzstufen I und II und denen, die über hoch ausgeprägte mathematische Kompetenzen verfügen. So nehmen Schüler:innen der Kompetenzstufen IV und V die *Klassenführung* etwas positiver, in der Tendenz

Abbildung 10.5: Wahrgenommene Unterrichtsqualität im Mathematikunterricht nach Kompetenzstufen (Anteile der Schüler:innen in Prozent)

Unterrichtsqualität im Mathematikunterricht		Mittlerer Skalenwert M (SE)		Unterrichtsqualität gering % (SE)		mittel % (SE)		hoch % (SE)	
Klassenführung	Kompetenzstufe IV + V	2.44	(0.02) ■	32.4	(1.4) ■	36.0	(1.4) ■	31.7	(1.7) ■
	Kompetenzstufe III	2.39	(0.02) ■	35.4	(1.3) ■	37.2	(1.3) ■	27.3	(1.3) ▼
	Kompetenzstufe I + II	2.21	(0.03) ▼	47.2	(2.1) ▲	31.2	(1.5) ▼	21.6	(1.6) ▼
Kognitive Aktivierung	Kompetenzstufe IV + V	3.13	(0.02) ■	4.5	(0.6) ■	32.1	(1.1) ■	63.4	(1.1) ■
	Kompetenzstufe III	3.18	(0.02) ■	4.9	(0.7) ■	25.5	(1.5) ▼	69.6	(1.6) ▲
	Kompetenzstufe I + II	3.17	(0.03) ■	6.9	(0.8) ▲	24.7	(1.4) ▼	68.4	(1.6) ▲
Konstruktive Unterstützung	Kompetenzstufe IV + V	3.43	(0.02) ■	3.0	(0.5) ■	12.2	(0.8) ■	84.8	(0.9) ■
	Kompetenzstufe III	3.28	(0.02) ▼	6.4	(0.7) ▲	16.7	(1.0) ▲	76.9	(1.3) ▼
	Kompetenzstufe I + II	3.24	(0.03) ▼	6.9	(0.8) ▲	19.1	(1.4) ▲	74.0	(1.7) ▼



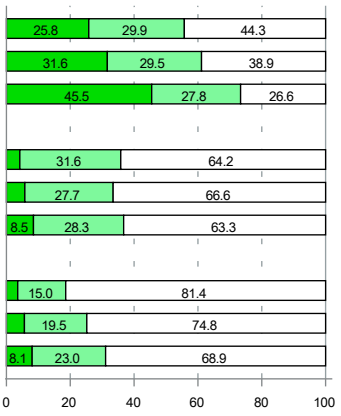
Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen, die die Unterrichtsqualität als gering einschätzen
% der Schüler:innen, die die Unterrichtsqualität als mittel einschätzen
% der Schüler:innen, die die Unterrichtsqualität als hoch einschätzen
▲ = Wert im Vergleich zur Kompetenzstufe IV + V statistisch signifikant höher ($p \leq .05$)
▼ = Wert im Vergleich zur Kompetenzstufe IV + V statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zur Kompetenzstufe IV + V nicht statistisch signifikant ($p > .05$)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Abbildung 10.6: Wahrgenommene Unterrichtsqualität im Sachunterricht nach Kompetenzstufen (Anteile der Schüler:innen in Prozent)

Unterrichtsqualität im Sachunterricht		Mittlerer Skalenwert M (SE)		Unterrichtsqualität gering % (SE)		mittel % (SE)		hoch % (SE)	
Klassenführung	Kompetenzstufe IV + V	2.59	(0.03) ■	25.8	(1.5) ■	29.9	(1.0) ■	44.3	(1.7) ■
	Kompetenzstufe III	2.51	(0.03) ■	31.6	(1.5) ▲	29.5	(1.3) ■	38.9	(1.6) ▼
	Kompetenzstufe I + II	2.27	(0.03) ▼	45.5	(1.9) ▲	27.8	(1.6) ■	26.6	(1.6) ▼
Kognitive Aktivierung	Kompetenzstufe IV + V	3.09	(0.02) ■	4.2	(0.5) ■	31.6	(1.1) ■	64.2	(1.2) ■
	Kompetenzstufe III	3.10	(0.02) ■	5.8	(0.8) ■	27.7	(1.2) ▼	66.6	(1.4) ■
	Kompetenzstufe I + II	3.07	(0.02) ■	8.5	(0.9) ▲	28.3	(1.4) ■	63.3	(1.5) ■
Konstruktive Unterstützung	Kompetenzstufe IV + V	3.36	(0.02) ■	3.5	(0.5) ■	15.0	(0.9) ■	81.4	(1.1) ■
	Kompetenzstufe III	3.25	(0.02) ▼	5.6	(1.0) ■	19.5	(1.4) ▲	74.8	(1.4) ▼
	Kompetenzstufe I + II	3.15	(0.02) ▼	8.1	(1.0) ▲	23.0	(1.6) ▲	68.9	(1.6) ▼



Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.

% der Schüler:innen, die die Unterrichtsqualität als gering einschätzen
% der Schüler:innen, die die Unterrichtsqualität als mittel einschätzen
% der Schüler:innen, die die Unterrichtsqualität als hoch einschätzen
▲ = Wert im Vergleich zur Kompetenzstufe IV + V statistisch signifikant höher ($p \leq .05$)
▼ = Wert im Vergleich zur Kompetenzstufe IV + V statistisch signifikant niedriger ($p \leq .05$)
■ = Unterschied zur Kompetenzstufe IV + V nicht statistisch signifikant ($p > .05$)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

allerdings immer noch als eher niedrig ausgeprägt wahr (z. B. auf Grund regelmäßig vorkommender Unterrichtsstörungen). Entsprechend fällt der Anteil der Schüler:innen, die die *Klassenführung* im Mathematikunterricht moderat oder hoch effizient einschätzen, bei denen mit besseren mathematischen Leistungen (Kompetenzstufe III und höher) größer aus als bei Schüler:innen der Kompetenzstufen I und II.

Für die Basisdimensionen *kognitive Aktivierung* und *konstruktive Unterstützung* fallen die Ergebnisse etwas anders aus. Unterschiede in den mittleren Einschätzungen von Schüler:innen verschiedener Kompetenzstufen finden sich für die *kognitive Aktivierung* nicht. Interessanterweise sind die Anteile der Schüler:innen der Kompetenzstufen I und II, die die *kognitive Aktivierung* als hoch (68.4 %) oder niedrig (6.9 %) ausgeprägt wahrnehmen, in beiden Fällen signifikant größer als die der Schüler:innen mit hohen mathematischen Leistungen auf Kompetenzstufe IV und V (hoch: 63.4 %, niedrig: 4.5 %). Eine eindeutige Tendenz lässt sich hier allerdings nicht ausmachen, die Unterschiede sind insgesamt marginal. Für die *konstruktive Unterstützung* zeigt sich, dass sich Schüler:innen niedriger Kompetenzstufen signifikant weniger durch ihre Mathematiklehrkraft unterstützt fühlen als solche mit hohen mathematischen Leistungen (Kompetenzstufen IV und V). Entsprechend ist auch der Anteil dieser Schüler:innen, die die *konstruktive Unterstützung* im Mathematikunterricht als gering wahrnehmen, niedriger (3.0 %), bzw. der Anteil derjenigen, der sie hoch einschätzt, ist größer (84.8 %). Die Einschätzungen zur *konstruktiven Unterstützung* und *kognitiven Aktivierung* fallen insgesamt für alle Kompetenzstufen positiv aus.

Ein ähnliches Bild zeigt sich für den naturwissenschaftsbezogenen Sachunterricht (siehe Abbildung 10.6): Die *Klassenführung* und *konstruktive Unterstützung* werden von Schüler:innen mit eher gering ausgeprägten naturwissenschaftlichen Kompetenzen signifikant niedriger eingeschätzt als von denjenigen, die hohe Kompetenzstufen erreichen. Für die *kognitive Aktivierung* zeigen sich im naturwissenschaftsbezogenen Sachunterricht wiederum keine bzw. nur marginale Unterschiede: Der Anteil der Schüler:innen, die sich gering kognitiv aktiviert fühlen, ist bei Schüler:innen der Kompetenzstufen I und II (8.5 %) signifikant größer als bei denen der Kompetenzstufen IV und V (4.2 %). Ferner ist der Anteil der Schüler:innen, die sich in mittlerem Maße kognitiv aktiviert fühlen, bei Schüler:innen der Kompetenzstufe III signifikant geringer (27.7 %) als bei denen der Kompetenzstufen IV und V (31.6 %).

10.3 Lehrkräftefortbildung

Lehrkräftefortbildungen als Bestandteil stetiger Professionalisierung im Lehrberuf nehmen vor dem Hintergrund aktueller Entwicklungen, wie der fortschreitenden Nutzung digitaler Medien im Unterricht, des Lehrkräftemangels und der zunehmenden Lehrkräftebelastungen, eine relevante Rolle für Schul- und Unterrichtsqualität ein (Eickelmann et al., 2019; KMK, 2020; Rothland & Klusmann, 2023). Auf die Bedeutsamkeit von Lehrkräftefortbildungen für die Unterrichtsqualität und, darüber vermittelt, für die Leistungen von Schüler:innen finden sich in vielen Studien Hinweise, auch wenn die Befundlage insgesamt heterogen ausfällt (Hattie, 2009; Lynch et al., 2019; Porsch & Wendt, 2016; Richter, 2023).

International finden sich unterschiedliche Regelungen zu Lehrkräftefortbildungen (Kelly et al., 2020; Reynolds et al., 2024). Diese reichen von keinen bzw. kaum Vorgaben bis hin zu sehr starken Reglementierungen, die teilweise an die Zulassungsvoraussetzungen für die Fortführung der Lehrtätigkeit gekoppelt sind. In Deutschland umfasst die staatliche Fortbildung von Lehrkräften der *Ständigen Konferenz der Kultusminister der Länder der Bundesrepublik Deutschland* (KMK) zufolge „berufsbegleitende Qualifizierungsmaßnahmen, die dem Erhalt, der Aktualisierung und der Weiterentwicklung der vorhandenen beruflichen Kompetenzen im Sinne des lebenslangen Lernens dienen“ (KMK, 2020,

S. 2). Generell besteht in Deutschland für Lehrkräfte eine allgemeine Pflicht zur Teilnahme an Fortbildungen, es zeigen sich aber regional Unterschiede, da die Rahmenbedingungen für die Fortbildung von den einzelnen Ländern der Bundesrepublik Deutschland festgelegt werden (KMK, 2017, 2020). Die Formate der angebotenen Lehrkräftefortbildungen in Deutschland sind sehr unterschiedlich und können von schulinternen bis zu länderübergreifenden sowie von halbtägigen bis zu mehrtägigen Präsenz- oder auch Online-Veranstaltungen (*blended-learning*-Formate) variieren (KMK, 2020). Damit geht ein stark uneinheitliches sowie unübersichtliches System der Lehrkräftefortbildung einher, das durch fehlende grundlegende Informationen zu beispielsweise Fortbildungsbedarfen und -angeboten geprägt ist. Entsprechend finden sich konkrete Richtlinien zu Umfang und Art der Fortbildung in nur wenigen Ländern, gleiches gilt für Fortbildungskriterien und Vorgaben zur Dokumentation der Teilnahme (Kuschel et al., 2020).

Mit dem aktuellen Zyklus TIMSS 2023 ist es erstmals im Rahmen der TIMS-Studie möglich, nicht nur die bisherige Teilnahme an Lehrkräftefortbildungen, sondern auch Angebote sowie zukünftige Bedarfe für die Domänen Mathematik und Naturwissenschaften nachzuzeichnen. Im Folgenden werden diese drei Aspekte international und national näher beleuchtet.

10.3.1 Fortbildungsteilnahme im internationalen Vergleich

In Tabelle 10.1 ist im internationalen Vergleich dargestellt, wie viel Prozent der Schüler:innen von Mathematiklehrkräften unterrichtet werden, die angeben, in den letzten zwei Jahren (also in den Jahren 2021 bis 2023) an Fortbildungsveranstaltungen teilgenommen zu haben. Dabei wird zwischen verschiedenen Schwerpunkten differenziert. Im internationalen Durchschnitt zeigt sich, dass über die verschiedenen Fortbildungsinhalte hinweg eine ähnliche Teilnahmequote vorliegt. Am seltensten werden Fortbildungen zum Thema *Leistungsfeststellung im Mathematikunterricht* (37.4 %) besucht, etwas häufiger wird an Veranstaltungen in den Bereichen *Integration von Informationstechnologien* (42.2 %) und *Lehrplan zum Mathematikunterricht* (42.6 %) teilgenommen. Die höchsten Teilnahmequoten ergeben sich wiederum bei Fortbildungen zur *Stärkung kritischen Denkens und Problemlösens* (44.0 %), zu *Mathematischen Inhalten* (44.8 %), zur *Mathematikdidaktik* (45.5 %) und zum *Eingehen auf die individuellen Bedürfnisse der Schüler:innen* (45.8 %). Im Vergleich zur Fortbildungsteilnahme in TIMSS 2019 zeigen sich damit kaum Veränderungen im internationalen Durchschnitt, einzig für die *Integration von Informationstechnologien* ist ein Anstieg um ca. 8 Prozent zu verzeichnen (siehe Guill & Wendt, 2020).

Spitzenreiter hinsichtlich der Fortbildungsteilnahme im internationalen Vergleich sind mit Macau (63.2–88.6 %), Usbekistan (72.5–84.5 %) und Schardscha (78.8–84.5 %) Teilnehmer, die 2023 das erste Mal an TIMSS teilgenommen haben. Die geringste Fortbildungsteilnahme weisen Norwegen (5.3–13.5 %) und Finnland (6.9–17.9 %) auf.

Für Deutschland zeigt sich im Vergleich zum internationalen Durchschnitt eine deutlich niedrigere und weniger gleichmäßige Teilnahme an Fortbildungen für alle Bereiche. Fast ein Drittel und damit der größte Teil der Schüler:innen wird von Lehrkräften unterrichtet, die in den letzten zwei Jahren Fortbildungen zu *Mathematischen Inhalten* besucht haben (29.8 %). Bei ca. einem Viertel trifft das auf Veranstaltungen zu *Mathematikdidaktik* (26.6 %) und *Integration von Informationstechnologien* (24.0 %) zu. Knapp ein Fünftel hat sich mit den

Tabelle 10.1: Fortbildungsteilnahme der Mathematiklehrkräfte nach Inhalten (Anteile der Schüler:innen mit entsprechenden Fortbildungen der Mathematiklehrkräfte in Prozent)

	Mathematische Inhalte		Mathematik-didaktik		Lehrplan zum Mathematik-unterricht		Stärkung kritischen Denkens oder Problemlösens		Integration von Informations-technologien		Leistungs-feststellung im Mathematik-unterricht		Eingehen auf die individuellen Bedürfnisse der Schüler:innen	
Teilnehmer	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)	%	(*SE*)
2 *Usbekistan*	80.4	(3.1)	84.5	(3.0)	72.5	(3.5)	79.6	(3.1)	75.1	(3.7)	81.9	(2.9)	74.7	(3.3)
Macau	85.9	(0.1)	88.6	(0.1)	77.1	(0.1)	69.7	(0.1)	78.0	(0.1)	63.2	(0.1)	66.7	(0.1)
Vereinigte Arabische Emirate (VAE)	75.1	(1.1)	76.3	(1.2)	71.6	(1.2)	78.2	(1.2)	76.9	(1.2)	72.4	(1.1)	78.4	(1.3)
Kuwait	77.0	(3.5)	76.7	(3.5)	69.0	(3.7)	72.2	(3.8)	72.1	(3.7)	66.2	(4.0)	76.7	(3.4)
Katar	74.9	(2.2)	76.1	(1.8)	65.0	(2.5)	74.0	(2.3)	71.2	(2.2)	60.0	(2.8)	70.5	(2.2)
3 *Hongkong*	72.8	(3.8)	76.6	(3.7)	73.0	(4.1)	59.5	(4.5)	73.8	(4.7)	57.1	(5.3)	61.3	(4.9)
2 *Kasachstan*	61.6	(3.0)	64.7	(2.8)	61.5	(2.9)	69.6	(2.6)	68.0	(3.0)	68.8	(2.4)	61.9	(2.8)
1 *Südafrika (5. Jgst.)*	79.1	(2.2)	53.7	(3.0)	84.6	(2.2)	63.6	(2.9)	44.5	(2.6)	70.9	(2.1)	56.4	(3.0)
2 *Singapur*	68.3	(2.2)	78.7	(1.9)	67.3	(2.2)	60.0	(2.2)	64.6	(2.2)	55.2	(2.0)	57.4	(2.2)
2 *Polen*	77.8	(2.9)	65.7	(3.6)	54.0	(3.6)	53.5	(3.6)	81.1	(2.9)	50.9	(3.5)	60.5	(3.5)
2 *Armenien*	57.8	(4.2)	57.5	(4.0)	61.9	(4.1)	62.4	(4.0)	67.2	(3.2)	64.0	(4.1)	68.8	(3.5)
Oman	60.7	(2.8)	69.0	(2.4)	52.4	(3.0)	61.0	(2.2)	62.4	(2.6)	64.1	(2.3)	58.0	(2.7)
2 *Georgien*	48.7	(3.8)	62.2	(3.4)	64.0	(3.5)	51.9	(3.7)	79.0	(3.0)	53.5	(3.5)	60.3	(3.4)
Iran	68.3	(2.6)	73.9	(2.6)	54.7	(2.9)	46.8	(2.9)	52.5	(3.0)	49.6	(3.0)	67.9	(2.9)
1 2 3 *England*	76.2	(3.8)	76.0	(4.4)	79.4	(3.9)	57.6	(4.8)	24.0	(4.0)	43.1	(5.0)	54.5	(5.2)
Bahrain	48.5	(3.3)	61.7	(3.2)	48.3	(3.0)	62.0	(2.7)	64.6	(3.1)	51.4	(2.3)	58.2	(2.8)
2 3 *USA*	65.7	(1.8)	56.2	(2.1)	67.4	(1.9)	52.0	(2.1)	48.5	(2.0)	42.9	(2.0)	59.2	(2.0)
2 *Litauen*	56.6	(3.2)	37.2	(3.0)	62.4	(2.8)	49.6	(3.0)	79.5	(2.3)	38.4	(3.1)	62.1	(2.9)
Taiwan	56.6	(2.8)	51.6	(3.0)	53.6	(3.2)	45.7	(2.6)	65.9	(2.6)	45.6	(2.6)	60.0	(2.4)
2 *Saudi-Arabien*	47.5	(3.0)	55.9	(2.8)	49.4	(3.1)	58.3	(3.0)	53.8	(3.1)	51.4	(2.9)	58.1	(3.1)
Lettland	57.9	(3.3)	45.2	(3.5)	57.3	(3.4)	51.4	(3.5)	47.9	(3.8)	53.9	(3.3)	51.5	(3.7)
Aserbaidschan	50.8	(3.4)	49.8	(3.7)	53.5	(3.5)	47.7	(3.5)	42.1	(3.1)	54.2	(3.4)	55.8	(3.2)
Jordanien	43.9	(3.0)	50.7	(3.3)	43.0	(3.1)	56.7	(3.7)	49.5	(3.5)	45.5	(3.3)	54.6	(3.4)
Australien	59.3	(3.2)	56.4	(3.2)	61.1	(3.0)	46.1	(3.1)	23.4	(2.8)	42.6	(3.4)	54.7	(3.0)
1 2 3 *Neuseeland*	59.4	(4.7)	61.8	(4.8)	52.9	(5.0)	57.2	(4.7)	25.4	(3.7)	33.9	(4.2)	46.2	(4.2)
Marokko	32.6	(3.1)	46.8	(2.7)	36.8	(2.3)	56.9	(2.6)	37.6	(2.6)	65.0	(2.4)	59.7	(2.7)
2 3 *Albanien*	40.2	(4.9)	52.0	(6.5)	38.9	(6.2)	61.6	(6.0)	50.6	(6.1)	50.7	(5.5)	33.5	(5.0)
2 3 *Brasilien*	55.6	(2.6)	44.3	(1.6)	35.1	(2.1)	43.6	(2.4)	34.4	(3.0)	41.8	(2.3)	48.1	(2.1)
Internationaler Mittelwert	**44.8**	**(3.1)**	**45.5**	**(3.3)**	**42.6**	**(3.2)**	**44.0**	**(3.3)**	**42.2**	**(3.2)**	**37.4**	**(3.0)**	**45.8**	**(3.3)**
2 3 *Kanada*	49.5	(1.7)	50.9	(1.9)	41.9	(1.3)	41.1	(1.8)	33.5	(1.4)	34.0	(1.7)	42.6	(1.6)
2 *Zypern*	55.2	(2.8)	47.3	(2.9)	39.6	(3.0)	33.1	(2.7)	28.3	(2.8)	34.0	(2.9)	33.8	(2.6)
2 3 *Rumänien*	28.3	(4.6)	42.4	(5.1)	34.4	(5.2)	39.1	(5.2)	40.6	(4.8)	26.8	(4.3)	57.5	(5.1)
1 2 3 *Türkei (5. Jgst.)*	40.4	(3.5)	36.8	(4.5)	34.8	(4.0)	44.7	(3.8)	49.2	(4.4)	27.0	(3.8)	28.2	(3.5)
2 *Spanien*	36.6	(2.1)	39.6	(2.2)	33.6	(2.3)	34.5	(2.1)	41.8	(2.3)	27.9	(1.9)	46.1	(2.5)
2 *Italien*	30.5	(3.1)	42.2	(3.1)	29.8	(3.6)	39.5	(3.1)	42.0	(3.3)	30.8	(3.3)	36.7	(3.1)
Nordmazedonien	34.0	(3.8)	30.4	(3.3)	50.1	(3.9)	44.1	(3.8)	28.3	(3.6)	30.4	(3.4)	33.6	(3.4)
VG OECD	**38.6**	**(3.2)**	**38.2**	**(3.4)**	**35.4**	**(3.2)**	**34.4**	**(3.2)**	**32.6**	**(3.1)**	**25.3**	**(2.9)**	**38.3**	**(3.4)**
Slowakei	19.4	(2.8)	36.4	(2.8)	43.7	(2.8)	37.9	(3.0)	47.5	(3.3)	20.3	(2.3)	34.9	(2.7)
2 *Serbien*	25.6	(3.4)	27.7	(3.7)	21.6	(3.5)	41.9	(3.8)	45.4	(4.2)	21.6	(3.3)	55.6	(3.8)
2 *Schweden*	28.8	(3.8)	29.9	(3.8)	60.6	(3.8)	21.8	(3.7)	17.5	(3.9)	37.2	(4.0)	32.4	(4.1)
2 3 *Chile*	45.0	(3.7)	30.0	(3.8)	21.9	(3.1)	38.8	(3.7)	31.6	(3.7)	18.6	(3.0)	39.8	(4.0)
2 *Frankreich*	51.7	(3.9)	57.7	(3.8)	30.8	(3.4)	41.8	(3.3)	11.0	(2.1)	7.4	(1.6)	22.0	(2.8)
VG EU	**33.4**	**(3.2)**	**33.2**	**(3.3)**	**31.1**	**(3.2)**	**30.4**	**(3.2)**	**33.0**	**(3.1)**	**23.2**	**(2.8)**	**36.4**	**(3.4)**
Republik Korea (Südkorea)	21.4	(3.4)	32.3	(4.0)	29.4	(3.8)	35.4	(3.4)	35.6	(3.8)	28.8	(3.6)	28.1	(4.0)
Japan	32.8	(3.6)	42.7	(3.8)	17.4	(3.2)	23.3	(3.3)	21.3	(3.4)	19.3	(3.3)	47.0	(4.4)
2 *Tschechien*	21.3	(2.0)	30.1	(2.2)	7.4	(1.2)	40.7	(2.4)	39.7	(2.4)	18.8	(1.9)	44.7	(2.5)
2 *Montenegro*	22.3	(2.4)	15.7	(1.7)	13.4	(1.6)	60.6	(3.1)	29.3	(2.3)	33.0	(2.2)	26.9	(2.3)
3 *Niederlande*	36.8	(4.2)	38.7	(5.1)	20.9	(4.2)	25.1	(3.8)	13.5	(3.0)	25.4	(4.0)	40.1	(4.5)
2 *Slowenien*	34.8	(3.3)	29.0	(3.1)	15.5	(2.4)	32.9	(3.4)	32.6	(3.2)	17.5	(2.8)	38.0	(3.4)
Portugal	31.7	(2.7)	29.2	(2.8)	28.5	(2.8)	23.2	(2.6)	40.2	(3.0)	15.4	(2.0)	31.0	(3.0)
Bulgarien	20.8	(3.0)	15.6	(2.7)	24.1	(3.0)	30.7	(3.5)	36.8	(3.6)	32.7	(3.6)	37.4	(3.9)
Irland	30.2	(3.8)	24.2	(3.5)	25.4	(3.7)	33.6	(4.1)	26.9	(3.4)	14.8	(2.9)	33.0	(4.0)
2 3 *Belgien (Fläm. Gem.)*	24.9	(3.1)	25.0	(3.0)	33.3	(4.0)	16.6	(3.0)	24.6	(3.3)	15.8	(2.7)	36.5	(4.0)
2 *Kosovo*	21.3	(3.4)	14.6	(2.6)	29.8	(3.6)	27.2	(3.2)	14.9	(2.8)	27.8	(3.4)	25.7	(3.5)
Deutschland	29.8	(2.7)	26.6	(2.8)	18.3	(2.2)	15.0	(2.2)	24.0	(2.5)	10.6	(1.7)	19.8	(2.4)
2 *Bosnien und Herzegowina*	13.3	(2.9)	15.6	(4.0)	24.4	(4.2)	22.4	(4.0)	15.3	(3.3)	14.6	(3.5)	20.9	(3.6)
2 *Belgien (Frz. Gem.)*	17.6	(2.5)	18.0	(2.3)	36.8	(3.1)	8.2	(1.7)	10.0	(1.9)	8.4	(1.6)	25.2	(2.5)
2 3 *Dänemark*	28.6	(3.9)	26.9	(3.9)	11.8	(3.0)	14.5	(2.8)	15.1	(2.8)	8.7	(2.3)	15.8	(3.1)
2 *Ungarn*	9.1	(2.1)	12.6	(2.7)	8.2	(2.2)	19.6	(3.4)	14.0	(2.8)	8.2	(1.7)	24.0	(3.7)
Finnland	6.9	(1.7)	11.4	(2.1)	7.7	(2.0)	7.2	(1.8)	10.5	(2.0)	7.1	(1.3)	17.9	(2.0)
1 2 *Norwegen (5. Jgst.)*	10.2	(2.0)	10.8	(2.5)	8.9	(2.4)	10.6	(2.6)	13.5	(2.9)	5.3	(1.6)	6.4	(1.8)
Benchmark-Teilnehmer														
Schardscha, VAE	81.4	(2.0)	84.3	(1.8)	81.1	(2.6)	84.5	(1.1)	84.5	(1.1)	78.8	(2.2)	82.8	(3.0)
Dubai, VAE	74.5	(1.9)	71.6	(2.8)	66.3	(3.0)	74.4	(3.0)	73.7	(2.2)	69.0	(1.3)	78.6	(1.8)
Abu Dhabi, VAE	70.5	(2.2)	73.2	(1.7)	69.2	(1.9)	74.5	(2.0)	74.3	(2.4)	67.3	(2.4)	73.9	(2.2)
2 *Ontario, Kanada*	55.1	(2.7)	49.5	(3.2)	52.9	(2.7)	42.1	(2.6)	37.3	(2.0)	31.0	(2.4)	42.7	(2.8)
2 3 *Québec, Kanada*	31.2	(3.4)	42.1	(3.9)	11.6	(1.7)	29.6	(3.3)	23.8	(2.8)	29.3	(3.3)	27.3	(2.5)

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

 © TIMSS 2023

Themen *Eingehen auf die individuellen Bedürfnisse der Schüler:innen* (19.8 %) und *Lehrplan zum Mathematikunterricht* (18.3 %) auseinandergesetzt. Am seltensten besucht wurden die Themen *Stärkung kritischen Denkens und Problemlösens* (15.0 %) sowie *Leistungsfeststellung im Mathematikunterricht* (10.6 %). Im Vergleich zu TIMSS 2019 zeigt sich damit für nahezu alle Bereiche eine Abnahme von bis zu 10 Prozent, nur die Teilnahme an Fortbildungen zur *Mathematikdidaktik* ist unverändert. Einzig für den Bereich *Integration von Informationstechnologien* zeigt sich eine deutliche Zunahme von 8.3 Prozent auf 24.0 Prozent (siehe auch Kapitel 5 in diesem Band). Staaten, die in TIMSS 2019 ähnliche Fortbildungsquoten über alle Bereiche hinweg und damit eine ähnliche Ausgangsbasis wie Deutschland aufwiesen, sind Irland, Ungarn und Schweden. Während sich für Irland ähnlich wie für Deutschland eher geringere Teilnahmequoten in TIMSS 2023 zeigen (Abnahme um ca. 5–10 %), bleiben die Werte für Ungarn relativ konstant. Schweden weist in TIMSS 2023 hingegen eine zum Teil deutlich höhere Fortbildungsteilnahme auf als in TIMSS 2019, insbesondere für Veranstaltungen zum *Lehrplan zum Mathematikunterricht* (Anstieg um 37.5 Prozentpunkte). Es ist somit – wie auch über alle anderen Staaten und Benchmark-Teilnehmer hinweg – keine eindeutige Tendenz in der Fortbildungsteilnahme von Mathematiklehrkräften zwischen TIMSS 2019 und TIMSS 2023 zu verzeichnen.

Wie viel Prozent der Schüler:innen von Sachunterrichtslehrkräften unterrichtet werden, die angeben, in den letzten zwei Jahren vor der Befragung an Fortbildungsveranstaltungen teilgenommen zu haben, ist in Tabelle 10.2 dargestellt. Neben zwei zusätzlichen Kategorien (*Naturwissenschaftlichen Sachunterricht fächerübergreifend gestalten* und *Integration von Umweltschutz und Nachhaltigkeit*) werden die gleichen Bereiche wie für die Fortbildungen in Mathematik erfragt.

Im internationalen Durchschnitt zeigt sich zwar ebenfalls eine gleichmäßige Verteilung über alle Bereiche hinweg, allerdings liegt die Teilnahmequote insgesamt rund 10 Prozent niedriger als für Mathematik. Auch im Vergleich zu 2019 fällt die Fortbildungsquote der Sachunterrichtslehrkräfte für den Großteil der Teilnehmer geringer aus (siehe Guill & Wendt, 2020). In TIMSS 2023 werden wie in Mathematik am seltensten Fortbildungen zum Thema *Leistungsfeststellung im naturwissenschaftlichen Sachunterricht* (28.4 %) und zur *Integration von Umweltschutz und Nachhaltigkeit* (29.1 %) sowie etwas häufiger zu *Naturwissenschaftlichen Sachunterricht fächerübergreifend gestalten* (31.4 %) und zur *Didaktik des naturwissenschaftlichen Sachunterrichts* (32.6 %) besucht.

Die höchsten Teilnahmequoten zeigen sich bei Fortbildungen zu *Naturwissenschaftlichen Inhalten*, zum *Lehrplan zum naturwissenschaftlichen Sachunterricht* und zur *Integration von Informationstechnologien* (jeweils 34.3 %), zur *Stärkung kritischen Denkens und Problemlösens* (35.9 %) und zum *Eingehen auf die individuellen Bedürfnisse der Schüler:innen* (36.4 %). Damit bestehen nur wenig Unterschiede zwischen den verschiedenen Inhalten und auch im Vergleich zu TIMSS 2019 zeigen sich kaum Veränderungen. Zur Kategorie *Integration von Umweltschutz und Nachhaltigkeit* können keine Aussagen zu Veränderungen getroffen werden, da sie in TIMSS 2023 das erste Mal erhoben worden ist.

Im internationalen Vergleich zeigt sich die höchste Fortbildungsteilnahme für Kuwait (68.1–77.2 %), Usbekistan (68.2–81.6 %) und Schardscha (74.0–84.7 %). Anders als für den internationalen Durchschnitt, der für die Naturwissenschaften rund 10 Prozent niedriger ausfällt als für Mathematik, zeigen sich für diese Teilnehmer keine nennenswerten Unterschiede in der Fortbildungsteilnahme zwischen den beiden Domänen. Das Schlusslicht im internationalen Vergleich bilden

Tabelle 10.2: Fortbildungsteilnahme der Sachunterrichtslehrkräfte nach Inhalten (Anteile der Schüler:innen mit entsprechenden Fortbildungen der Sachunterrichtslehrkräfte in Prozent)

Teilnehmer	Naturwiss. Inhalte % (*SE*)	Didaktik des naturwiss. Sachunterrichts % (*SE*)	Lehrplan zum naturwiss. Sachunterricht % (*SE*)	Stärkung kritischen Denkens oder Problemlösens % (*SE*)	Integration von Informationstechnologien % (*SE*)	Leistungsfeststellung im naturwiss. Sachunterricht % (*SE*)	Eingehen auf die individuellen Bedürfnisse der Schüler:innen % (*SE*)	Naturwiss. Sachunterricht fächerübergreifend gestalten % (*SE*)	Integration von Umweltschutz und Nachhaltigkeit % (*SE*)
2 *Usbekistan*	78.5 (3.5)	77.5 (3.5)	79.0 (3.5)	68.2 (3.9)	69.9 (4.0)	78.1 (3.7)	75.5 (3.3)	81.6 (3.2)	75.2 (3.7)
Kuwait	74.5 (3.3)	76.4 (3.3)	77.2 (3.6)	75.8 (3.5)	76.0 (3.9)	72.4 (3.8)	75.9 (3.6)	68.5 (3.7)	68.1 (3.6)
Vereinigte Arabische Emirate (VAE)	70.1 (1.1)	70.5 (1.0)	71.8 (1.0)	77.6 (0.9)	73.4 (0.9)	69.3 (1.0)	75.4 (1.0)	71.2 (0.9)	63.4 (0.8)
Katar	65.2 (2.6)	68.9 (2.2)	63.8 (2.8)	75.1 (2.0)	70.2 (2.1)	56.2 (2.6)	65.8 (2.3)	66.3 (2.4)	46.1 (3.1)
Macau	77.7 (0.1)	67.1 (0.1)	66.2 (0.1)	55.0 (0.1)	62.6 (0.1)	37.7 (0.1)	62.8 (0.1)	62.9 (0.1)	44.3 (0.1)
1 *Südafrika (5. Jgst.)*	66.8 (2.6)	49.7 (2.8)	73.3 (2.4)	56.0 (3.5)	53.5 (3.0)	67.0 (2.8)	56.6 (2.9)	64.2 (3.2)	43.9 (3.5)
2 *Kasachstan*	54.7 (3.3)	59.3 (3.1)	59.3 (3.4)	62.2 (3.0)	58.0 (3.3)	58.7 (3.2)	60.0 (3.4)	51.3 (3.2)	49.9 (3.2)
2 *Singapur*	68.6 (1.9)	80.1 (2.0)	70.5 (2.0)	58.0 (2.3)	62.2 (2.2)	63.2 (2.1)	49.8 (2.4)	31.9 (2.0)	27.1 (1.8)
Bahrain	42.2 (3.0)	58.0 (3.5)	39.5 (3.3)	66.0 (3.0)	70.6 (2.9)	56.2 (2.9)	53.9 (3.4)	61.3 (3.4)	53.8 (2.9)
2 *Saudi-Arabien*	51.1 (3.4)	54.4 (3.2)	52.6 (3.2)	58.6 (3.5)	52.0 (3.4)	47.0 (3.4)	56.4 (3.5)	46.2 (3.3)	46.9 (3.3)
Iran	59.1 (2.7)	63.4 (2.7)	52.1 (2.9)	46.9 (3.0)	47.9 (2.9)	50.0 (2.8)	52.8 (2.9)	46.8 (2.9)	46.2 (2.8)
Taiwan	68.1 (2.9)	51.8 (2.6)	58.6 (3.0)	39.8 (2.9)	60.2 (2.9)	39.3 (2.8)	45.0 (2.9)	37.7 (2.7)	46.9 (2.9)
Oman	42.3 (2.4)	59.2 (2.5)	43.2 (2.7)	58.1 (2.5)	49.4 (2.7)	55.4 (2.5)	50.4 (2.5)	47.0 (2.6)	38.0 (2.5)
2 *Georgien*	47.9 (3.8)	55.2 (3.8)	56.7 (3.5)	48.8 (3.6)	51.6 (3.3)	41.8 (3.8)	44.9 (3.8)	39.1 (3.3)	51.4 (3.5)
2 3 *Albanien*	33.5 (5.8)	44.4 (5.4)	40.1 (5.3)	60.7 (5.7)	50.6 (5.2)	42.0 (6.4)	41.9 (5.9)	58.0 (5.6)	54.7 (6.3)
3 *Hongkong*	52.2 (4.5)	42.6 (5.6)	38.2 (4.4)	57.3 (5.8)	52.9 (5.4)	34.9 (5.5)	59.5 (4.8)	39.2 (4.9)	40.1 (4.8)
Jordanien	39.3 (3.7)	49.8 (3.6)	42.0 (3.8)	47.9 (3.4)	49.2 (3.4)	43.9 (3.4)	46.1 (3.2)	53.4 (3.2)	45.2 (3.5)
2 *Armenien*	35.1 (4.4)	35.0 (3.9)	44.1 (4.3)	42.1 (4.0)	50.1 (4.8)	42.1 (4.3)	60.9 (4.5)	59.6 (4.4)	46.0 (4.5)
2 *Polen*	65.1 (4.1)	41.3 (3.7)	37.2 (3.8)	41.7 (3.6)	75.9 (3.4)	29.2 (3.6)	59.1 (3.9)	21.9 (3.3)	35.7 (3.4)
Lettland	52.7 (3.1)	44.9 (3.7)	45.9 (3.2)	54.9 (3.7)	42.9 (3.9)	47.2 (3.7)	38.3 (3.6)	44.6 (4.0)	29.5 (3.6)
Aserbaidschan	39.2 (3.2)	34.1 (3.1)	46.3 (3.2)	43.7 (3.6)	42.8 (3.4)	45.9 (3.4)	50.7 (3.5)	43.3 (3.1)	43.4 (3.1)
1 2 3 *England*	52.9 (4.9)	56.4 (4.5)	52.7 (5.1)	43.5 (4.9)	21.1 (3.5)	32.1 (4.7)	36.8 (4.5)	30.7 (3.7)	20.2 (3.9)
2 *Litauen*	40.4 (3.2)	24.1 (2.3)	48.3 (3.0)	36.9 (2.7)	52.6 (3.0)	24.6 (2.5)	34.2 (2.8)	42.6 (3.1)	31.0 (2.9)
Marokko	21.4 (2.6)	28.2 (2.6)	26.0 (2.4)	27.3 (2.3)	23.1 (2.6)	43.6 (2.9)	59.3 (2.6)	45.2 (2.9)	49.4 (2.7)
1 2 3 *Türkei (5. Jgst.)*	40.2 (4.0)	31.0 (3.9)	32.4 (3.8)	39.0 (4.3)	42.7 (4.0)	28.6 (3.4)	28.2 (3.2)	27.9 (3.5)	51.6 (4.7)
2 *Zypern*	54.4 (3.1)	49.7 (3.2)	45.5 (3.4)	38.8 (3.2)	33.2 (3.2)	33.1 (2.9)	24.8 (2.8)	17.9 (2.6)	19.7 (2.8)
Internationaler Mittelwert	**34.3 (3.1)**	**32.6 (3.0)**	**34.4 (3.1)**	**35.9 (3.2)**	**34.3 (3.0)**	**28.4 (2.9)**	**36.4 (3.2)**	**31.4 (3.0)**	**29.1 (3.0)**
Republik Korea (Südkorea)	33.9 (3.8)	40.3 (4.1)	33.6 (3.9)	30.7 (3.6)	33.0 (3.7)	28.7 (3.5)	23.9 (3.3)	26.4 (3.5)	35.6 (3.7)
2 3 *Brasilien*	35.2 (2.8)	31.2 (2.6)	32.7 (2.5)	29.4 (2.2)	22.1 (2.4)	27.1 (2.2)	32.6 (2.2)	31.0 (2.3)	31.8 (2.1)
Nordmazedonien	29.8 (3.5)	27.0 (3.7)	40.3 (3.5)	29.3 (3.5)	23.9 (3.2)	25.2 (3.2)	31.5 (3.3)	30.9 (3.3)	29.6 (3.3)
2 *Spanien*	23.2 (2.1)	21.8 (1.7)	25.5 (1.8)	28.0 (1.6)	41.1 (2.4)	23.8 (2.2)	44.0 (2.3)	29.2 (2.1)	26.3 (1.8)
2 3 *Rumänien*	25.6 (5.0)	26.3 (4.9)	24.5 (5.2)	35.0 (5.6)	32.2 (5.1)	16.4 (3.8)	48.1 (5.5)	31.4 (5.5)	21.1 (4.9)
2 3 *USA*	35.2 (2.0)	28.7 (1.9)	40.8 (2.2)	35.2 (2.1)	29.3 (2.2)	20.0 (1.8)	31.1 (2.0)	23.1 (1.7)	11.9 (1.5)
2 *Serbien*	24.6 (3.7)	18.8 (3.1)	17.8 (3.0)	34.7 (3.5)	43.7 (3.9)	15.7 (3.2)	40.9 (3.9)	26.1 (3.6)	26.2 (3.8)
Slowakei	16.4 (2.5)	16.9 (2.4)	36.8 (3.2)	27.3 (3.1)	39.2 (3.5)	18.0 (2.5)	25.2 (2.7)	17.1 (2.7)	17.7 (2.5)
Bulgarien	15.8 (2.5)	12.1 (2.2)	22.8 (3.4)	25.9 (3.5)	30.9 (3.6)	15.3 (2.6)	28.8 (3.8)	29.6 (3.6)	27.8 (3.3)
2 *Schweden*	28.9 (4.3)	20.3 (4.1)	55.3 (4.2)	13.1 (3.2)	11.5 (2.7)	24.2 (3.6)	16.4 (2.9)	11.8 (2.9)	17.7 (3.9)
VG EU	**22.3 (2.8)**	**18.2 (2.7)**	**22.5 (2.9)**	**22.7 (3.0)**	**23.9 (2.8)**	**14.6 (2.4)**	**24.7 (3.0)**	**18.2 (2.8)**	**29.1 (3.0)**
2 3 *Chile*	20.1 (3.1)	10.2 (2.5)	18.2 (3.1)	26.6 (3.0)	13.3 (2.4)	9.9 (2.3)	41.1 (3.5)	33.4 (3.1)	20.5 (2.5)
2 *Slowenien*	23.3 (3.1)	15.5 (2.6)	14.4 (2.6)	28.1 (3.1)	28.3 (3.2)	12.2 (2.0)	24.7 (2.8)	22.8 (3.0)	21.6 (2.8)
VG OECD	**23.3 (2.9)**	**19.9 (2.7)**	**22.9 (2.9)**	**23.6 (3.0)**	**22.4 (2.7)**	**14.8 (2.4)**	**25.0 (3.0)**	**18.7 (2.6)**	**18.2 (2.7)**
2 *Italien*	15.6 (2.3)	19.6 (2.2)	18.4 (2.7)	20.1 (2.5)	23.1 (2.8)	20.5 (2.7)	28.7 (2.8)	20.4 (2.5)	19.8 (2.7)
2 *Kosovo*	20.0 (3.8)	18.7 (3.5)	33.2 (3.8)	19.4 (3.3)	13.8 (3.2)	23.8 (3.4)	18.0 (2.9)	20.0 (3.8)	18.1 (3.1)
Portugal	22.5 (2.6)	16.1 (2.2)	18.1 (2.6)	21.2 (2.8)	22.0 (2.3)	12.7 (2.2)	22.8 (2.8)	20.9 (2.7)	25.1 (2.8)
2 *Montenegro*	18.4 (2.4)	8.5 (1.4)	9.0 (1.4)	53.8 (3.1)	22.6 (2.2)	16.9 (2.1)	19.1 (2.0)	16.0 (1.7)	14.1 (1.7)
Australien	15.5 (2.0)	18.5 (2.3)	16.8 (2.2)	28.8 (3.1)	15.1 (1.7)	8.7 (1.5)	35.4 (3.2)	16.9 (2.1)	13.0 (2.2)
2 *Tschechien*	17.7 (1.9)	11.4 (1.7)	4.0 (1.2)	28.1 (2.4)	24.3 (2.5)	6.6 (1.5)	32.2 (2.6)	14.4 (2.3)	16.7 (2.3)
1 2 3 *Neuseeland*	16.2 (3.1)	14.4 (2.7)	15.6 (3.1)	26.9 (4.2)	15.2 (3.3)	3.3 (1.7)	19.9 (4.0)	23.7 (3.8)	18.6 (3.9)
2 *Bosnien und Herzegowina*	20.0 (4.3)	14.1 (3.6)	21.6 (4.1)	26.2 (4.8)	12.1 (2.9)	14.4 (3.7)	18.6 (4.4)	14.2 (3.9)	9.8 (3.5)
Japan	28.5 (3.8)	32.4 (4.1)	16.0 (3.4)	12.2 (3.0)	12.4 (2.9)	18.1 (3.5)	19.4 (3.5)	5.9 (2.1)	1.6 (1.1)
Irland	15.3 (2.5)	14.7 (2.9)	14.4 (2.6)	22.0 (3.4)	10.1 (1.9)	5.5 (1.7)	21.8 (3.1)	17.4 (3.0)	16.3 (2.7)
2 3 *Kanada*	14.3 (1.2)	12.0 (1.0)	12.9 (1.2)	16.8 (1.4)	15.2 (1.1)	8.8 (0.8)	20.9 (1.5)	15.4 (1.1)	12.2 (1.1)
2 3 *Belgien (Fläm. Gem.)*	10.7 (2.2)	10.1 (2.0)	20.9 (3.4)	13.1 (2.9)	18.0 (3.2)	5.7 (1.7)	17.9 (3.1)	12.6 (2.1)	12.8 (2.6)
2 *Belgien (Frz. Gem.)*	13.1 (2.0)	13.1 (2.0)	29.0 (2.8)	6.8 (1.3)	4.2 (1.4)	4.6 (1.2)	7.4 (1.8)	7.4 (1.5)	5.4 (1.5)
2 *Ungarn*	5.9 (1.9)	6.8 (2.0)	6.4 (2.2)	17.7 (3.2)	4.8 (1.7)	4.4 (1.7)	24.9 (3.7)	7.0 (2.1)	12.4 (2.6)
3 *Niederlande*	10.0 (2.8)	10.5 (2.9)	7.2 (2.1)	16.8 (3.8)	6.6 (2.3)	3.4 (1.7)	9.1 (3.1)	9.8 (2.6)	5.4 (1.7)
Deutschland	13.7 (2.2)	6.2 (1.5)	9.3 (1.7)	8.5 (1.8)	12.7 (2.2)	5.1 (1.1)	6.7 (1.5)	7.6 (1.4)	8.6 (1.8)
2 3 *Dänemark*	13.5 (2.7)	9.7 (2.2)	2.2 (1.0)	8.1 (2.1)	5.8 (2.2)	2.8 (1.2)	7.6 (2.0)	6.0 (2.1)	6.6 (2.3)
2 *Frankreich*	6.0 (1.7)	7.6 (2.0)	4.8 (1.6)	6.0 (2.0)	3.3 (1.4)	1.4 (0.9)	7.9 (2.3)	4.3 (1.6)	10.2 (2.4)
Finnland	1.2 (0.6)	1.3 (0.6)	4.3 (1.6)	2.3 (0.9)	4.2 (1.5)	5.2 (1.3)	13.0 (2.0)	3.0 (0.9)	4.3 (1.3)
1 2 *Norwegen (5. Jgst.)*	4.0 (1.6)	3.0 (1.4)	2.8 (0.9)	2.8 (1.4)	3.6 (1.5)	1.9 (1.2)	4.5 (1.6)	5.5 (1.5)	2.7 (1.2)
Benchmark-Teilnehmer									
Schardscha. VAE	79.9 (2.1)	79.6 (2.2)	79.0 (1.0)	82.5 (1.0)	83.9 (1.4)	76.5 (1.4)	84.7 (1.4)	80.3 (1.6)	74.0 (2.3)
Dubai. VAE	66.0 (2.2)	65.8 (1.9)	67.8 (1.6)	76.7 (1.8)	68.4 (1.9)	66.4 (1.8)	71.4 (2.3)	64.6 (2.1)	58.7 (1.9)
Abu Dhabi. VAE	67.2 (2.0)	68.1 (1.9)	70.3 (2.2)	72.0 (2.2)	69.5 (1.7)	64.3 (1.9)	69.5 (2.0)	66.3 (1.9)	53.9 (1.7)
2 *Ontario. Kanada*	18.2 (1.8)	14.0 (1.4)	19.0 (1.9)	18.7 (2.1)	17.7 (1.8)	10.2 (1.2)	21.4 (2.5)	18.0 (1.9)	16.3 (1.8)
2 3 *Québec. Kanada*	8.7 (1.8)	10.1 (2.3)	3.5 (1.3)	6.0 (1.4)	9.8 (2.0)	5.8 (1.9)	12.2 (2.2)	8.6 (1.9)	5.1 (1.6)

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

 © TIMSS 2023

wie in Mathematik Norwegen (1.9–5.5 %) und Finnland (1.2–13.0 %), allerdings weisen Frankreich (1.4–10.2 %) und Québec (3.5–12.2 %) über alle Bereiche hinweg nur nominell höhere Werte in der Fortbildungsteilnahme auf.

Im internationalen Vergleich gehört Deutschland ebenfalls zu den Teilnehmerstaaten, die sehr geringe Fortbildungsquoten für die Naturwissenschaften aufweisen. Am häufigsten werden Schüler:innen von Lehrkräften unterrichtet, die in den letzten zwei Jahren Veranstaltungen zu *Naturwissenschaftlichen Inhalten* (13.7 %) und zur *Integration von Informationstechnologien* (12.7 %) besucht haben. Je knapp 10 Prozent entfallen auf die Bereiche *Lehrplan zum naturwissenschaftlichen Sachunterricht* (9.3 %)*, Integration von Umweltschutz und Nachhaltigkeit* (8.6 %) sowie *Stärkung kritischen Denkens und Problemlösens* (8.5 %). Eine sehr geringe Teilnahmequote zeigt sich für die Fortbildungsthemen *Naturwissenschaftlichen Sachunterricht fächerübergreifend gestalten* (7.6 %), *Eingehen auf die individuellen Bedürfnisse der Schüler:innen* (6.7 %), *Didaktik des naturwissenschaftlichen Sachunterrichts* (6.2 %) und *Leistungsfeststellung im naturwissenschaftlichen Sachunterricht* (5.1 %). Im Vergleich zu TIMSS 2019 zeigt sich in Deutschland somit eine zum Teil gravierende Abnahme in der Fortbildungsteilnahme um bis zu 19 Prozentpunkte (siehe Guill & Wendt, 2020). Lediglich für die *Integration von Informationstechnologien* (12.7 %) zeigen sich nahezu unveränderte Werte (siehe auch Kapitel 5 in diesem Band).

Vergleichbare Fortbildungsquoten und damit eine ähnliche Ausgangsbasis wie Deutschland in TIMSS 2019 weisen die Staaten Irland und Japan auf. Für beide Staaten ist die Teilnahmequote in TIMSS 2023 in nahezu allen Bereichen gesunken, womit sie eine ähnliche Entwicklung wie Deutschland zeigen. Für Sachunterrichtslehrkräfte kann somit anders als für Mathematiklehrkräfte im Vergleich von 2019 zu 2023 für viele Teilnehmer eine deutlichere Abnahme der Fortbildungsteilnahme festgestellt werden.

10.3.2 Fortbildungsangebote und -bedarfe in Deutschland

In den Tabellen 10.3 und 10.4 sind neben der Fortbildungsteilnahme in Deutschland in den letzten zwei Jahren auch die in diesem Zeitraum wahrgenommenen, d. h. die den Lehrkräften bekannten, Angebote sowie aktuellen Bedarfe an Lehrkräftefortbildungen dargestellt. Diese werden wie in den internationalen Vergleichen für verschiedene Inhalte von Mathematik und Naturwissenschaften differenziert.

Wie in Abschnitt 10.3.1 beschrieben zeigt sich für Mathematik, dass Schüler:innen am häufigsten von Lehrkräften unterrichtet werden, die Fortbildungen zu *Mathematischen Inhalten* und *Mathematikdidaktik* besucht haben (siehe Tabelle 10.3). Am seltensten nahmen Lehrkräfte an Fortbildungen zur *Leistungsfeststellung im Mathematikunterricht* und zur *Stärkung kritischen Denkens oder Problemlösens* teil. Hinsichtlich der Fortbildungsangebote zeigt sich, dass mit Abstand am häufigsten Angebote zu *Mathematischen Inhalten* (70.2 %), zur *Mathematikdidaktik* (65.1 %) und zur *Integration von Informationstechnologien* (57.1 %) bekannt sind. Alle anderen Inhalte weisen Werte von ca. 45 Prozent auf. Darüber hinaus lässt sich erkennen, dass der höchste Bedarf an Fortbildungen in den Bereichen *Integration von Informationstechnologien* (68.1 %) sowie etwas geringer für die *Stärkung kritischen Denkens oder Problemlösens* (58.0 %) und das *Eingehen auf die individuellen Bedürfnisse der Schüler:innen* (57.2 %)

Tabelle 10.3: Fortbildungsteilnahme, -angebote und -bedarfe der Mathematiklehrkräfte (Anteile der Schüler:innen mit entsprechenden Fortbildungen der Mathematiklehrkräfte in Prozent)

	Fortbildungs-teilnahme		Fortbildungs-angebote		Fortbildungs-bedarfe	
	%	(*SE*)	%	(*SE*)	%	(*SE*)
Mathematische Inhalte	29.8	(2.7)	70.2	(2.7)	37.8	(2.7)
Mathematikdidaktik	26.6	(2.8)	65.1	(2.9)	46.0	(2.8)
Lehrplan zum Mathematikunterricht	18.3	(2.2)	44.7	(3.1)	20.6	(2.2)
Stärkung kritischen Denkens oder Problemlösens	15.0	(2.2)	47.4	(2.9)	58.0	(3.1)
Integration von Informationstechnologien	24.0	(2.5)	57.1	(3.1)	68.1	(2.9)
Leistungsfeststellung im Mathematikunterricht	10.6	(1.7)	44.8	(3.0)	36.8	(2.8)
Eingehen auf die individuellen Bedürfnisse der Schüler:innen	19.8	(2.4)	46.9	(3.3)	57.2	(2.4)
Sprachsensibler Fachunterricht	20.2	(2.2)	46.9	(2.8)	37.7	(3.1)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

besteht. Der geringste Bedarf zeigt sich mit 20.6 Prozent für Fortbildungen zum *Lehrplan zum Mathematikunterricht*.

Werden die Ergebnisse zu Teilnahme, Angeboten und Bedarfen für die einzelnen Inhalte einander gegenübergestellt, lassen sich Aussagen zur Passung dieser drei Kategorien treffen. In Tabelle 10.3 zeigen sich dabei zum Teil auffällige Diskrepanzen, die insbesondere zwei Muster erkennen lassen. Das erste Muster zeichnet sich durch im Verhältnis eher geringe Teilnahmen und Bedarfe bei gleichzeitig sehr hoher Bekanntheit von Angeboten aus. Dies betrifft insbesondere die Fortbildungen zu *Mathematischen Inhalten*, in geringerem Maße auch Fortbildungen zum *Lehrplan zum Mathematikunterricht*. Das zweite Muster lässt sich dadurch charakterisieren, dass die aktuellen Bedarfe die bisherigen Angebote übersteigen und zeigt sich für die Fortbildungsinhalte *Integration von Informationstechnologien*, *Eingehen auf die individuellen Bedürfnisse der Schüler:innen* und *Stärkung kritischen Denkens oder Problemlösens*. Für die übrigen Inhalte sind die Unterschiede zwischen den Kategorien geringer bzw. uneinheitlich. Insgesamt kann für Fortbildungen im Bereich Mathematik festgestellt werden, dass die Teilnahme an Lehrkräftefortbildungen deutlich niedriger ausfällt als Angebote bekannt sind und Bedarfe angegeben werden.

In Tabelle 10.4 sind die Ergebnisse für Fortbildungen zu den Naturwissenschaften dargestellt. Betreffend der Teilnahme lässt sich feststellen, dass mit etwas über 10 Prozent Schüler:innen am häufigsten von Lehrkräften unterrichtet werden, die Fortbildungen zu *Naturwissenschaftlichen Inhalten*, zur *Integration von Informationstechnologien* und zum *Sprachsensiblen Fachunterricht* besucht haben. Die geringste Teilnahme zeigt sich für *Leistungsfeststellung im naturwissenschaftlichen Sachunterricht*, zur *Didaktik des naturwissenschaftlichen Sachunterrichts* und zum *Eingehen auf die individuellen Bedürfnisse der Schüler:innen* (siehe auch Tabelle 10.2). Für die Angebote an Fortbildungen zeigt sich, dass, wie in Mathematik, mit Abstand am häufigsten Angebote zu fachspezifischen Inhalten (*Naturwissenschaftliche Inhalte*; 57.6 %) von Lehrkräften bemerkt werden. Eine etwas geringere Bekanntheit kann für die Inhalte *Inte-*

Tabelle 10.4: Fortbildungsteilnahme, -angebote und -bedarfe der Sachunterrichtslehrkräfte (Anteile der Schüler:innen mit entsprechenden Fortbildungen der Sachunterrichtslehrkräfte in Prozent)

	Fortbildungs-teilnahme		Fortbildungs-angebote		Fortbildungs-bedarfe	
	%	(*SE*)	%	(*SE*)	%	(*SE*)
Naturwissenschaftliche Inhalte	13.7	(2.2)	57.6	(3.3)	50.2	(3.2)
Didaktik des naturwiss. Sachunterrichts	6.2	(1.5)	43.2	(3.2)	42.9	(3.2)
Lehrplan zum naturwiss. Sachunterricht	9.3	(1.7)	37.7	(3.6)	28.8	(3.1)
Stärkung kritischen Denkens oder Problemlösens	8.5	(1.8)	39.0	(3.6)	46.5	(3.3)
Integration von Informationstechnologien	12.7	(2.2)	44.9	(3.3)	51.2	(3.2)
Leistungsfeststellung im naturwiss. Sachunterricht	5.1	(1.2)	30.8	(3.2)	28.8	(2.9)
Eingehen auf die individuellen Bedürfnisse der Schüler:innen	6.7	(1.5)	29.5	(3.4)	34.8	(3.1)
Naturwiss. Sachunterricht fächerübergreifend gestalten	7.6	(1.5)	36.2	(3.5)	45.9	(3.1)
Integration von Umweltschutz und Nachhaltigkeit	8.6	(1.8)	47.9	(3.3)	49.5	(3.1)
Sprachsensibler Fachunterricht	10.3	(2.0)	35.0	(3.4)	34.7	(3.1)

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

gration von Umweltschutz und Nachhaltigkeit (47.9 %), *Integration von Informationstechnologien* (44.9 %) sowie *Didaktik des naturwissenschaftlichen Sachunterrichts* (43.2 %) festgestellt werden. Für die weiteren Fortbildungsinhalte ergeben sich jeweils Werte zwischen ca. 30 und 40 Prozent.

Hinsichtlich des Bedarfes an Fortbildungen zeigen sich die höchsten Werte mit jeweils ca. 50 Prozent in den Bereichen *Integration von Informationstechnologien* (51.2 %), *Naturwissenschaftliche Inhalte* (50.2 %) und *Integration von Umweltschutz und Nachhaltigkeit* (49.5 %). Der geringste Bedarf ergibt sich mit jeweils 28.8 Prozent für Fortbildungen zum *Lehrplan zum naturwissenschaftlichen Sachunterricht* und zur *Leistungsfeststellung im naturwissenschaftlichen Sachunterricht.*

Bei der Gegenüberstellung von Teilnahme, Angeboten und Bedarfen an Fortbildung zeigen sich in den Naturwissenschaften, ähnlich wie für Mathematik, Unterschiede. Das erste Muster lässt sich durch eine im Verhältnis sehr geringe Teilnahme an Fortbildungen, die im Kontrast zu den hohen Werten für Angebote und Bedarfe steht, charakterisieren (*Didaktik des naturwissenschaftlichen Sachunterrichts* und *Integration von Umweltschutz und Nachhaltigkeit*). Das zweite Muster zeichnet sich ebenfalls durch eine hohe Differenz zwischen Teilnahme und Angeboten aus. Allerdings übersteigen zudem in diesem Muster die Bedarfe an Fortbildungsveranstaltungen die festgestellten Angebote. Dies betrifft insbesondere Fortbildungen zur *Integration von Informationstechnologien,* zur *Stärkung kritischen Denkens oder Problemlösens* und zu *Naturwissenschaftlichen Sachunterricht fächerübergreifend gestalten*.

Darüber hinaus lässt sich auch das erste Muster aus dem Bereich Mathematik in den Naturwissenschaften erkennen, bei dem die Angebote die Bedarfe übersteigen. Dies zeigt sich in den Naturwissenschaften für Fortbildungen zu den

Naturwissenschaftlichen Inhalten und zum *Lehrplan zum naturwissenschaftlichen Sachunterricht* und damit für die gleichen Bereiche wie in Mathematik. Die Diskrepanzen sind hier jedoch nicht so deutlich wie in Mathematik. Insgesamt lässt sich feststellen, dass die Teilnahme von Sachunterrichtslehrkräften an Fortbildungen, auch im Vergleich zu Mathematik, deutlich niedriger ausfällt als Angebote festgestellt und Bedarfe angegeben werden.

10.4 Zusammenfassung

Im Folgenden werden zunächst die zentralen Ergebnisse zur Unterrichtsqualität in Deutschland und im internationalen Vergleich und anschließend die Befunde zur Teilnahme an Lehrkräftefortbildung sowie zu den Angeboten und den aktuellen Bedarfen zusammengefasst.

Unterrichtsqualität

Im ersten Teil des Kapitels wurden Einschätzungen von Viertklässler:innen zur Unterrichtsqualität in Mathematik und im naturwissenschaftsbezogenen Sachunterricht in Deutschland sowie im internationalen Vergleich berichtet. Die Einschätzungen der Viertklässler:innen zu ihrem Mathematik- bzw. Sachunterricht wurden national hinsichtlich der drei Basisdimensionen *Klassenführung*, *kognitive Aktivierung* und *konstruktive Unterstützung* erfasst. Im internationalen Vergleich wurden dagegen lediglich die Einschätzungen der Schüler:innen zur *Klassenführung* erhoben.

Eine große Mehrheit der Schüler:innen in Deutschland nimmt die Häufigkeit von Störungen in ihrem Mathematik- und Sachunterricht als *moderat* ausgeprägt wahr, etwa 17 Prozent der Schüler:innen fühlen sich häufig gestört. Dieses Ergebnis findet sich tendenziell auch in den Durchschnittswerten der EU- und OECD-Staaten, jedoch ist der Anteil der Schüler:innen, die kaum oder keine Unterrichtsstörungen wahrnehmen, in den genannten Vergleichsgruppen höher ausgeprägt. In der Literatur wird regelmäßig diskutiert, inwiefern die Wahrnehmung von Unterrichtsstörungen über verschiedene kulturelle Kontexte hinweg verglichen werden kann (z. B. Senden et al., 2023). In vertiefenden Analysen wäre demnach zu klären, ob es sich bei den hier gezeigten Differenzen nicht auch um kulturell geprägte, staatenspezifische Antworttendenzen handelt.

Die nationalen Ergebnisse zur Unterrichtsqualität verweisen ebenfalls darauf, dass Viertklässler:innen in Deutschland die *Klassenführung* in ihrem Mathematik- und naturwissenschaftsbezogenen Sachunterricht im Mittel als *moderat* effizient einschätzen. Die Angaben zur *kognitiven Aktivierung* und zur *konstruktiven Unterstützung* fallen deutlich positiver aus. Erfreulich ist, dass etwa zwei Drittel (*kognitive Aktivierung*) bzw. drei Viertel (*konstruktive Unterstützung*) der Schüler:innen eine hohe Unterrichtsqualität hinsichtlich dieser Basisdimensionen wahrnehmen. Dieser Befund fand sich in ähnlicher Weise auch in TIMSS 2019 (Stang et al., 2020) und ist insofern bemerkenswert, als dass zwischen den Einschätzungen von Schüler:innen zur Unterrichtsqualität vor und nach der COVID-19-Pandemie kaum Unterschiede zu verzeichnen sind. Erste Ergebnisse zur Unterrichtswirklichkeit während der COVID-19-Pandemie liegen bereits vor und könnten Erklärungsansätze liefern (im Überblick Voss & Wittwer, 2020).

Schließlich wurde in diesem Kapitel auch untersucht, inwieweit die von Viertklässler:innen in Deutschland wahrgenommene Unterrichtsqualität mit ihrer

mathematischen oder naturwissenschaftlichen Testleistung zusammenhängt. Dazu wurden die Einschätzungen der Schüler:innen zu ihrem Mathematik- und Sachunterricht in Beziehung mit den in TIMSS 2023 erreichten Kompetenzstufen gesetzt. Wichtige Unterschiede zeigen sich vor allem für die von den Schüler:innen wahrgenommene *Klassenführung* und *konstruktive Unterstützung*. Demnach schätzen Viertklässler:innen, die in TIMSS 2023 hohe Kompetenzstufen erreichen, die *Klassenführung* und die *konstruktive Unterstützung* im Unterricht etwas besser ein als ihre Mitschüler:innen auf den Kompetenzstufen I und II. Einschränkend ist anzumerken, dass die querschnittliche Anlage von TIMSS keine gesicherte Aussage über die Richtung der Zusammenhänge zulässt. So ist es zwar einerseits denkbar, dass hier die theoretische Annahme (d. h. Unterrichtsqualität beeinflusst Leistung) bestätigt wird, anderseits sind auch umgedrehte oder reziproke Beziehungen denkbar (Leistung beeinflusst wahrgenommene Unterrichtsqualität bzw. Unterrichtsqualität und Leistung beeinflussen sich wechselseitig).

Fortbildung

Im zweiten Teil des Kapitels wurde das Thema Lehrkräftefortbildung in den Blick genommen. In der Gesamtheit zeigt sich, dass die Schüler:innen in Deutschland von Mathematik- und Sachunterrichtslehrkräften unterrichtet werden, die in den letzten zwei Jahren (2021 bis 2023) im internationalen Vergleich eher selten Fortbildungen besucht haben. Über alle Staaten und Benchmark-Teilnehmer hinweg lässt sich erkennen, dass es für beide Domänen große Unterschiede in den Teilnahmequoten gibt. In einigen Staaten betragen diese bis zu ca. 90 Prozent, womit sie deutlich über dem internationalen Mittelwert liegen. Staaten wie die USA, Hongkong, Taiwan oder Neuseeland, in denen Mindestvorgaben zum Umfang der Fortbildung existieren oder Fortbildungen zum Erhalt der Lehrlizenz besucht werden müssen, weisen dabei zum Teil deutlich höhere Teilnahmequoten auf als Deutschland und weitere Teilnehmer, in denen es wenige oder keine offiziellen Vorgaben zur Fortbildungsteilnahme gibt (Kelly et al., 2020; Reynolds et al., 2024). Für Deutschland ist darüber hinaus anzumerken, dass sich die Vorgaben zwischen den einzelnen Ländern der Bundesrepublik Deutschland noch einmal voneinander unterscheiden.

Über alle Staaten hinweg zeigt sich, dass die Teilnahmequoten für naturwissenschaftsspezifische Fortbildungen niedriger ausfallen als die für Mathematik, was angesichts der zentralen Bedeutung des Fachs Mathematik für die Schule erklärbar ist. In Deutschland ist im Vergleich zu TIMSS 2019 über fast alle Themenfelder hinweg eine Abnahme der Teilnahme zu verzeichnen. Die Abnahme in den Teilnahmequoten ist für die Naturwissenschaften stärker als für Mathematik, was vermutlich mit der Fokussierung auf die Kernfächer und Basiskompetenzen während der COVID-19-Pandemie zusammenhängt. Einzig für den Inhalt *Integration von Informationstechnologien* zeigen sich höhere bzw. gleichbleibende Teilnahmequoten. Dieser Befund ist vor dem Hintergrund der zunehmenden Bedeutung digitaler Medien im schulischen Kontext, insbesondere durch die COVID-19-Pandemie, erwartungskonform.

Ein möglicher Erklärungsansatz für die grundsätzlich eher geringe Fortbildungsteilnahme in Deutschland in TIMSS 2019 stellte das Fehlen von Fortbildungsangeboten dar, weshalb in TIMSS 2023 zusätzlich nach Angeboten in den letzten zwei Jahren sowie aktuellen Bedarfen gefragt wurde. In den national vertiefenden Analysen wird ersichtlich, dass es teils große Diskrepanzen zwischen der Teilnahme, den Angeboten und den Bedarfen an Fortbildung gibt. Das

Fehlen von Fortbildungsangeboten scheint dabei allerdings nicht ursächlich für die geringen Teilnahmequoten zu sein, da die Angebote durchgängig die realisierten Teilnahmen übersteigen. Stattdessen deuten die Ergebnisse eher auf strukturelle Hürden hin, da sowohl Angebote als auch Bedarfe berichtet werden, die sich jedoch nicht in entsprechend hohen Teilnahmequoten niederschlagen. Es scheint somit andere als die in diesem Kapitel untersuchten Gründe zu geben, die die Lehrkräfte daran hindern, häufiger an Fortbildungen teilzunehmen. Vertiefende Analysen sowie Forschungen sind nötig, um weitere Einblicke zu gewinnen und die Gründe für die verhältnismäßig geringen Teilnahmequoten besser zu verstehen.

Mögliche strukturelle Hürden könnten sich zum einen in einer unzureichenden Passung von Fortbildungsangeboten und -bedarfen zeigen (Richter et al., 2018). Zum anderen könnten auch aktuelle Herausforderungen im Lehrberuf (wie beispielsweise Lehrkräftemangel oder gestiegene Anforderungen) verhindern, dass Lehrkräfte die von ihnen formulierten Bedarfe an Fortbildungen erfüllen können (Robert Bosch Stiftung, 2024). Hier wäre es notwendig, auf struktureller sowie organisatorischer Ebene adäquate Voraussetzungen zu schaffen, die es den Lehrkräften ermöglichen, Fortbildungen zu besuchen.

Literatur

Ditton, H. (2002). Unterrichtsqualität – Konzeptionen, methodische Überlegungen und Perspektiven. *Unterrichtswissenschaft, 30*(3), 197–212. https://doi.org/10.25656/01:7685

Eickelmann, B., Bos, W., Gerick, J., Goldhammer, F., Schaumburg, H., Schwippert, K., Senkbeil, M., & Vahrenhold, J. (Hrsg.). (2019). *ICILS 2018 #Deutschland. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern im zweiten internationalen Vergleich und Kompetenzen im Bereich Computational Thinking*. Waxmann. https://doi.org/10.25656/01:18166

Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school. Dimensions and prediction of student outcomes. *Learning and Instruction*, *29*, 1–9. https://doi.org/10.1016/j.learninstruc.2013.07.001

Guill, K., & Wendt, H. (2020). Privater Nachhilfeunterricht und Lehrkräftefortbildung am Ende der Grundschulzeit. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 209–221). Waxmann. https://doi.org/10.31244/9783830993193

Hattie, J. A. C. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Helmke, A., & Schrader, F.-W. (2008). Merkmale der Unterrichtsqualität: Potenzial, Reichweite und Grenzen. *Seminar - Lehrerbildung und Schule*, *3*, 17–47.

Helmke, A., & Weinert, F. E. (2022). *Unterrichtsqualität und Professionalisierung: Diagnostik von Lehr-Lern-Prozessen und evidenzbasierte Unterrichtsentwicklung*. Klett-Kallmeyer.

Henschel, S., Rjosk, C., & Heinschel, A. (2023). Merkmale der Unterrichtsqualität im Fach Deutsch. In P. Stanat, S. Schipolowski, R. Schneider, S. Weirich, S. Henschel & K. A. Sachse (Hrsg.), *IQB-Bildungstrend 2022. Sprachliche Kompetenzen am Ende der 9. Jahrgangsstufe im dritten Ländervergleich* (S. 359–388). Waxmann. https://doi.org/10.31244/9783830997771

Kelly, D. L., Centurino, V. A. S., Martin, M. O., & Mullis, I. V. S. (2020). *TIMSS 2019 Encyclopedia: Education Policy and Curriculum in Mathematics and Science*. https://timssandpirls.bc.edu/timss2019/encyclopedia/index.html

Kleickmann, T., Steffensky, M., & Praetorius, A.-K. (2020). Quality of teaching in science education: More than three basic dimensions? *Zeitschrift für Pädagogik, 66*, 37–55. https://doi.org/10.3262/ZPB2001037

Klieme, E., Schümer, G., & Knoll, S. (2001). TIMSS – Impulse für Schule und Unterricht. In E. Klieme & J. Baumert (Hrsg.), *Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur" und Unterrichtsgestaltung im internationalen Vergleich* (S. 43–57). BMBF.

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2017). *Lehrkräftefortbildung in den Ländern*. https://www.kmk.org/fileadmin/Dateien/pdf/Bildung/AllgBildung/2017-12-19_Lehrerfortbildung_in_den_Laendern__003_.pdf

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2020). *Ländergemeinsame Eckpunkte zur Fortbildung von Lehrkräften als ein Bestandteil ihrer Professionalisierung in der dritten Phase der Lehrerbildung*. Beschluss der Kultusministerkonferenz vom 12.03.2020. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2020/2020_03_12-Fortbildung-Lehrkraefte.pdf

Kounin, J. S. (1976). *Techniken der Klassenführung*. Klett.

Kunter, M. (2016). Bedingungen und Effekte von Unterricht: Aktuelle Forschungsperspektiven aus der pädagogischen Psychologie. In N. McElvany, W. Bos, H. G. Holtappels, M. Gebauer & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (S. 9–32). Waxmann.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Waxmann. https://doi.org/10.31244/9783830974338

Kunter, M., & Trautwein, U. (2013). *Psychologie des Unterrichts*. Schöningh/UTB. https://doi.org/10.36198/9783838538952

Kuschel, J., Richter, D., & Lazarides, R. (2020). Wie relevant ist die gesetzliche Fortbildungsverpflichtung für Lehrkräfte? Eine empirische Untersuchung zur Fortbildungsteilnahme in verschiedenen deutschen Bundesländern. *Zeitschrift für Bildungsforschung*, *10*(2), 211–229. https://doi.org/10.1007/s35834-020-00274-3

Lipowsky, F. (2020). Unterricht. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (3. überarb. Aufl., S. 69–118). Springer. https://doi.org/10.1007/978-3-662-61403-7_4

Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, *19*, 527–537. https://doi.org/10.1016/j.learninstruc.2008.11.001

Lynch, K., Hill, H. C., Gonzalez, K. E., & Pollard, C. (2019). Strengthening the research base that informs STEM instructional improvement efforts: A meta-analysis. *Educational Evaluation and Policy Analysis, 41*(3), 260–293. https://doi.org/10.3102/0162373719849044

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109–119. https://doi.org/10.3102/0013189X09332374

Porsch, R., & Wendt, H. (2016). Aus- und Fortbildung von Mathematik und Sachunterrichtslehrkräften. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 189–204). Waxmann. https://doi.org/10.25656/01:14022

Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: the German framework of Three Basic Dimensions. *ZDM – Mathematics Education*, *50*, 407–426. https://doi.org/10.1007/s11858-018-0918-4

Reynolds, K. A., Aldrich, C. E. A., Bookbinder, A., Gallo, A., von Davier, M., & Kennedy, A. (Hrsg.). (2024). *TIMSS 2023 Encyclopedia: Education policy and curriculum in mathematics and science*. TIMSS & PIRLS International Study Center.

Richter, E., Richter, D., & Marx, A. (2018). Was hindert Lehrkräfte an Fortbildungen teilzunehmen? *Zeitschrift für Erziehungswissenschaft*, *21*(5), 1021–1043. https://doi.org/10.1007/s11618-018-0820-4

Richter, D. (2023). Lehrerinnen und Lehrer lernen: Fort- und Weiterbildung im Lehrer:innenberuf. In M. Rothland (Hrsg.), *Beruf Lehrer:in. Ein Studienbuch* (2. Aufl.). (S. 277–293). Waxmann. https://doi.org/10.36198/9783838588216

Rieser, S., Stahns, R., Walzebug, A., & Wendt, H. (2016). Einblicke in die Gestaltung des Mathematik- und Sachunterrichts. In H. Wendt, W. Bos, C. Selter, O. Köller,

K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 205–224). Waxmann. https://doi.org/10.25656/01:14022

Robert Bosch Stiftung (2024). *Deutsches Schulbarometer: Befragung Lehrkräfte. Ergebnisse zur aktuellen Lage an allgemein- und berufsbildenden Schulen.* Robert Bosch Stiftung.

Rothland, M., & Klusmann, U. (2023). Belastung und Beanspruchung im Lehrer:innenberuf. In M. Rothland (Hrsg.), *Beruf Lehrer:in. Ein Studienbuch* (2. Aufl.). (S. 405-431). Waxmann. https://doi.org/10.36198/9783838588216

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.

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, *77*(4), 454–499. https://doi.org/10.3102/0034654307310317

Senden, B., Nilsen, T., & Blömeke, S. (2022). Instructional Quality: A Review of Conceptualizations, Measurement Approaches, and Research Findings. In M. Blikstad-Balas, K. Klette & M. Tengberg, (Hrsg.). *Ways of analyzing teaching quality. Potential and pitfalls* (S. 140–172). Scandinavian University Press. https://doi.org/10.18261/9788215045054-2021

Senden, B., Teig, N., & Nilsen, T. (2023). Studying the comparability of student perceptions of teaching quality across 38 countries. *International Journal of Educational Research Open, 5*, 100309. https://doi.org/10.1016/j.ijedro.2023.100309

Slavin, R. E. (2000). *Educational psychology: Theory and practice* (6. Aufl.). Allyn and Bacon.

Sliwka, A., Klopsch, B., & Dumont, H. (2019). *Konstruktive Unterstützung im Unterricht.* (Bd. 3). Institut für Bildungsanalysen Baden-Württemberg.

Stang, J., Lepper, C., Steffensky, M., & McElvany, N. (2020). Einblicke in die Gestaltung des Mathematik- und naturwissenschaftsbezogenen Sachunterrichts an Grundschulen in Deutschland. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 187–208). Waxmann. https://doi.org/10.31244/9783830993193

Steffensky, M., & Neuhaus, B. J. (2018). Unterrichtsqualität im naturwissenschaftlichen Unterricht. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Theorien in der naturwissenschaftsdidaktischen Forschung* (S. 299–313). Springer. https://doi.org/10.1007/978-3-662-56320-5_18

von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U., & Yin, L. (2024). *TIMSS 2023 international results in mathematics and science*. TIMSS & PIRLS International Study Center.

Voss, T., & Wittwer, J. (2020). Unterricht in Zeiten von Corona: Ein Blick auf die Herausforderungen aus der Sicht von Unterrichts- und Instruktionsforschung. *Unterrichtswissenschaft, 48,* 601–627. https://doi.org/10.1007/s42010-020-00088-2

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, *63*(3), 249–294. https://doi.org/10.3102/00346543063003249

Weinert, F. E., Schrader, F.-W., & Helmke, A. (1989). Quality of instruction and achievement outcomes. *International Journal of Educational Research*, *13*(8), 895–914. https://doi.org/10.1016/0883-0355(89)90072-4

Kapitel 11
Privater Nachhilfeunterricht und Förderungen zum Ausgleich pandemiebedingter Lernrückstände am Ende der Grundschulzeit

Lisa Benckwitz, Karin Guill und Christin Beese

11.1 Einleitung

Neben dem regulären Schulunterricht nehmen viele Schüler:innen in Deutschland privaten Nachhilfeunterricht in Anspruch (Hille et al., 2016). Dieser zeichnet sich dadurch aus, dass Schüler:innen regelmäßigen, auf sie abgestimmten Unterricht außerhalb der eigenen Familie und außerhalb des Schulunterrichts erhalten, der häufig von Lehrkräften, Studierenden oder älteren Schüler:innen gegeben wird. Nachhilfeunterricht ist mit dem Ziel verbunden, den akademischen Lernerfolg der Schüler:innen abzusichern (Krüger, 1977; Park et al., 2016).

Die Ergebnisse aus der *Trends in International Mathematics and Science Study* (TIMSS) 2015 und 2019 weisen darauf hin, dass am Ende der Grundschulzeit nur wenige Schüler:innen in Deutschland Nachhilfeunterricht nutzen (Guill & Wendt, 2016, 2020). Allerdings zeigen Forschungssynthesen zu den Folgen der Schulschließungen im Rahmen der COVID-19-Pandemie, dass vor allem die Leistung von Grundschüler:innen negativ durch die Schulschließungen beeinflusst wurde (Hammerstein et al., 2021; Helm, Huber & Postlbauer, 2021; König & Frey, 2022). Diese pandemiebedingten Lernrückstände könnten möglicherweise dazu geführt haben, dass mehr Familien privaten Nachhilfeunterricht in Anspruch nehmen. Folglich wird in diesem Kapitel untersucht, ob der Anteil an Grundschüler:innen, die privaten Nachhilfeunterricht nutzen, in TIMSS 2023 angestiegen ist. Darüber hinaus werden pandemiebezogene Prädiktoren der Nutzung von Nachhilfe betrachtet, um zu überprüfen, ob sich Hinweise auf einen Zusammenhang zwischen pandemiebezogener Schulschließung und Nutzung von privatem Nachhilfeunterricht zeigen.

Offen ist jedoch die Frage, ob privat bezahlter Nachhilfeunterricht im Vergleich zu staatlich finanzierten Förderprogrammen eine eher untergeordnete Rolle im Ausgleich negativer Folgen der COVID-19-Pandemie für Schüler:innen spielt. So hat z. B. die Bundesregierung das Aktionsprogramm *Aufholen nach Corona*

für die Jahre 2021 und 2022 beschlossen (BMFSFJ, 2022), um die sozial-emotionalen und kognitiven Folgen der pandemiebedingten Schulschließungen einzugrenzen (siehe auch Kapitel 2 in diesem Band). Im Rahmen des Programms wurden 2 Milliarden Euro für Projekte zur Förderung der sozial-emotionalen Entwicklung (z. B. Selbstwirksamkeitstrainings) und Fördermaßnahmen zum Abbau pandemiebedingter Lernrückstände bereitgestellt (z. B. Nachhilfe in den Ferien). Die *Ständige Wissenschaftliche Kommission der Kultusministerkonferenz* (SWK, 2021) empfiehlt, dass die im Rahmen der Aufholprogramme zur Verfügung gestellten Ressourcen gezielt an Bildungsübergängen eingesetzt werden, also auch am Übergang von der Grundschule in die Sekundarstufe I. Folglich untersucht dieses Kapitel den Anteil der Schüler:innen am Ende der Grundschulzeit, die an verschiedenen schulischen und außerschulischen Förderungen zum Ausgleich pandemiebedingter Lernrückstände teilgenommen haben (z. B. unterrichtsbegleitende Fördermaßnahmen).[1]

Die SWK (2021) empfiehlt weiterhin, dass sich Förderprogramme vor allem auf besonders stark von der COVID-19-Pandemie betroffene Schüler:innen fokussieren sollten, z. B. auf Schüler:innen, die lediglich die unteren Kompetenzstufen erreichen oder sich in benachteiligten schulischen Lernkontexten befinden. Es ist bisher allerdings nicht empirisch geklärt, inwiefern jene Schüler:innen tatsächlich von den Fördermaßnahmen erreicht wurden, da bei der Umsetzung von *Aufholen nach Corona* auf eine standardisierte, bundesweite Lernstandserhebung verzichtet werden musste (SWK, 2024). Den Schulen und Lehrkräften wurden zwar diagnostische Verfahren zur Verfügung gestellt, jedoch ist unklar, inwiefern diese tatsächlich eingesetzt wurden, um Schüler:innen mit besonders hohem Unterstützungsbedarf zu identifizieren (Helbig et al., 2022). Dementsprechend untersucht dieses Kapitel außerdem, inwiefern Schüler:innen mit besonders hohem Unterstützungsbedarf tatsächlich von den Aufholprogrammen erreicht wurden.

In diesem Kapitel werden somit die Verbreitung und die möglichen Prädiktoren der Teilnahme an privatem Nachhilfeunterricht sowie an Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände in den Blick genommen. Der zweite Abschnitt dieses Kapitels bildet Trends in der Nutzung von privatem Nachhilfeunterricht auf Grundlage der Daten aus der *Internationalen Grundschul-Lese-Untersuchung* (IGLU) 2006 sowie TIMSS 2011 bis TIMSS 2023 ab. Als potenzielle Prädiktoren der Inanspruchnahme von privatem Nachhilfeunterricht werden zum einen im Anschluss an die Untersuchungen in TIMSS 2015 (Guill & Wendt, 2016) und TIMSS 2019 (Guill & Wendt, 2020) die Leistung der Grundschüler:innen sowie die *Armutsgefährdung* der Familien betrachtet. Zum anderen werden pandemiebedingte Faktoren (subjektiv wahrgenommene Lernrückstände durch die COVID-19-Pandemie und Länge der Schulschließungen) berücksichtigt, die in TIMSS 2023 erstmals erhoben wurden. Im dritten Abschnitt dieses Kapitels wird die Teilnahme an Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände abgebildet. Die Darstellung erfolgt hier nicht im Trend, sondern nur für den aktuellen Zyklus TIMSS 2023, da die Teilnahme an Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände in TIMSS 2023 erstmals erhoben wurde. Als mögliche Prädiktoren der Nutzung von För-

1 Neben schulisch organisierten Angeboten wurden im Kontext von *Aufholen nach Corona* Gutscheine für privaten Nachhilfeunterricht über die Schulen nach Bedarfsklärung mit den Erziehungsberechtigten ausgegeben. Dementsprechend ist es nicht möglich, eindeutig zwischen privater Nachhilfe und staatlich geförderter Nachhilfe zum Ausgleich pandemiebedingter Lernrückstände zu unterscheiden.

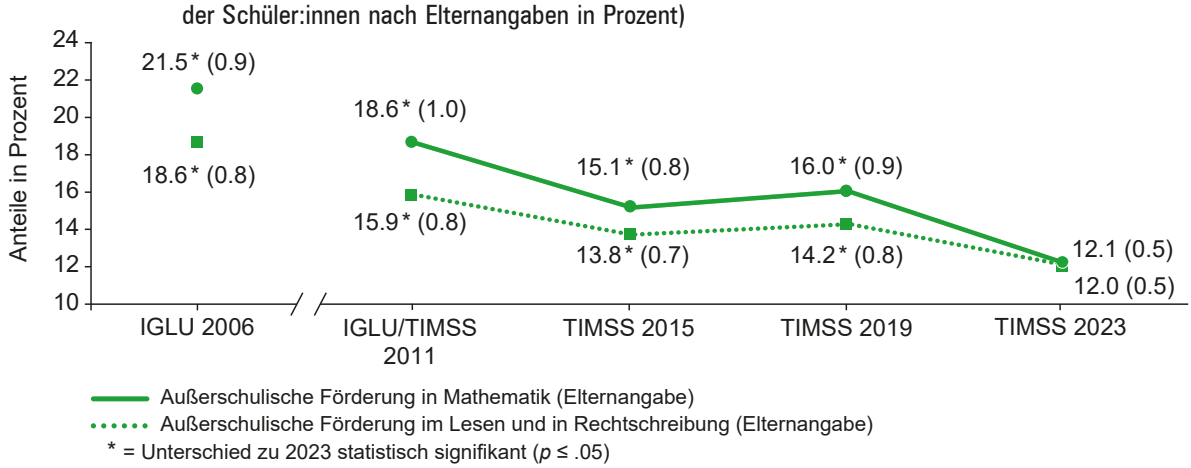
derprogrammen werden auf Grundlage der Empfehlungen der SWK (2021) neben den Kompetenzen der Schüler:innen Indikatoren der *Armutsgefährdung* sowie die Länge der Schulschließungen untersucht.

11.2 Verbreitung von Nachhilfeunterricht

In der Verbreitung von Nachhilfeunterricht zeigen sich weltweit große Unterschiede (Guill & Wendt, 2020; Klemm & Hollenbach-Biele, 2016). Obwohl die Nutzung in Deutschland im OECD-Vergleich eher gering ausfällt, zeigt sich seit Beginn der 2000er-Jahre auch in Deutschland eine steigende Anzahl von Schüler:innen, die Nachhilfe in Anspruch nehmen. So bilden die Daten des *Sozio-oekonomischen Panels* (SOEP) ab, dass immer mehr 17-Jährige angeben, mindestens einmal in ihrer Schullaufbahn privaten Nachhilfeunterricht in Anspruch genommen zu haben (Hille et al., 2016).

In der Gruppe der Grundschüler:innen zeigt sich allerdings ein gegenläufiger Trend, der in Abbildung 11.1 dargestellt wird. Die Daten von IGLU 2006 sowie TIMSS 2011 bis TIMSS 2023 bilden eine abnehmende Inanspruchnahme von Nachhilfeunterricht in der vierten Klasse ab. In IGLU 2006 nahmen nach Angaben der Erziehungsberechtigten 21.5 bzw. 18.6 Prozent der Schüler:innen Nachhilfe in Mathematik bzw. in Lesen und Rechtschreibung in Anspruch. Im Vergleich zu IGLU 2006 zeigte sich in TIMSS 2011 und 2015 eine rückläufige Tendenz, während sich die Teilnahmequote in TIMSS 2019 stabilisierte. Hier gaben 16.0 bzw. 14.2 Prozent Nachhilfe in Mathematik bzw. in Lesen und Rechtschreibung an. Durch die pandemiebedingten Schulschließungen, die zwischen 2020 und 2022 stattfanden, kam es besonders bei Grundschüler:innen zu Lernrückständen (Hammerstein et al., 2021; Helm, Huber & Postlbauer, 2021; König & Frey, 2022). Da das vorrangige Ziel von Nachhilfeunterricht die Sicherung des Lernerfolgs ist (Krüger, 1977), könnten die pandemiebedingten Lernrückstände zu einer erhöhten Nachfrage nach privater Nachhilfe geführt haben. Entgegen dieser Vermutung zeigt sich in TIMSS 2023 allerdings ein signifikanter Rückgang in der Inanspruchnahme von Nachhilfe seit TIMSS 2019; so geben 12.1 bzw. 12.0 Prozent der Erziehungsberechtigten Nachhilfe in Mathematik bzw. in Lesen und Rechtschreibung in den letzten zwölf Monaten vor der Befragung an.

Abbildung 11.1: Außerschulische Förderung im Vergleich von IGLU 2006 und TIMSS 2011 bis TIMSS 2023 (Anteil der Schüler:innen nach Elternangaben in Prozent)



IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

Für die abnehmende Nutzung von Nachhilfeunterricht am Ende der Grundschulzeit können unterschiedliche Erklärungsansätze herangezogen werden. Zum einen könnte die Umstrukturierung des Schulsystems in einer zunehmenden Anzahl von Ländern der Bundesrepublik Deutschland dazu geführt haben, dass weniger Viertklässler:innen Nachhilfeunterricht in Anspruch nehmen. Durch den Ausbau von Schulen mit mehreren Bildungsgängen haben mehr Schüler:innen die Möglichkeit, später auf den gymnasialen Zweig zu wechseln. Zusammen mit der zunehmenden Abschaffung der verbindlichen Übergangsempfehlung am Ende der Grundschulzeit könnte dies dazu beigetragen haben, dass Familien seltener auf Nachhilfeunterricht zurückgreifen, um eine Gymnasialempfehlung zu erhalten. Diese Annahme wird durch eine Studie von Guill und Lintorf (2019) gestützt, die zeigt, dass in Ländern mit verbindlichen Übergangsempfehlungen (siehe Kapitel 12 in diesem Band) Schüler:innen eher Nachhilfeunterricht nutzen. Zum anderen ist es denkbar, dass der voranschreitende Ausbau von Ganztagsangeboten zu einer sinkenden Nachfrage nach privater Nachhilfe geführt hat. Ein häufiger Fokus von Nachhilfeunterricht in Deutschland ist beispielsweise die Erledigung von Hausaufgaben (Guill et al., 2021). Da die meisten Ganztagsschulen eine Hausaufgabenbetreuung am Nachmittag anbieten (StEG-Konsortium, 2019), könnten viele Familien diese Angebote als Alternative zur Hausaufgabenbetreuung durch eine Nachhilfelehrkraft ansehen.

Da privater Nachhilfeunterricht in der Regel mit Kosten verbunden ist, ist es außerdem denkbar, dass Familien durch steigende Lebenshaltungskosten weniger finanzielle Ressourcen zur Verfügung stehen und sie deshalb auf kostenfreie oder kostengünstigere Alternativen (z. B. Nachhilfe durch Eltern oder ältere Geschwister, Lern-Apps) zurückgreifen. Ein Umstieg auf Alternativen zum privaten Nachhilfeunterricht könnte auch dadurch begünstigt worden sein, dass Familien durch die pandemiebedingten Kontaktbeschränkungen nach anderen Fördermöglichkeiten suchen mussten. Da sich die Angabe zum Besuch von Nachhilfeunterricht in TIMSS 2023 auf einen Zeitraum bezieht, in dem teilweise noch Kontaktbeschränkungen herrschten (Frühjahr 2022 bis Frühjahr 2023), haben viele Familien zu diesem Zeitpunkt möglicherweise weiterhin Alternativen zum Besuch von Nachhilfeunterricht genutzt. Bei anderen außerfamiliären Aktivitäten, wie z. B. der Mitgliedschaft in Sportvereinen, zeigt sich allerdings, dass die Teilnahmequoten von Kindern und Jugendlichen nach einem Tief während der Pandemie im Jahr 2023 wieder anstiegen (DOSB, 2020, 2021, 2022, 2023). Dementsprechend ist es möglich, dass die Teilnahmequote an privatem Nachhilfeunterricht in den nächsten Jahren ebenfalls wieder das Niveau von vor der COVID-19-Pandemie erreicht.

11.2.1 Wissen der Lehrkräfte um den Besuch von privatem Nachhilfeunterricht

In TIMSS 2019 wurde erstmals untersucht, ob die Lehrkräfte über die Nutzung von privatem Nachhilfeunterricht informiert sind (Guill & Wendt, 2020). Dies ist insofern relevant, weil das Wissen der Lehrkräfte um den Besuch von privater Nachhilfe eine Voraussetzung dafür darstellt, dass sich Lehrkräfte und Nachhilfelehrkräfte bezüglich der Lerninhalte und der spezifischen Lerndefizite des jeweiligen Kindes absprechen können. Im ungünstigeren Fall könnte das Wissen der Lehrkräfte darüber, dass Schüler:innen Nachhilfeunterricht besuchen, allerdings auch zu höheren Erwartungen an die Schulleistungen führen und somit Leistungs-

druck bei den Schüler:innen verursachen (Benckwitz et al., 2022; Bray, 2003; Hajar, 2018). In TIMSS 2019 zeigte sich, dass die Lehrkräfte bei Schüler:innen mit Nachhilfeunterricht zu großen Teilen entweder davon ausgingen, dass kein Nachhilfeunterricht genutzt wurde oder dass sie dazu keine Aussage treffen konnten (Guill & Wendt, 2020). Dieses Bild zeigt sich ebenfalls in TIMSS 2023. Wenn Kinder nach Angabe ihrer Eltern private Nachhilfe in Mathematik nutzen, sind die Lehrkräfte in 25 Prozent der Fälle informiert (siehe Tabelle 11.1). Hingegen gehen sie in 46 Prozent der Fälle fälschlicherweise davon aus, dass keine Nachhilfe in Anspruch genommen wird. In 28 Prozent der Fälle mit Nachhilfe ist es den Lehrkräften nicht bekannt, ob Nachhilfeunterricht genutzt wird und 1 Prozent macht dazu keine Angabe. Hinsichtlich der Nachhilfe in Deutsch ist das Befundmuster nahezu identisch. Insgesamt kann es daher wie in TIMSS 2019 als bislang zu wenig genutzte Chance gesehen werden, Lehrkräfte über die Inanspruchnahme von Nachhilfe zu informieren, um einen Austausch zwischen Lehrkräften und Nachhilfelehrkräften zu ermöglichen (Guill & Wendt, 2020).

Tabelle 11.1: Übereinstimmung der Angaben von Erziehungsberechtigten und Lehrkräften zur Nutzung von privatem Nachhilfeunterricht in Mathematik und Deutsch (Lesen und Rechtschreibung)

	Mathematik					Deutsch				
Lehrkräfteangabe / **Elternangabe**	Ja, Nachhilfe (%)	Keine Nachhilfe (%)	Nicht bekannt (%)	Keine Angabe (%)	**Summe (*n*)**	Ja, Nachhilfe (%)	Keine Nachhilfe (%)	Nicht bekannt (%)	Keine Angabe (%)	**Summe (*n*)**
Ja, Nachhilfe (%)	25	46	28	1	**312**	25	45	29	1	**326**
Nein, keine Nachhilfe (%)	1	62	35	1	**2237**	2	62	35	1	**2212**
Kein Fragebogen ausgefüllt/Keine Angabe (%)	6	56	36	3	**1893**	8	54	36	3	**1905**
Gesamt	**5**	**58**	**35**	**2**	**4442**	**6**	**57**	**35**	**2**	**4442**

Differenzen zu 100 Prozent und zur Summe (*n*) ergeben sich durch das Rundungsverfahren der gewichteten Daten.

Lesebeispiel: Wenn Kinder nach Angaben ihrer Erziehungsberechtigten Nachhilfe in Mathematik erhalten, sind Lehrkräfte in 25 Prozent dieser Fälle informiert, in 46 Prozent der Fälle gehen sie davon aus, dass das Kind keine Nachhilfe erhält, in 28 Prozent der Fälle ist es ihnen nicht bekannt und 1 Prozent hat dazu keine Angabe gemacht.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

11.2.2 Differenzierte Betrachtung der Nutzung von Nachhilfe

Studien zeigen, dass individuelle Merkmale von Schüler:innen mit der Nutzung von privatem Nachhilfeunterricht zusammenhängen (im Überblick Park et al., 2016). In diesem Abschnitt werden im Anschluss an TIMSS 2015 (Guill & Wendt, 2016) und TIMSS 2019 (Guill & Wendt, 2020) die Schulleistung und die *Armutsgefährdung* sowie pandemiebedingte Faktoren, die in TIMSS 2023 erstmals erhoben wurden, untersucht. In Abbildung 11.2 werden die Teilnahmequoten an privatem Nachhilfeunterricht in Mathematik differenziert nach diesen Faktoren dargestellt.

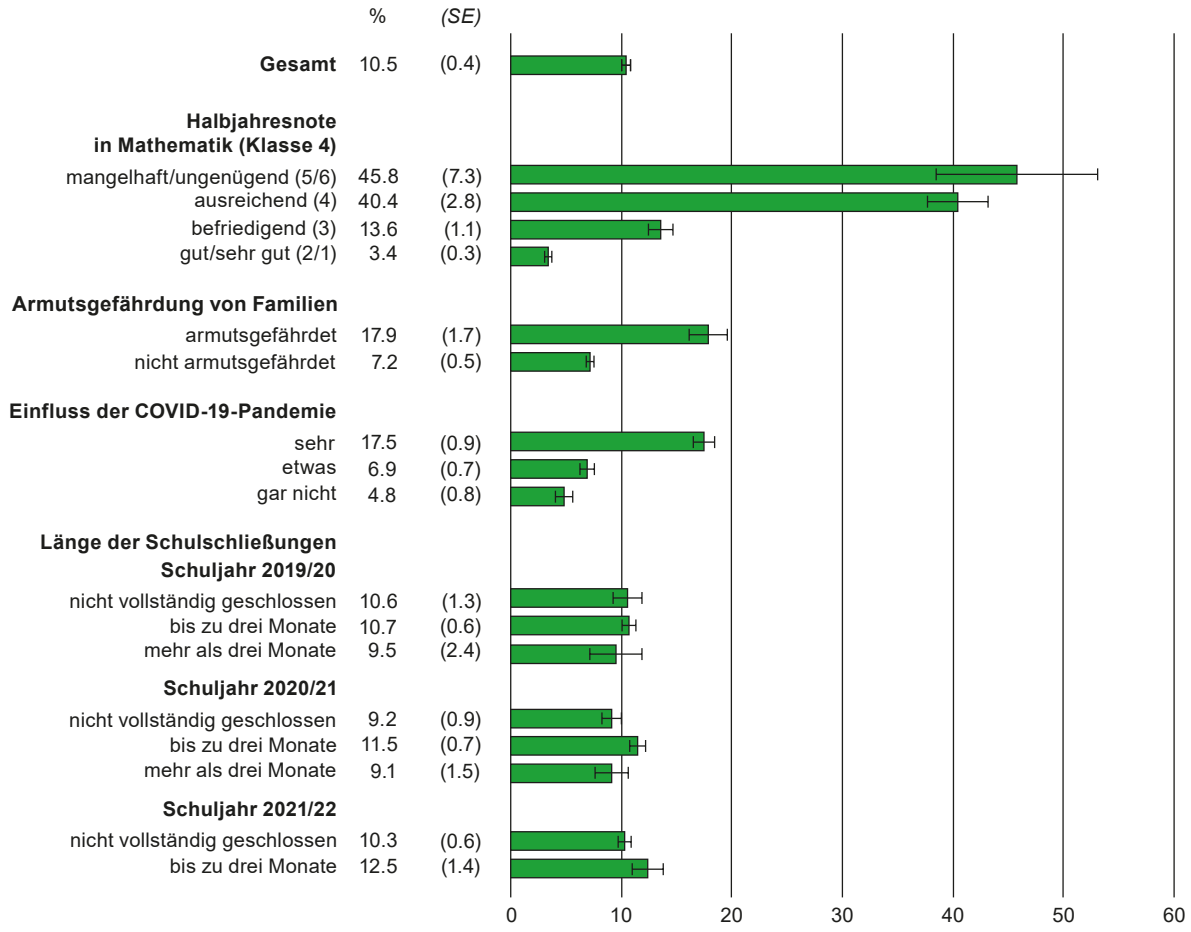
Obwohl Nachhilfeunterricht primär bei unzureichenden Schulleistungen in Anspruch genommen wird, zeigt sich in verschiedenen Studien ein substanzieller Anteil an Schüler:innen, die Nachhilfe nutzen, um besonders gute Noten zu erzielen (Hille et al., 2016; Klemm & Hollenbach-Biele, 2016). In TIMSS 2015 und 2019 nutzten am Ende der Grundschule allerdings vor allem Schüler:in-

nen mit mangelhaften oder ungenügenden Zeugnisnoten Nachhilfeunterricht in Mathematik (Guill & Wendt, 2016, 2020). In TIMSS 2023 zeichnet sich ein vergleichbares Bild ab (siehe Abbildung 11.2). Unter den Schüler:innen mit mangelhaften bzw. ungenügenden Halbjahresnoten in Mathematik nehmen 45.8 Prozent Nachhilfe in Anspruch. Bei ausreichenden Schulnoten beträgt die Teilnahmequote 40.4 Prozent, während bei befriedigenden Noten 13.6 Prozent Nachhilfe nutzen. Der Anteil an Schüler:innen mit Nachhilfeunterricht bei guten oder sehr guten Noten fällt sehr gering aus (3.4 %). Insgesamt deuten die Ergebnisse von TIMSS 2015 bis TIMSS 2023 darauf hin, dass Nachhilfeunterricht in der vierten Klasse vor allem genutzt wird, um unzureichende Schulnoten zu verbessern.

Die Nutzung von privatem Nachhilfeunterricht wird oft im Zusammenhang mit dem sozioökonomischen Status diskutiert (Entrich & Lauterbach, 2023; Park et al., 2016). Hierbei wird angenommen, dass Schüler:innen aus benachteiligten sozialen Lagen einen erschwerten Zugang zu privatem Nachhilfeunterricht haben. Dies betrifft finanzielle Möglichkeiten, aber auch die Kapazitäten von Eltern, außerunterrichtliche Unterstützung zu organisieren. Studien zeigen allerdings sowohl signifikant positive als auch nicht signifikante und signifikant negative Zusammenhänge zwischen dem sozioökonomischen Status und der Nutzung von Nachhilfe (Entrich, 2021; Guill & Lintorf, 2019; Hille et al., 2016; Luplow & Schneider, 2014). In TIMSS 2015 und TIMSS 2019 zeigte sich nicht, dass Schüler:innen aus armutsgefährdeten Familien seltener Nachhilfe in Anspruch nahmen. Ein vergleichbares Bild zeichnet sich in TIMSS 2023 ab (siehe Abbildung 11.2). Während 7.2 Prozent aller Schüler:innen aus Familien ohne *Armutsgefährdung* Nachhilfe nutzen, berichten 17.9 Prozent der Schüler:innen aus armutsgefährdeten Familien die Nutzung von Nachhilfe. Die höhere Teilnahmequote unter den armutsgefährdeten Schüler:innen wird auch dadurch bedingt, dass es in Deutschland weiterhin einen starken Zusammenhang zwischen dem familiären Hintergrund und der Schulleistung gibt (Stanat et al., 2022; Stubbe et al., 2020; siehe auch Kapitel 7 in diesem Band). Dementsprechend sind leistungsschwächere Schüler:innen in armutsgefährdeten Familien überrepräsentiert und es ist nicht auszuschließen, dass bei gleicher Schulleistung armutsgefährdete Schüler:innen seltener Nachhilfe in Anspruch nehmen als Schüler:innen ohne *Armutsgefährdung*. Zusätzlich ist unklar, ob Schüler:innen aus armutsgefährdeten Familien Nachhilfe in derselben Quantität und Qualität erhalten wie Schüler:innen, die nicht von *Armutsgefährdung* betroffen sind (Guill & Wendt, 2020). Darüber hinaus sind einkommensschwächere Familien relativ zu ihrem Einkommen stärker finanziell durch Nachhilfe belastet (Schröder et al., 2015). Dementsprechend ist es weiterhin erstrebenswert, dass niedrigschwellige Förderangebote gemacht werden, zu denen Schüler:innen unabhängig von den familiären Ressourcen Zugang haben.

Ein mögliches Motiv zur Inanspruchnahme von Nachhilfeunterricht ist der Abbau von Lernlücken, die durch die Schulschließungen während der COVID-19-Pandemie entstanden sind. In TIMSS 2023 wurden die Erziehungsberechtigten gefragt, wie stark ihrer Meinung nach die Pandemie den Lernfortschritt ihres Kindes negativ beeinflusst hat. Die Daten weisen darauf hin, dass die subjektive Wahrnehmung einer starken Beeinträchtigung durch die Pandemie mit der Nutzung von Nachhilfe einhergeht. So nutzen 17.5 Prozent der Schüler:innen Nachhilfe, deren Eltern einen *sehr* starken negativen Einfluss der Pandemie auf den Lernfortschritt berichten (siehe Abbildung 11.2). Unter den Schüler:innen, die die Pandemie *etwas* beeinflusst hat, nehmen 6.9 Prozent private Nachhilfe in

Abbildung 11.2: Privater Nachhilfeunterricht in Mathematik differenziert nach Halbjahresnote, *Armutsgefährdung* und pandemiebedingten Faktoren



IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Anspruch und es gibt auch unter den Schüler:innen, die die Pandemie *gar nicht* beeinflusst hat, einen Anteil mit Nachhilfeunterricht (4.8 %).

Analysen von Wößmann et al. (2020, 2021) zeigen, dass sich die Lernzeit von Schüler:innen während der Schulschließungen deutlich reduziert hat. Dementsprechend ist es denkbar, dass Schüler:innen eher Nachhilfeunterricht in Anspruch nehmen, wenn ihre Schulen länger geschlossen waren. Diese Tendenz bildet sich in TIMSS 2023 allerdings nicht ab (siehe Abbildung 11.2). Beispielsweise nahmen im Schuljahr 2020/21 9.2 Prozent der Schüler:innen, deren Schulen nicht vollständig geschlossen waren, Nachhilfe in Anspruch. Unter den Schüler:innen, deren Schulen bis zu drei Monate bzw. mehr als drei Monate geschlossen waren, sind die Teilnahmequoten vergleichbar (11.5 % bzw. 9.1 %). Dass sich kein klarer Zusammenhang zwischen der Länge der Schulschließungen und der Nutzung von Nachhilfe zeigt, kann auf unterschiedliche Faktoren zurückgeführt werden. So gibt es individuelle Unterschiede darin, wie Schüler:innen auf die pandemiebedingten Schulschließungen reagiert haben. Beispielsweise zeigten Wößmann et al. (2021), dass leistungsschwache Schüler:innen und Kinder aus nicht akademischen Haushalten zu Hause weniger effektiv gelernt haben. Darüber hinaus gab es Unterschiede darin, wie effizient der Distanzunterricht gestaltet wurde. Zum Beispiel bewerteten Schüler:innen, Eltern, Lehrkräfte und

Schulleitungen die Qualität des Distanzunterrichts unterschiedlich (Köller et al., 2021). Folglich könnten individuelle Merkmale der Schüler:innen sowie die Qualität des Distanzunterrichts für die Entscheidung, Nachhilfe zu nutzen, bedeutsamer sein als die Länge der Schulschließungen.

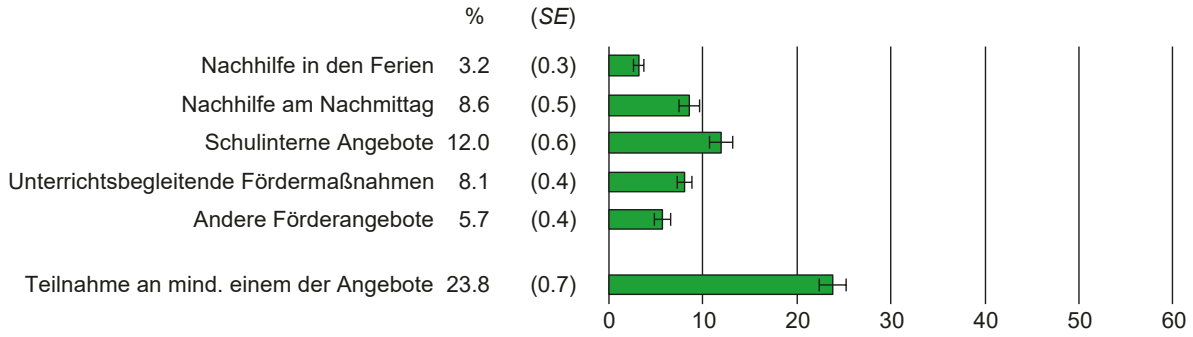
11.3 Teilnahme an Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände

Zum Ausgleich der durch die COVID-19-Pandemie entstandenen Lern- und Entwicklungsrückstände wurden verschiedene Förderprogramme initiiert (z. B. *Aufholen nach Corona*). Um ein erfolgreiches Weiterlernen der Schüler:innen nach der Pandemie sicherzustellen, sollten den Schüler:innen passgenaue, bedarfsorientierte Angebote gemacht werden, die sich an den Empfehlungen der SWK (2021) orientieren. Es wurden bestehende Förderprogramme weiterentwickelt sowie neue Angebote erstellt, deren Durchführung von qualifiziertem Personal übernommen und mit dem Regelunterricht abgestimmt werden sollte. Die Implementation der Angebote war auf unterschiedlichen Ebenen angesetzt (SWK, 2021, 2024):

- Außerunterrichtliche Angebote, z. B. Nachhilfe in den Ferien in Sommercamps, Lernwerkstätten oder Nachhilfe am Nachmittag, um Unterrichtsinhalte nachzuholen
- Unterrichtsergänzende Angebote, z. B. schulinterne Förderkurse am Nachmittag in den Kernfächern wie Mathematik und Sprachen
- Unterrichtsbegleitende Fördermaßnahmen, z. B. Schulbegleitung durch pädagogische Fachkräfte, Studierende oder Freiwilligendienstleistende

In TIMSS 2023 wurde die Teilnahme an Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände erfragt. Insgesamt haben 23.8 Prozent der Grundschüler:innen an mindestens einem Angebot teilgenommen (siehe Abbildung 11.3). Dabei haben die meisten Schüler:innen (12.0 %) schulinterne Angebote wie etwa Förderkurse in Anspruch genommen, gefolgt von Nachhilfe am Nachmittag (8.6 %) und unterrichtsbegleitenden Maßnahmen (z. B. Schulbegleitung: 8.1 %). Eine relativ geringe Anzahl an Schüler:innen hat Nachhilfe in den Ferien besucht (3.2 %). Darüber hinaus haben 5.7 Prozent andere Angebote genutzt.

Abbildung 11.3: Teilnahme an Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände differenziert nach Angeboten



IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

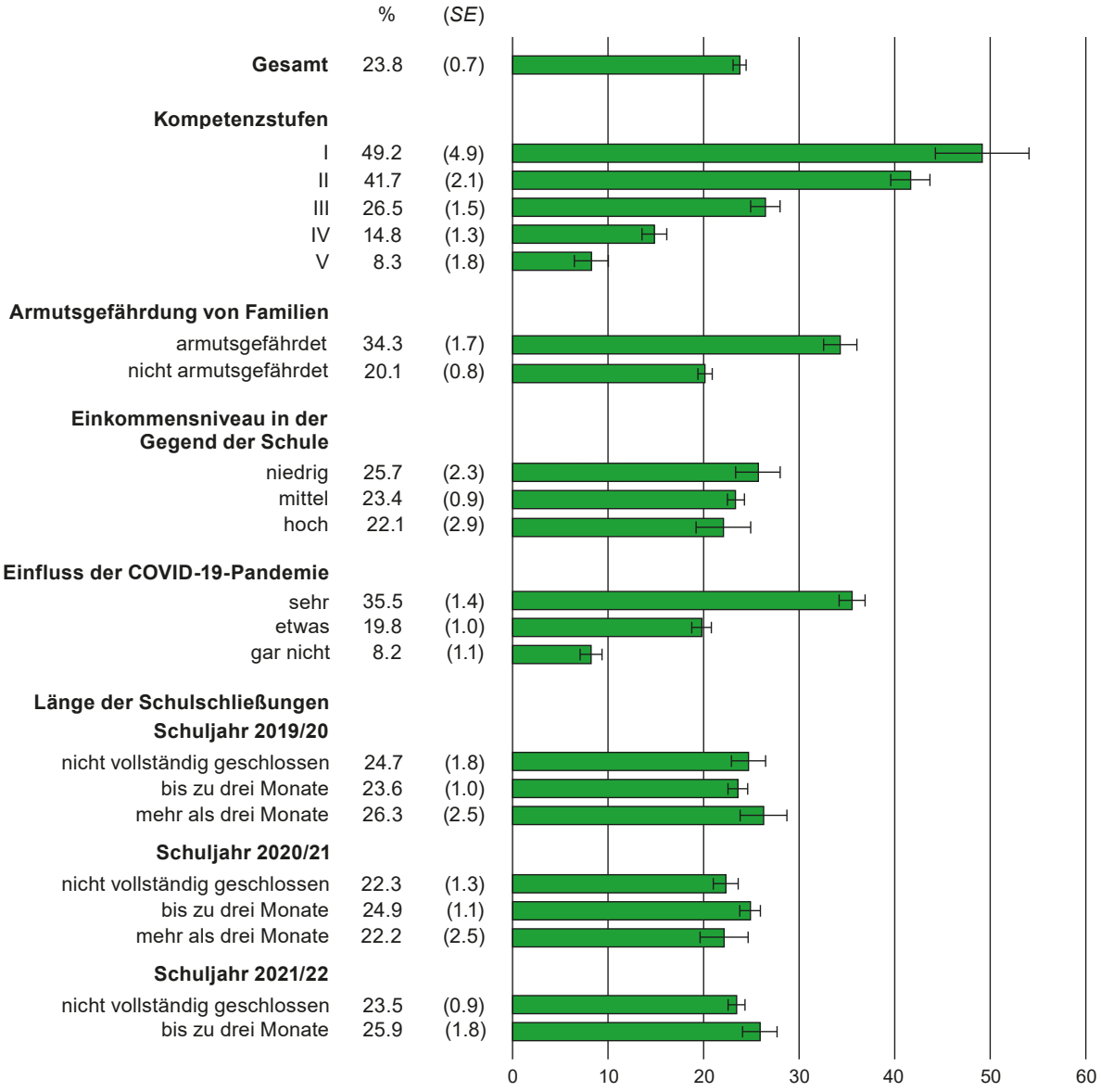
Bei der Umsetzung der Förderprogramme sollte der Fokus nach den Empfehlungen der SWK (2021) auf besonders betroffene und vulnerable Gruppen von Schüler:innen gelegt werden. Entsprechend der Empfehlungen werden in diesem Abschnitt die Schulleistung, Indikatoren der *Armutsgefährdung* sowie pandemiebedingte Faktoren als potenzielle Prädiktoren der Teilnahme an Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände betrachtet. In Abbildung 11.4 wird die Teilnahme an den Programmen differenziert nach diesen Faktoren dargestellt.

In TIMSS 2023 kann erstmals untersucht werden, inwiefern die Schüler:innen in den unteren Kompetenzstufen tatsächlich von den Förderprogrammen zum Ausgleich der Lernrückstände erreicht wurden. Abbildung 11.4 zeigt, dass sich unter den Schüler:innen in den Kompetenzstufen I und II mit 49.2 bzw. 41.7 Prozent tatsächlich die höchsten Teilnahmequoten zeigten. Allerdings weisen die Zahlen auch auf eine große Gruppe von Schüler:innen in den unteren Kompetenzstufen hin, die keine Förderung erhalten haben, obwohl bei diesen Schüler:innen das Risiko besteht, dass sich die bestehenden Lernrückstände weiter akkumulieren. In Kompetenzstufe III nahmen 26.5 Prozent aller Schüler:innen an Förderprogrammen teil und auch in den höchsten Kompetenzstufen (IV und V) gab es Schüler:innen mit Förderungen (14.8 % bzw. 8.3 %). Die Unterscheidung der Teilnahmequoten nach der elterlichen Wahrnehmung des Einflusses der Pandemie auf den Lernfortschritt der Kinder deutet ebenfalls darauf hin, dass auch Kinder, die nicht negativ von der Pandemie beeinflusst wurden, an Förderprogrammen teilgenommen haben. So war der prozentuale Anteil unter den Schüler:innen, die *sehr* oder *etwas* negativ durch die Pandemie beeinflusst wurden, zwar deutlich höher (35.5 % bzw. 19.8 %), allerdings nahmen auch 8.2 Prozent der Schüler:innen, die die Pandemie *gar nicht* negativ beeinflusst hatte, an Förderprogrammen teil. Es ist also möglich, dass die Ressourcen nicht optimal eingesetzt wurden. Allerdings ist auch denkbar, dass die elterliche Einschätzung über die Notwendigkeit einer Förderung nicht in jedem Fall mit der Einschätzung der Lehrkräfte übereinstimmte.

Schon vor der Pandemie bestand in Deutschland ein enger Zusammenhang zwischen dem sozioökonomischen Hintergrund von Schüler:innen und ihren Schulleistungen (Stanat et al., 2022; Stubbe et al., 2020). Zusätzlich ist es mittlerweile gut belegt, dass die Schulschließungen tendenziell belastender für sozial benachteiligte Schüler:innen waren (Engzell et al., 2021; Helm, Huber & Loisinger, 2021; Vodafone Stiftung Deutschland, 2020; Wößmann et al., 2020, 2021). Dementsprechend empfahl die SWK (2021), dass die Förderprogramme zum Ausgleich pandemiebedingter Lernrückstände vor allem in Bildungsinstitutionen in benachteiligten Kontexten implementiert werden sollten. Auf individueller Ebene zeigen die Daten aus TIMSS 2023, dass 20.1 Prozent der Schüler:innen ohne *Armutsgefährdung* an einer Förderung teilnahmen, während unter den Schüler:innen aus armutsgefährdeten Familien 34.3 Prozent Förderangebote in Anspruch nahmen. Da sich allerdings in armutsgefährdeten Familien leistungsschwächere Schüler:innen häufen (siehe auch Kapitel 7 in diesem Band), kann nicht eindeutig geklärt werden, welches Kriterium (d. h. *Armutsgefährdung* und/oder Leistungsschwäche) für die Teilnahme an Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände herangezogen wurde.

Auf Schulebene treten hingegen nur geringe Unterschiede in der Förderungsquote nach Einkommensniveau auf. So nahmen bei hohem bzw. mittlerem Einkommensniveau in der Gegend der Schule 22.1 bzw. 23.4 Prozent der Schüler:innen an Förderprogrammen teil, während die Quote bei niedrigem Ein-

Abbildung 11.4: Förderprogramme zum Ausgleich pandemiebedingter Lernrückstände differenziert nach Kompetenzstufen, *Armutsgefährdung* und pandemiebedingten Faktoren



IEA: Trends in International Mathematics and Science Study © TIMSS 2023

kommensniveau bei 25.7 Prozent lag. Die Daten auf Schulebene weisen also darauf hin, dass es kaum gelungen ist, Förderprogramme vorrangig in benachteiligten schulischen Kontexten zu implementieren. Dies ist unter anderem dadurch bedingt, dass in vielen Ländern der Bundesrepublik Deutschland keine geeigneten Daten und Instrumente (wie z. B. ein Sozialindex der Schulen) verfügbar sind, die eine zentrale Voraussetzung dafür darstellen, die Mittel gezielt in sozial benachteiligten Schulen einzusetzen (Helbig et al., 2022). Angesichts der Häufung leistungsschwacher Schüler:innen in sozial benachteiligten Schulkontexten könnten die Ergebnisse darauf hindeuten, dass vorhandene Lernstandserhebungen (z. B. *Vergleichsarbeiten*; VERA) nicht systematisch eingesetzt wurden, um Unterstützungsbedarfe zu identifizieren.

Da während der Schulschließungen die aktive Lernzeit der Schüler:innen abgenommen hat (Wößmann et al., 2020, 2021), hängt der Bedarf an Förderun-

gen zum Ausgleich pandemiebedingter Lernrückstände möglicherweise positiv mit der Länge der Schulschließungen zusammen. Die Daten aus TIMSS 2023 liefern allerdings keine Hinweise auf einen klaren Zusammenhang zwischen der Länge der Schulschließungen und der Teilnahme an Förderprogrammen. Im Schuljahr 2020/21 bekamen beispielsweise 22.3 Prozent der Schüler:innen, deren Schulen nicht vollständig geschlossen waren, eine Förderung zum Ausgleich der pandemiebedingten Lernrückstände. In den Schulen, die im Schuljahr 2020/21 bis zu drei Monate bzw. mehr als drei Monate geschlossen waren, lag der Anteil bei 24.9 bzw. 22.2 Prozent. Daher kann, ähnlich wie im Bereich der privaten Nachhilfe, davon ausgegangen werden, dass andere Prädiktoren (z. B. Schulleistungen, Qualität des Distanzunterrichts, Teilnahme an der Notbetreuung) eine entscheidendere Rolle für die Teilnahme an Förderprogrammen zum Abbau pandemiebedingter Lernrückstände spielten als die Länge der Schulschließungen.

11.4 Zusammenfassung

In Bezug auf die Nutzung privater Nachhilfe am Ende der Grundschulzeit zeigt dieses Kapitel, dass sich der Trend rückläufiger Teilnahmequoten auch in TIMSS 2023 fortsetzt. So nehmen in TIMSS 2023 nur 12.1 bzw. 12.0 Prozent der Viertklässler:innen Nachhilfe in Mathematik bzw. in Lesen und Rechtschreibung in Anspruch und damit signifikant weniger Schüler:innen als in TIMSS 2019. Ähnlich wie in TIMSS 2019 ist ein substanzieller Teil der Lehrkräfte nicht über den Nachhilfeunterricht der Schüler:innen informiert. Offen bleibt an dieser Stelle, welche Faktoren zum Rückgang der Nachhilfenutzung in der Grundschule beitragen.

Hinsichtlich der Nachhilfenutzung differenziert nach individuellen Merkmalen der Grundschüler:innen zeichnet sich ein vergleichbares Bild wie in TIMSS 2019 ab; am stärksten genutzt wird die Nachhilfe bei unzureichenden Noten und bei *Armutsgefährdung* der Familie. In Bezug auf pandemiebezogene Prädiktoren der Nutzung von Nachhilfe zeigt sich, dass vor allem bei der subjektiven Wahrnehmung eines starken negativen Effekts der Pandemie auf den Lernfortschritt Nachhilfe in Anspruch genommen wird. Die Länge der pandemiebedingten Schulschließungen scheint hingegen nur eine untergeordnete Rolle zu spielen.

Hinsichtlich der Nutzung von Förderprogrammen zum Ausgleich der pandemiebedingten Lernrückstände nahm knapp jedes vierte Kind mindestens ein Angebot in Anspruch. Bedauerlicherweise gibt es einen großen Anteil an Schüler:innen in den untersten Kompetenzstufen, die nicht von einem Förderprogramm erreicht wurden. Um ein erfolgreiches Lernen dieser Schüler:innen in der Sekundarstufe sicherzustellen, sollte gewährleistet werden, dass ihnen in der Zukunft niedrigschwellige Unterstützungsangebote gemacht werden.

In Bezug auf die Prädiktoren der Teilnahme an Förderprogrammen zum Ausgleich pandemiebedingter Lernrückstände zeigt sich ein ähnliches Bild wie bei der Nutzung von privatem Nachhilfeunterricht. Schüler:innen in den untersten Kompetenzstufen, aus armutsgefährdeten Familien und mit großen subjektiv wahrgenommenen Lernverlusten durch die Pandemie nahmen am häufigsten an den Förderprogrammen teil. Auf Schulebene weisen die Daten darauf hin, dass weder das Einkommensniveau in der Gegend der Schule noch die Länge der Schulschließungen eine zentrale Rolle für die Nutzung der Förderprogramme spielten. Dementsprechend deuten die Ergebnisse an, dass die Empfehlungen der SWK (2021) für die Implementation der Förderprogramme eher auf der Indi-

vidual- als auf der Schulebene befolgt wurden. Schließlich bleibt zwar offen, inwiefern die Teilnahme an den Förderprogrammen tatsächlich förderlich für die Schulleistungen war, allerdings könnte sie möglicherweise zu den gleichbleibenden Leistungswerten zwischen TIMSS 2019 und TIMSS 2023 beigetragen haben.

Literatur

Benckwitz, L., Guill, K., Roloff, J., Ömeroğulları, M., & Köller, O. (2022). Investigating the relationship between private tutoring, tutors' use of an individual frame of reference, reasons for private tutoring, and students' motivational-affective outcomes. *Learning and Individual Differences*, *95*, 102137. https://doi.org/10.1016/j.lindif.2022.102137

BMFSFJ (Bundesministerium für Familie, Senioren, Frauen und Jugend). (2022). *Aktionsprogramm „Aufholen nach Corona für Kinder und Jugendliche"*. https://www.bmfsfj.de/bmfsfj/aktuelles/alle-meldungen/aktionsprogramm-aufholen-nach-corona-fuer-kinder-und-jugendliche--178422

Bray, M. (2003). *Adverse effects of private supplementary tutoring: Dimensions, implications and government responses.* UNESCO International Institute for Educational Planning. https://etico.iiep.unesco.org/en/adverse-effects-private-supplementary-tutoring-dimensions-implications-and-government-responses

DOSB (Deutscher Olympischer Sportbund). (2020). *Bestandserhebung 2020. Korrigierte Fassung.* https://cdn.dosb.de/user_upload/www.dosb.de/uber_uns/Bestandserhebung/BE-Heft_2020.pdf

DOSB (Deutscher Olympischer Sportbund). (2021). *Bestandserhebung 2021.* https://cdn.dosb.de/user_upload/www.dosb.de/uber_uns/Bestandserhebung/BE-Heft_2021.pdf

DOSB (Deutscher Olympischer Sportbund). (2022). *Bestandserhebung 2022.* https://cdn.dosb.de/user_upload/www.dosb.de/uber_uns/Bestandserhebung/BE-Heft_2022.pdf

DOSB (Deutscher Olympischer Sportbund). (2023). *Bestandserhebung 2023.* https://cdn.dosb.de/user_upload/www.dosb.de/uber_uns/Bestandserhebung/Bestandserhebung_2023.pdf

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *PNAS, 118*(17). https://doi.org/10.1073/pnas.2022376118

Entrich, S. R. (2021). Worldwide shadow education and social inequality: Explaining differences in the socioeconomic gap in access to shadow education across 63 societies. *International Journal of Comparative Sociology, 61*(6), 441–475. https://doi.org/10.1177 %2F0020715220987861

Entrich, S. R., & Lauterbach, W. (2023). *German-style shadow education: The role of private tutoring for school allocation in a highly stratified education system*. https://doi.org/10.31235/osf.io/m9fe5

Guill, K., & Lintorf, K. (2019). Private tutoring when stakes are high: Insights from the transition from primary to secondary school in Germany. *International Journal of Educational Development, 65*, 172–182. https://doi.org/10.1016/j.ijedudev.2018.08.001

Guill, K., Ömeroğulları, M., & Köller, O. (2021). Intensity and content of private tutoring lessons during German secondary schooling: Effects on students' grades and test achievement. *European Journal of Psychology of Education*, *37*(4), 1093–1114. https://doi.org/10.1007/s10212-021-00581-x

Guill, K., & Wendt, H. (2016). Außerschulischer Nachhilfeunterricht am Ende der Grundschulzeit. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 247–256). Waxmann. https://doi.org/10.25656/01:14022

Guill, K., & Wendt, H. (2020). Privater Nachhilfeunterricht und Lehrkräftefortbildung am Ende der Grundschulzeit. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und natur-*

wissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich (S. 209–222). Waxmann. https://doi.org/10.31244/9783830993193

Hajar, A. (2018). Exploring Year 6 pupils' perceptions of private tutoring: Evidence from three mainstream schools in England. *Oxford Review of Education, 44*(4), 514–531. https://doi.org/10.1080/03054985.2018.1430563

Hammerstein, S., König, C., Dreisörner, T., & Frey, A. (2021). *Effects of COVID-19-related school closures on student achievement. A systematic review*. https://doi.org/10.31234/osf.io/mcnvk

Helbig, M., Edelstein, B., Fickermann, D., & Zink, C. (2022). *Aufholen nach Corona? Maßnahmen der Länder im Kontext des Aktionsprogramms von Bund und Ländern*. Waxmann. https://doi.org/10.31244/9783830996033

Helm, C., Huber, S. G., & Loisinger, T. (2021). Was wissen wir über schulische Lehr-Lern-Prozesse im Distanzunterricht während der Corona-Pandemie? – Evidenz aus Deutschland, Österreich und der Schweiz. *Zeitschrift für Erziehungswissenschaft, 24*, 237–311. https://doi.org/10.1007/s11618-021-01000-z

Helm, C., Huber, S. G., & Postlbauer, A. (2021). Lerneinbußen und Bildungsbenachteiligung durch Schulschließungen während der Covid-19-Pandemie im Frühjahr 2020. Eine Übersicht zur aktuellen Befundlage. In D. Fickermann, B. Edelstein, J. Gerick & K. Racherbäumer (Hrsg.), *Schule und Schulpolitik während der Corona-Pandemie: Nichts gelernt?* (S. 59–81). Waxmann. https://doi.org/10.31244/9783830994589.03

Hille, A., Spies, C. K., & Staneva, M. (2016). Immer mehr Schülerinnen und Schüler nehmen Nachhilfe, besonders in Haushalten mit mittleren Einkommen. *DIW Wochenbericht*, *6*, 111–120. https://www.econstor.eu/handle/10419/127439

Klemm, K., & Hollenbach-Biele, N. (2016). *Nachhilfeunterricht in Deutschland: Ausmaß – Wirkung – Kosten.* Bertelsmann Stiftung.

Köller, O., Jansen, T., & Meyer, J. (2021). *Distanzlernen in Schleswig-Holstein (DiScH). Eine empirische Studie zum Lernen in der Covid-19-Pandemie. Erste Ergebnisse aus der Befragung vom März/April 2021*. IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik. https://www.leibniz-ipn.de/en/the-ipn/archive/ipn_disch.pdf/@@download/file

König, C., & Frey, A. (2022). The impact of COVID-19-related school closures on student achievement – A meta-analysis. *Educational Measurement, 41*(1), 16–22. https://doi.org/10.1111/emip.12495

Krüger, R. (1977). Nachhilfe – Chance oder Skandal? 17 Antworten auf Fragen zu einem vernachlässigten Problem. *Die Deutsche Schule, 69*(9), 545–558.

Luplow, N., & Schneider, T. (2014). Nutzung und Effektivität privat bezahlter Nachhilfe im Primarbereich. *Zeitschrift für Soziologie, 43*(1), 31–49. https://doi.org/10.1515/zfsoz-2014-0104

Park, H., Buchmann, C., Choi, J., & Merry, J. J. (2016). Learning beyond the school walls: Trends and implications. *Annual Review of Sociology, 42*, 231–252. https://doi.org/10.1146/annurev-soc-081715-074341

Schröder, C., Spieß. C. K., & Storck, J. (2015). Private Bildungsausgaben für Kinder: Einkommensschwache Familien sind relativ stärker belastet. *DIW Wochenbericht, 8*, 158–169.

Stanat, P., Schipolowski, S., Schneider, R., Sachse, K. A., Weirich, S., & Henschel, S. (Hrsg.). (2022). *IQB-Bildungstrend 2021: Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich*. Waxmann. https://doi.org/10.31244/9783830996064

StEG-Konsortium. (2019). *Ganztagsschule 2017/2018. Deskriptive Befunde einer bundesweiten Befragung. Studie zur Entwicklung von Ganztagsschulen, StEG.* DIPF, DJI, IFS, Justus-Liebig-Universität. https://doi.org/10.25656/01:17105

Stubbe, T. C., Krieg, M., Beese, C., & Jusufi, D. (2020). Soziale Disparitäten in den mathematischen und naturwissenschaftlichen Kompetenzen von Viertklässlerinnen und Viertklässlern. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 263–289). Waxmann. https://doi.org/10.31244/9783830993193

SWK (Ständige Wissenschaftliche Kommission der Kultusministerkonferenz). (2021). *Pandemiebedingte Lernrückstände aufholen – Unterstützungsmaßnahmen fokussieren, verknüpfen und evaluieren.* https://www.kmk.org/fileadmin/Dateien/pdf/KMK/SWK/2021/2021_06_11-Pandemiebedingte-Lernruckstaende-aufholen.pdf

SWK (Ständige Wissenschaftliche Kommission der Kultusministerkonferenz). (2024). *Abschließendes Berichtsblatt zum Aktionsprogramm „Aufholen nach Corona für Kinder und Jugendliche" für die Jahre 2021 und 2022. Maßnahme „Abbau von Lernrückständen"*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2024/2024-Corona-abschl-Berichtsblatt.pdf

Vodafone Stiftung Deutschland (Hrsg.). (2020). *Unter Druck. Die Situation von Eltern und ihren schulpflichtigen Kindern während der Schulschließungen*. https://www.vodafone-stiftung.de/wp-content/uploads/2020/04/Vodafone-Stiftung-Deutschland_Studie_Unter_Druck.pdf

Wößmann, L., Freundl, V., Grewenig, E., Lergetporer, P., Werner, K., & Zierow, L. (2020). *Bildung in der Coronakrise: Wie haben die Schulkinder die Zeit der Schulschließungen verbracht, und welche Bildungsmaßnahmen befürworten die Deutschen?* ifo Schnelldienst, *73*, 25–39.

Wößmann, L., Freundl, V., Grewenig, E., Lergetporer, P., Werner, K., & Zierow, L. (2021). *Bildung erneut im Lockdown: Wie verbrachten Schulkinder die Schulschließungen Anfang 2021?* ifo Schnelldienst, *74*, 36–52.

Kapitel 12
Schullaufbahnpräferenzen am Übergang in die Sekundarstufe und der Zusammenhang mit leistungsrelevanten und sozialen Merkmalen

Tobias C. Stubbe, Annika Beier und Martje Lott

12.1 Einleitung

In Deutschland findet der Übergang von der Grundschule auf eine Schulform der Sekundarstufe I in den meisten Ländern der Bundesrepublik Deutschland nach dem vierten Grundschuljahr statt. Grundgedanke dieser institutionellen Leistungsdifferenzierung der Schüler:innen ist, dass der Unterricht in leistungshomogeneren Gruppen besser auf die Lernvoraussetzungen abgestimmt und somit eine optimale Förderung aller Schüler:innen erreicht werden kann (Baumert, 2006).

Diese im internationalen Vergleich frühe Differenzierung in unterschiedliche Bildungsgänge wird von vielen Expert:innen kritisiert (Maaz, 2020). Bei der Prognose der zukünftigen Leistungsfähigkeit eines Kindes in dieser Lebensphase ist mit Ungenauigkeiten zu rechnen, sodass Schullaufbahnempfehlungen nicht selten fehlerbehaftet sind (Cremer & Tobisch, 2021; Klapproth et al., 2013; Tiedemann & Billmann-Mahecha, 2010). Neben dem Problem der prognostischen Zuverlässigkeit weisen sowohl die Empfehlungen von Grundschullehrkräften als auch die Schulformentscheidungen von Erziehungsberechtigten deutliche Zusammenhänge mit dem sozialen Hintergrund und einer möglichen Migrationsgeschichte der Familie auf (im Überblick Maaz, 2020; Glock et al., 2015; van Leest et al., 2021). Befürworter:innen führen als Legitimierung dieser frühen Differenzierung nach dem Leistungsniveau der Kinder die grundsätzliche Möglichkeit der nachträglichen Korrektur der ursprünglichen Schullaufbahnentscheidung im Verlauf der Sekundarstufe an.

Das vorliegende Kapitel berichtet zunächst aktuelle Forschungsbefunde zu Schullaufbahnpräferenzen am Übergang in die Sekundarstufe und stellt anschließend die übergeordneten Fragestellungen dieses Kapitels vor (Abschnitt 12.2). Im Ergebnisteil werden zunächst deskriptiv die Schullaufbahnpräferenzen von Lehrkräften und Eltern bzw. Erziehungsberechtigten dargestellt (Abschnitt 12.3), um anschließend Befunde zum Zusammenhang zwischen den Schullaufbahnpräferen-

zen der Lehrkräfte und den leistungsrelevanten Merkmalen von Viertklässler:innen zu präsentieren (Abschnitt 12.4). Es folgen Ergebnisse zum Zusammenhang zwischen den Schullaufbahnpräferenzen von Lehrkräften bzw. Erziehungsberechtigten und sozialen Merkmalen der Schüler:innen (Abschnitt 12.5). Abschließend werden die zentralen Befunde dieses Kapitels zusammenfassend diskutiert (Abschnitt 12.6).

12.2 Theorie und Forschungsstand zu Schullaufbahnpräferenzen am Übergang in die Sekundarstufe

Unterschiede im Übergangsverhalten am Ende der Primarstufe lassen sich u. a. auf Leistungsdisparitäten zurückführen, die im Zusammenhang mit der familiären Herkunft stehen. Insbesondere die Gliederung des Sekundarschulbereiches in lehrplanmäßig unterschiedlich anspruchsvolle Bildungsprogramme verschärft den Einfluss der sozialen Herkunft auf die Schulleistungen. Dabei wirken zwei Mechanismen zusammen: Zum einen ist der Übergang in die verschiedenen Bildungsgänge des Sekundarschulbereiches mit dem sozialen Hintergrund assoziiert. Kinder aus bildungsfernen Familien haben selbst bei gleichen Testleistungen und Schulnoten eine geringere Chance, eine Gymnasialempfehlung zu erhalten, als Kinder aus bildungsnahen Familien (sekundäre Herkunftseffekte; siehe Kapitel 7 in diesem Band). Zum anderen entstehen durch diese Aufteilung relativ homogene Entwicklungsmilieus, d. h., Schüler:innen einer Klasse weisen einen ähnlichen sozialen Hintergrund und eine ähnliche Leistungsfähigkeit auf. Die daraus resultierenden Kompositionseffekte führen gepaart mit institutionellen Unterschieden (z. B. unterschiedlich anspruchsvolle Lehrpläne) zu differenziellen Lernzuwächsen in den einzelnen Bildungsgängen (Maaz, 2020). Es ergibt sich ein sogenannter *Schereneffekt*: Die Lernzuwächse auf dem Gymnasium fallen am stärksten, in der Hauptschule am geringsten aus (Maaz & Dumont, 2019). Je größer diese *Schereneffekte* in der Sekundarstufe ausfallen, desto stärker vergrößern sich die mittleren Leistungsunterschiede auf den verschiedenen Schulformen. Dieser Leistungsunterschied hat wiederum Auswirkungen auf die soziale Mobilität zwischen den Schulformen (Trautwein et al., 2008).

Der angeführte Wechsel zwischen den Schulformen und damit die Erhöhung der Durchlässigkeit im Sekundarbereich I wird als horizontale Durchlässigkeit bezeichnet. Empirische Daten zeigen, dass Schulformwechsel insbesondere in Form eines Wechsels in Richtung der anspruchsniedrigeren Schulformen stattfinden und darüber hinaus mit dem sozialen Status und einem möglichen Migrationshintergrund zusammenhängen (im Überblick Stubbe, 2009; s. a. Albrecht et al., 2018; Pfost, et al., 2018).

Die Möglichkeit, einen höherwertigen Schulabschluss auf einer Schulform zu erwerben, die herkömmlich nicht zu diesem Abschluss führt (z. B. Realschulabschluss an einer Hauptschule), wird vertikale Durchlässigkeit genannt. Die Optionen der vertikalen Durchlässigkeit wurden im deutschen Schulsystem in den vergangenen Jahren verstärkt genutzt, allerdings existieren Unterschiede hinsichtlich der Verwertbarkeit und Anerkennung dieser Abschlüsse auf dem Arbeitsmarkt. Für weibliche Hochschulabsolventinnen zeigt sich beispielsweise, dass bei einer direkt erworbenen Hochschulreife im Vergleich zur alternativ erworbenen Hochschulreife Berufe mit einem höheren durchschnittlichen Berufsprestige ausgeübt werden (Schuchart & Schimke, 2019).

Traditionell werden in der Bundesrepublik Deutschland in der Sekundarstufe I drei Schulformen unterschieden: Hauptschule, Realschule und Gymnasium. Seit den 1970er-Jahren wurden in einigen Ländern der Bundesrepublik Deutschland Gesamtschulen eingeführt, die diese drei Schulformen in einer Institution vereinen (Graalmann et al., 2023). In den vergangenen Jahren erfuhr das klassische dreigliedrige Schulsystem zahlreiche Reformen, wodurch bestehende Schulformen abgeschafft und neue Schulformen eingeführt wurden. Dabei hat das föderale System der Bundesrepublik zu einer sehr uneinheitlichen Struktur der Schullandschaft geführt, in der das Gymnasium mittlerweile die einzige Schulform auf Sekundarschulebene ist, die in allen Ländern existiert. Seit etwa 20 Jahren besuchen relativ konstant 35 Prozent der Schüler:innen der Sekundarstufe I ein Gymnasium (Rabenstein et al., 2020).

Tabelle 12.1: Übersicht über die Schulformen der Sekundarstufe I in den Ländern der Bundesrepublik Deutschland (Schuljahr 2022/23)

	Schulen mit einem Bildungsgang			Schulen mit mehreren Bildungsgängen		
				ohne Gymnasialbildung	mit Gymnasialbildung	
Länder	**Hauptschule**	**Realschule**	**Gymnasium**		*ohne gymn. Oberstufe*	*mit gymn. Oberstufe*
Baden-Württemberg	Hauptschule Werkrealschule[A]	Realschule	Gymnasium	–	–	Gemeinschaftsschule*
Bayern	Mittelschule[A]	Realschule	Gymnasium	–	(Gesamtschule)[B]	–
Berlin	–	–	Gymnasium	–	–	Integrierte Sekundarschule* Gemeinschaftsschule*
Brandenburg	–	–	Gymnasium	Oberschule[CD]	–	Gesamtschule
Bremen	–	–	Gymnasium	Werkschule[E]	–	Oberschule
Hamburg	–	–	Gymnasium	–	–	Stadtteilschule
Hessen	Hauptschule	Realschule	Gymnasium	Mittelstufenschule	–	Gesamtschule*
Meckl.-Vorpommern	–	–	Gymnasium	Regionale Schule[D]	–	Gesamtschule*
Niedersachsen	Hauptschule	Realschule	Gymnasium	Oberschule	Oberschule	Gesamtschule*
Nordrhein-Westfalen	Hauptschule	Realschule	Gymnasium	–	Sekundarschule	Gesamtschule*
Rheinland-Pfalz	–	–	Gymnasium	Realschule plus[CD]	–	Gesamtschule*
Saarland	–	–	Gymnasium	–	–	Gemeinschaftsschule*
Sachsen	–	–	Gymnasium	Oberschule[D]	–	Gemeinschaftsschule*
Sachsen-Anhalt	–	–	Gymnasium	Sekundarschule	–	Gesamtschule Gemeinschaftsschule*
Schleswig-Holstein	–	–	Gymnasium	–	–	Gemeinschaftsschule
Thüringen	–	–	Gymnasium	Regelschule[D]	–	Gesamtschule* Gemeinschaftsschule

* = Eine gymnasiale Oberstufe kann an der jeweiligen Schule oder als Kooperation mit anderen Schulen angeboten werden.
A = An dieser Schulform kann neben dem Hauptschulabschluss auch ein mittlerer Schulabschluss erworben werden, der jedoch nicht dem Realschulabschluss entspricht.
B = Insgesamt 5 Schulen, die als ‚Schulen besonderer Art' geführt werden.
C = An dieser Schulform ist der Erwerb der Fachoberschulreife möglich.
D = Wechsel an ein Gymnasium/Übertritt in die gymnasiale Oberstufe bei entsprechendem Notendurchschnitt oder dem Besuch spezieller Kurse (Sachsen) explizit möglich
E = Schulform für lernbenachteiligte Jugendliche (Jahrgangsstufe 9 bis 11)
Die Angaben basieren auf verfügbaren Veröffentlichungen der Kultusministerien der Länder der Bundesrepublik Deutschland (Stand: Schuljahr 2022/23).

 © TIMSS 2023

Tabelle 12.1 gibt einen Überblick über die allgemeinbildenden Schulformen der Sekundarstufe (ohne Förderschulen) in den einzelnen Ländern der Bundesrepublik Deutschland für das Schuljahr der Erhebung der *Trends in International Mathematics and Science Study* (TIMSS) 2023 (Schuljahr 2022/23). Im linken Teil der Tabelle finden sich die drei ‚klassischen' Schulformen. Auf den ersten Blick ist zu erkennen, dass in der Mehrzahl der Länder die Haupt- und Realschulen nicht mehr existieren und in allen Ländern – abgesehen von Bayern – Schulformen mit mehreren Bildungsgängen eingeführt wurden. Diese neuen Schulformen – sowie die Gesamtschulen – finden sich im rechten Teil der Tabelle. Unterschieden wird dabei zwischen Schulen ohne bzw. mit Gymnasialbildung und bei Letzteren zusätzlich zwischen Schulen mit eigener gymnasialer Oberstufe und ohne eigene gymnasiale Oberstufe.

In den Stadtstaaten Berlin und Hamburg, im Saarland sowie in Schleswig-Holstein wird eine Zweigliedrigkeit des Schulsystems umgesetzt, d. h. neben dem Gymnasium existiert nur eine weitere Schulform. Auch Bremen verfügt nach der Grundschule über ein zweigliedriges Schulsystem, das ab der neunten Jahrgangsstufe jedoch durch eine praxisnahe Schulform mit Berufsorientierung ergänzt wird. Die Benennung von Schulformen folgt dabei keiner bundesweit einheitlichen Regelung. Entsprechend finden sich für neue Schulformen ohne Gymnasialbildung in den neun Ländern, in denen eine derartige Schulform existiert, sieben unterschiedliche Bezeichnungen. Daneben unterscheiden sich auch die gesetzlichen Vorgaben zum Übergang von der Primar- in die Sekundarstufe zwischen den Ländern zum Teil erheblich. In einigen Ländern (z. B. Bayern) ist die Empfehlung der Grundschule für eine weiterführende Schulform verbindlich – auch wenn es Möglichkeiten wie Probeunterricht gibt, um dennoch eine höhere Schulform zu wählen. In anderen Ländern (z. B. Niedersachsen) liegt die Schulformwahl allein bei den Erziehungsberechtigten. In einer dritten Gruppe von Ländern (z. B. Hamburg) entscheiden zwar auch die Erziehungsberechtigten, diese Entscheidung ist dann aber ggf. mit einer Probezeit verbunden (Cremer & Tobisch, 2021; KMK, 2015).

Um diese länderspezifischen Unterschiede zu berücksichtigen, sollten Analysen zum Übergang in die Sekundarstufe im Idealfall auf Länderebene durchgeführt werden. Da dies mit den Daten aus TIMSS 2023 jedoch nicht möglich ist, werden in diesem Kapitel alle Befunde auf Bundesebene dargestellt, um zumindest ein übergreifendes Bild zeichnen zu können.

Die Analysen in diesem Kapitel schließen an die Berichtslegung im Rahmen der bisherigen Zyklen der TIMS-Studie 2011, 2015[1] und 2019 sowie der *Internationalen Grundschul-Lese-Untersuchung* (IGLU) 2001, 2006, 2011, 2016 und 2021 an (Arnold et al., 2007, 2010; Bos et al., 2004; Stubbe et al., 2012; Stubbe et al., 2017; Stubbe et al., 2020; Stubbe et al., 2016; Stubbe et al., 2023).

1 Die Werte, die in diesem Kapitel für TIMSS 2015 berichtet werden, weichen ggf. von den Ergebnissen aus dem entsprechenden Berichtsband (Stubbe et al., 2016) ab, da damals alle Analysen mit imputierten Daten durchgeführt worden sind (Wendt et al., 2016). Da dies für die übrigen TIMSS-Zyklen sowie für IGLU nicht gemacht wurde, werden aus Gründen der Vergleichbarkeit in diesem Kapitel auch die TIMSS-2015-Analysen mit nicht-imputierten Daten gerechnet.

Im Folgenden werden diese übergeordneten Forschungsfragen untersucht:

1. Wie verteilen sich die Schullaufbahnpräferenzen der Lehrkräfte und Erziehungsberechtigten auf die unterschiedlichen Schulformen?
2. Welche Zusammenhänge gibt es zwischen leistungsrelevanten Merkmalen der Schüler:innen (mathematische Kompetenz, naturwissenschaftliche Kompetenz, Anstrengungsbereitschaft, Leistungsangst, Schulnoten, Selbstkonzept) und Gymnasialpräferenzen von Lehrkräften?
3. Welche Zusammenhänge gibt es zwischen sozialen Merkmalen der Schüler:innen und Gymnasialpräferenzen von Lehrkräften und Erziehungsberechtigten?

12.3 Schullaufbahnpräferenzen von Lehrkräften und Erziehungsberechtigten: Verteilung und Ausmaß der Übereinstimmung

Lehrkräfte wurden bei TIMSS 2023 wie schon bei den beiden vorangegangenen Erhebungen und in Übereinstimmung mit dem IGLU-Vorgehen nach dem erwarteten Schulabschluss der einzelnen Viertklässler:innen gefragt. Dieses Vorgehen ist der Abfrage der ausgesprochenen Übergangsempfehlung vorzuziehen, da die Empfehlungspraktiken und deren Verbindlichkeit zwischen den Ländern der Bundesrepublik Deutschland erheblich variieren (siehe Abschnitt 12.2). Mit der Frage nach dem erwarteten Schulabschluss (im Folgenden als *Schullaufbahnpräferenzen der Lehrkräfte* bezeichnet) erhält man einen guten Hinweis auf die prognostizierte Entwicklung der einzelnen Schüler:innen aus Sicht ihrer Grundschullehrkräfte.

Wie in den vergangenen Erhebungen von IGLU und TIMSS wurden die *Schullaufbahnpräferenzen der Erziehungsberechtigten* über deren Angabe zu der Schulform erfasst, die ihre Kinder im kommenden Schuljahr voraussichtlich besuchen werden. Da die Erziehungsberechtigten in fast allen Ländern die Schulanmeldungen ihrer Kinder selbst vornehmen, kann die im Anschluss an die Grundschule besuchte Schule als Indikator für deren Präferenz interpretiert werden.

In Tabelle 12.2 finden sich die relativen Häufigkeiten der Schullaufbahnpräferenzen der Lehrkräfte und der Erziehungsberechtigten für die fünf Erhebungszeitpunkte von IGLU (2011 inkl. TIMSS) sowie für TIMSS 2015, 2019 und 2023. Hinsichtlich der Angaben der Lehrkräfte fällt auf, dass die Werte seit 2011 relativ konstant ausfallen. Für jeweils rund 40 Prozent der Viertklässler:innen erwarten die befragten Lehrkräfte einen Realschulabschluss bzw. die allgemeine Hochschulreife als Schulabschluss. Für etwa 20 Prozent wird ein Hauptschulabschluss (bzw. *erster allgemeinbildender Schulabschluss*; ESA) erwartet. Blickt man bis ins Jahr der ersten IGLU-Erhebung 2001 zurück, zeigt sich ein Rückgang der Hauptschulpräferenzen um 11.7 Prozentpunkte, während der Anteil der Gymnasialpräferenzen um 8.8 Prozentpunkte gestiegen ist.

Bei den Erziehungsberechtigten sind die Werte seit 2015 relativ konstant: Etwas weniger als die Hälfte präferiert das Gymnasium (bei IGLU 2021 ist der Anteil mit 53.0 % etwas höher), jeweils ca. 20 bis 25 Prozent eine Realschule bzw. eine Schule mit mehreren Bildungsgängen und nur knapp über 5 Prozent eine Hauptschule (bei IGLU 2021 ist der Anteil mit 4.7 % etwas geringer). Innerhalb der vergangenen 22 Jahre hat es bei diesen Präferenzen aber deutlichere Veränderungen gegeben als bei den Lehrkräften. Grund dafür ist vor allem die

Tabelle 12.2: Schullaufbahnpräferenzen von Lehrkräften und Erziehungsberechtigten von IGLU 2001 bis TIMSS 2023 (Angaben in Zeilenprozent)

Schullaufbahnpräferenz	Hauptschule	Realschule	Gymnasium	Schulen mit mehreren Bildungsgängen
Lehrkräfte				
2023 (TIMSS)	17.6	38.7	43.7	–
2021 (IGLU)	17.6	38.5	43.9	–
2019 (TIMSS)	19.1	40.1	40.8	–
2016 (IGLU)	17.4	39.6	43.0	–
2015 (TIMSS)	20.4	39.3	40.3	–
2011 (IGLU/TIMSS)	21.6	36.6	41.8	–
2006 (IGLU)	24.8	35.5	39.7	–
2001 (IGLU)	29.3	35.7	34.9	–
Erziehungsberechtigte				
2023 (TIMSS)	5.7	25.7	49.3	19.3
2021 (IGLU)	4.7	22.0	53.0	20.3
2019 (TIMSS)	6.8	25.2	46.9	21.0
2016 (IGLU)	6.1	21.3	47.6	25.0
2015 (TIMSS)	7.0	26.3	46.0	20.7
2011 (IGLU/TIMSS)	11.9	28.3	45.5	14.3
2006 (IGLU)	14.7	27.0	47.3	11.0
2001 (IGLU)	22.1	29.2	40.8	7.8

Differenzen zu 100 Prozent ergeben sich durch das Rundungsverfahren.
Abweichungen zu den im Rahmen von TIMSS 2015 berichteten Werten (Stubbe et al., 2016) ergeben sich daraus, dass dort für 2015 mit imputierten Daten gearbeitet wurde.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

zunehmende Verbreitung von Schulen mit mehreren Bildungsgängen, bei denen entsprechend ein Anstieg um rund 150 Prozent zu verzeichnen ist. Bei den Gymnasialpräferenzen hat es nach 2001 zwar zunächst einen Anstieg um 6.5 Prozentpunkte gegeben, seit 2006 gibt es bei diesem Wert (abgesehen von dem bereits erwähnten Ausreißer 2021) aber kaum Veränderungen. Einen Rückgang bei den Präferenzen hat es vor allem hinsichtlich der Hauptschulen gegeben (um 74.2 % seit 2001), aber auch die Realschulpräferenzen sind etwas rückläufig. Dass es nach wie vor relativ viele Übergänge auf Realschulen gibt, liegt darin begründet, dass es diese Schulform zwar nur noch in wenigen, dafür aber in den bevölkerungsreichen Ländern der Bundesrepublik Deutschland gibt.

Die Tabelle 12.3 zeigt für die drei Präferenzen der Lehrkräfte den jeweiligen Anteil der Schüler:innen nach Präferenz der Erziehungsberechtigten. Abweichend von der entsprechenden Darstellung in den vergangenen Studienzyklen wird hier – entsprechend dem Vorgehen bei IGLU 2021 – als weitere Ausprägung der Spaltenvariable das Fehlen der Angabe (*missing*) berücksichtigt. Der Grund für dieses Vorgehen ist der geringe Rücklauf der Elternfragebögen bei TIMSS 2023 (siehe Kapitel 2 in diesem Band). Deutlich wird in dieser Tabelle der systematische Ausfall bei den Angaben der Erziehungsberechtigten (siehe Kapitel 7 in diesem Band). Für knapp zwei Drittel der Schüler:innen, für die von Seiten der Grundschullehrkräfte ein Hauptschulabschluss erwartet wird, liegt keine Angabe zur Schullaufbahnpräferenz der Erziehungsberechtigten vor. Für den Realschulabschluss sind es knapp 50 Prozent und für das Abitur rund ein Drittel.

Darüber hinaus ergeben sich aus dieser Tabelle keine grundlegend anderen Befunde als in den vergangenen TIMSS-Zyklen. Viertklässler:innen mit einer Schullaufbahnpräferenz der Lehrkräfte für das Gymnasium wechseln in den meisten Fällen auf diese Schulform. Bei einer Realschulpräferenz von Seiten der

Tabelle 12.3: Übereinstimmung der Schullaufbahnpräferenzen von Lehrkräften und Erziehungsberechtigten (Angaben in Zeilenprozent)

		Schullaufbahnpräferenz der Erziehungsberechtigten					
		missing	Hauptschule	Realschule	Gymnasium	Schule mit mehreren Bildungsgängen	*n*
Schullaufbahn-präferenz der Lehrkräfte	Hauptschule	67.7	12.2	11.8	0.6	7.7	713
	Realschule	47.2	1.6	25.9	9.4	15.9	1567
	Gymnasium	32.5	0.3	5.3	55.1	6.9	1771
	n	1797	118	584	1126	425	4051

Differenzen zu 100 Prozent und zur Summe (*n*) ergeben sich durch das Rundungsverfahren der gewichteten Daten.

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Grundschule ist die Realschule die meistgewählte Schulform, gefolgt von Schulen mit mehreren Bildungsgängen und dem Gymnasium. Liegt von Seiten der Lehrkräfte eine Präferenz für die Hauptschule vor, entscheiden sich die Erziehungsberechtigten am häufigsten für eine Hauptschule, eine Realschule oder eine Schule mit mehreren Bildungsgängen.

12.4 Schullaufbahnpräferenzen von Lehrkräften: Leistungsrelevante Merkmale der Schüler:innen

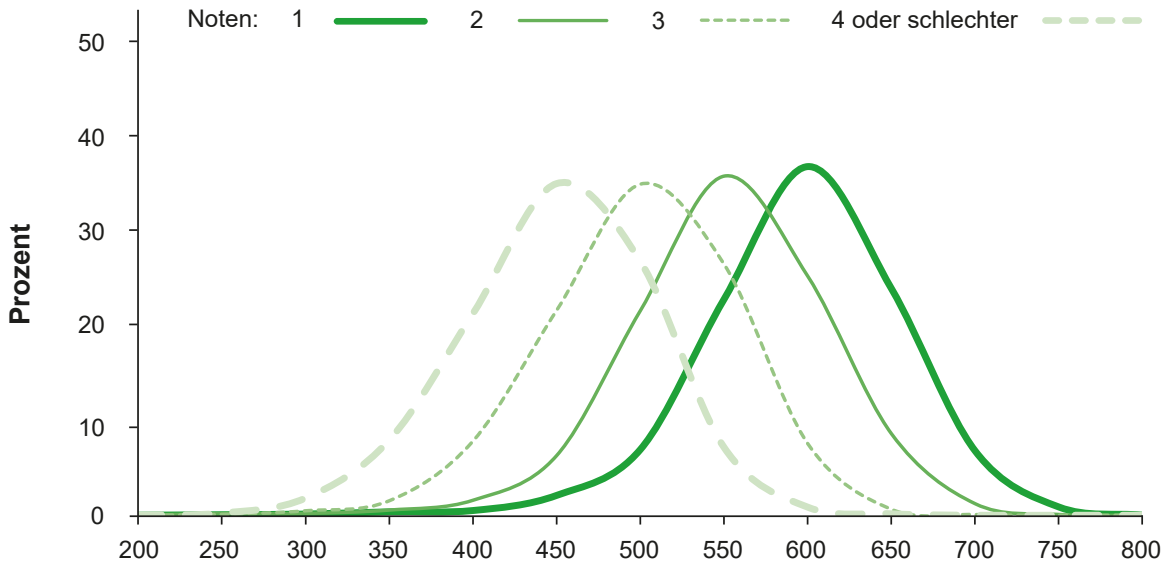
Wie bereits betont, unterscheiden sich die Regelungen zum Übergang von der Primar- in die Sekundarstufe zwischen den einzelnen Ländern der Bundesrepublik Deutschland. Gemein ist den jeweiligen Vorgaben aber, dass unterschiedliche leistungsrelevante Merkmale der Schüler:innen die Schullaufbahnempfehlung bestimmen sollen (KMK, 2015). Das wichtigste Kriterium sind dabei die Noten (insbesondere in den Hauptfächern), ergänzt durch weitere Aspekte wie Lernentwicklung sowie Arbeits- und Sozialverhalten.

Da aus unterschiedlichen Studien bekannt ist, dass die Korrelation zwischen Schulnoten und standardisierten Leistungstests relativ gering ausfällt (im Überblick Lintorf, 2012), zeigt Abbildung 12.1 zunächst die Mathematikkompetenz[2] der Viertklässler:innen differenziert nach den Mathematiknoten, bevor der Zusammenhang der Schullaufbahnpräferenz der Lehrkräfte mit der Mathematikkompetenz, den Mathematiknoten und weiteren ausgewählten leistungsrelevanten Merkmalen betrachtet wird.

Zunächst wird erwartungskonform deutlich, dass sich die mittleren Mathematikkompetenzen je nach erhaltener Mathematiknote substanziell unterscheiden. Zugleich zeigen die Überlappungen der vier Kurven aber auch, dass Viertklässler:innen mit mittleren Testleistungen (ca. 500 – 550 Punkte) je nach Lehrkraft alle Noten von ausreichend (oder schlechter) bis sehr gut erhalten können. Noten hängen somit nicht nur von den Leistungen der Schüler:innen ab, sondern auch von den Bewertungsmaßstäben der jeweiligen Lehrkräfte. In den folgenden Analysen muss also immer bedacht werden, dass sich Noten im Gegensatz

2 Da die Befunde für die Abbildungen 12.1, 12.2 sowie 12.3 für Mathematik und Naturwissenschaften ähnlich ausfallen, werden die Ergebnisse lediglich exemplarisch für Mathematik dargestellt.

Abbildung 12.1: Testleistungen der Schüler:innen differenziert nach Mathematiknoten – Gesamtskala Mathematik

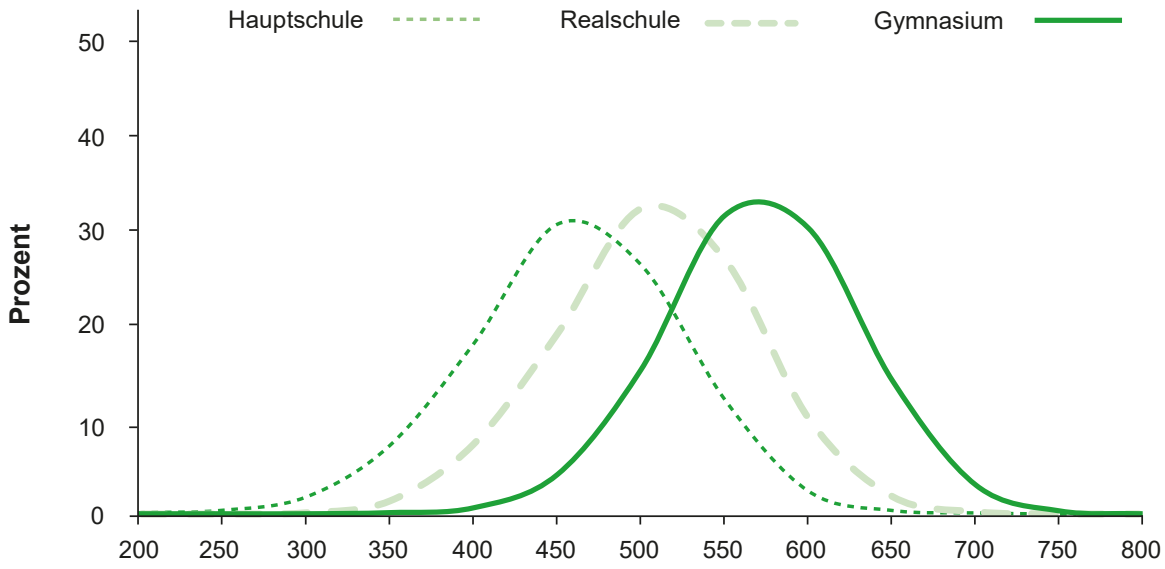


IEA: Trends in International Mathematics and Science Study © TIMSS 2023

zu standardisierten Tests nur bedingt für den Leistungsvergleich zwischen unterschiedlichen Klassen eignen.

In Abbildung 12.2 wird die Mathematikkompetenz der Viertklässler:innen differenziert nach der Schullaufbahnpräferenz der Lehrkräfte dargestellt. Auch in dieser Abbildung lassen sich neben den offensichtlichen Mittelwertsdifferenzen die deutlichen Überlappungen der drei Kurven erkennen, was darauf hinweist, dass die Lehrkräfte von Schüler:innen mit mittleren Testleistungen – insbesondere von denjenigen mit Kompetenzstufe III (475 – 550 Punkte) – Präferenzen für alle drei Schulformen aufweisen.

Abbildung 12.2: Testleistungen der Schüler:innen differenziert nach Schullaufbahnpräferenz der Lehrkräfte – Gesamtskala Mathematik



IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Tabelle 12.4 zeigt den Zusammenhang der Schullaufbahnpräferenz der Lehrkräfte mit den leistungsrelevanten Merkmalen der Viertklässler:innen, die im Rahmen von TIMSS 2023 erhoben wurden: die mathematische und die naturwissenschaftliche Kompetenz (erhoben über den TIMSS-Test), die Anstrengungsbereitschaft und die Leistungsangst (erhoben über den Elternfragebogen), die Noten in den Fächern Deutsch, Mathematik und Sachunterricht (erhoben über die Schülerteilnahmeliste) sowie das Selbstkonzept in Mathematik und im Sachunterricht (erhoben über den Schülerfragebogen). Berichtet werden standardisierte Regressionskoeffizienten (β) mit der dichotomen Variable *Gymnasialpräferenz der Lehrkräfte* als abhängige Variable. Bei den Modellen in der ersten Spalte wird jeweils nur eine unabhängige Variable berücksichtigt. Bei den Modellen in der zweiten Spalte werden zusätzlich zu der jeweiligen Zeilenvariable die Kompetenzen in den Domänen Mathematik und Naturwissenschaften kontrolliert. Bei den Modellen in der dritten Spalte werden zusätzlich zu der jeweiligen Zeilenvariable wiederum die Noten in den Fächern Deutsch, Mathematik und Sachunterricht kontrolliert.

Tabelle 12.4: Bivariate Zusammenhänge zwischen der Schullaufbahnpräferenz der Lehrkräfte und leistungsrelevanten Merkmalen der Schüler:innen (ohne und mit Berücksichtigung der Kompetenzen und der Schulnoten)

	Zusammenhang mit der Gymnasialpräferenz der Lehrkräfte		Zusammenhang mit der Gymnasialpräferenz der Lehrkräfte unter Kontrolle der Mathematik- und Naturwissenschaftskompetenz		Zusammenhang mit der Gymnasialpräferenz der Lehrkräfte unter Kontrolle der Deutsch-, Mathematik- und Sachunterrichtsnoten	
	b_{std}	(*SE*)	b_{std}	(*SE*)	b_{std}	(*SE*)
Mathematische Kompetenz	0.54**	(0.01)			0.16**	(0.02)
Naturwissenschaftliche Kompetenz	0.50**	(0.01)			0.11**	(0.02)
Anstrengungsbereitschaft	0.43**	(0.01)	0.30**	(0.02)	0.14**	(0.01)
Leistungsangst	-0.29**	(0.02)	-0.13**	(0.02)	-0.07**	(0.01)
Note in Deutsch	-0.66**	(0.01)	-0.52**	(0.01)		
Note in Mathematik	-0.63**	(0.01)	-0.47**	(0.02)		
Note im Sachunterricht	-0.59**	(0.01)	-0.43**	(0.01)		
Selbstkonzept Mathematik	0.39**	(0.01)	0.18**	(0.01)	0.08**	(0.01)
Selbstkonzept Sachunterricht	0.30**	(0.01)	0.14**	(0.02)	0.06**	(0.01)

b_{std} = Regressionsgewicht (standardisiert)
Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p \leq .05$); ** = signifikant ($p \leq .01$)

IEA: Trends in International Mathematics and Science Study

© TIMSS 2023

Es zeigt sich, dass alle Effekte signifikant ausfallen. Als besonders eng erweist sich (ohne Kontrolle weiterer Variablen) erwartungskonform der Zusammenhang mit den Schulnoten, aber auch mit den standardisierten Leistungstests, gefolgt von Anstrengungsbereitschaft, den beiden Selbstkonzepten und der Leistungsangst.

Unter Kontrolle der mathematischen und der naturwissenschaftlichen Kompetenz lassen sich für die übrigen Merkmale weiterhin relativ große Effekte feststellen, wobei die Schulnoten wiederum den engsten Zusammenhang mit der Schullaufbahnpräferenz aufweisen. Deutlich kleinere Regressionskoeffizienten ergeben sich für die Modelle, in denen die Noten kontrolliert werden.

Im Vergleich mit TIMSS 2019 zeigen sich keine relevanten Veränderungen (Stubbe et al., 2020).

12.5 Schullaufbahnpräferenzen von Lehrkräften und Erziehungsberechtigten: Soziale Merkmale der Schüler:innen

Während die Berücksichtigung leistungsrelevanter Merkmale der Schüler:innen beim Übergang in die Sekundarstufe bildungspolitisch vorgegeben ist, dürften Hintergrundmerkmale der Kinder sowie ihrer Familien, die nicht unmittelbar mit Leistung in Verbindung stehen, keinen Zusammenhang mit den Präferenzen der Grundschullehrkräfte sowie der eigenen Erziehungsberechtigten aufweisen. Dies widerspräche dem gesellschaftlichen Anspruch auf gleiche Chancen für höhere Bildungsabschlüsse unabhängig vom sozialen Status des Elternhauses. Zahlreiche Studien weisen diese Zusammenhänge jedoch regelmäßig nach (siehe Kapitel 7 in diesem Band), sodass im Folgenden Merkmale des sozialen Status analysiert werden. Die soziale Lage von Familien wird dabei mithilfe der *Erikson-Goldthorpe-Portocarero-Klasseneinteilung* (EGP-Klassen; Erikson et al., 1979) beschrieben, die Personen ausgehend von Angaben zu ihrer beruflichen Stellung distinkten Gruppen zuordnet. Bei der Interpretation der Befunde muss beachtet werden, dass durch ein verändertes Erhebungsverfahren bei TIMSS 2023 die Anteile in den EGP-Klassen II und III in diesem Zyklus überschätzt werden. In den vorangegangenen Zyklen wurden die Anteile in den EGP-Klassen I, IV, V, VI und VII überschätzt. Daher sind die Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar. Für nähere Informationen dazu siehe Kapitel 7 in diesem Band.

In Tabelle 12.5 sind die relativen Chancen (*odds ratios*) für eine Gymnasialpräferenz der Lehrkräfte und der Erziehungsberechtigten in Abhängigkeit von der EGP-Klasse der Familie dargestellt. Als Referenzgruppe werden jeweils Kinder, deren Lehrkräfte bzw. Erziehungsberechtigte eine Präferenz für eine andere Schulform als das Gymnasium haben und deren Erziehungsberechtigte der Gruppe der (Fach-)Arbeiter:innen angehören, herangezogen.

Tabelle 12.5: Relative Chancen (*odds ratios*) für eine Gymnasialpräferenz der Lehrkräfte bzw. der Erziehungsberechtigten nach sozialer Lage (EGP) der Schüler:innen

Sozioökonomische Stellung der Familie (EGP-Klasse)	Gymnasialpräferenz der Lehrkräfte			Gymnasialpräferenz der Erziehungsberechtigten		
	Modell I	Modell II	Modell III	Modell I	Modell II	Modell III
Obere Dienstklasse (I)	3.65**	2.79**	2.17**	4.47**	3.61**	2.98**
Untere Dienstklasse (II)	2.13**	1.64**	–ns	2.56**	2.05**	1.76*
Routinedienstleistungen (III)	–ns	–ns	–ns	1.56*	–ns	–ns
Selbstständige (IV)	–ns	–ns	–ns	1.72*	–ns	1.92*
(Fach-)Arbeiter:innen (V, VI)	Referenzgruppe (*odds ratio* = 1)					
Un- und angelernte Arbeiter:innen (VII)	0.57*	0.52**	–ns	0.59*	0.57*	–ns
Nagelkerkes R^2	0.08	0.23	0.41	0.09	0.21	0.33

Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (für nähere Informationen siehe Kapitel 7 in diesem Band).
Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p \leq .05$); ** = signifikant ($p \leq .01$)
Modell I: ohne Kontrolle von Kovariaten
Modell II: Kontrolle von kognitiven Fähigkeiten
Modell III: Kontrolle von kognitiven Fähigkeiten und der Mathematik- und Naturwissenschaftskompetenzen

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Ohne Berücksichtigung weiterer Einflussfaktoren (Modell I) ist die Chance von Schüler:innen aus der oberen Dienstklasse, eine Gymnasialpräferenz der Lehrkraft zu erhalten, mehr als dreieinhalbmal so hoch wie die eines (Fach-)Arbeiterkindes. Immerhin mehr als doppelt so groß ist diese Chance für Schüler:innen aus der unteren Dienstklasse. Viertklässler:innen, deren Erziehungsberechtigte der Gruppe der un- und angelernten Arbeiter:innen angehören, haben verglichen mit (Fach-)Arbeiterkindern hingegen nur eine halb so große Chance auf eine Gymnasialpräferenz ihrer Grundschullehrkräfte. In den EGP-Klassen III und IV gibt es keine signfikanten Unterschiede zur Referenzgruppe. Die Modellgüte ist allerdings gering (Nagelkerkes R^2: 0.08). Da ein Zusammenhang zwischen dem sozialen Status von Familien und den schulischen Leistungen der Kinder besteht (siehe Kapitel 7 in diesem Band), werden in Modell II zunächst die kognitiven Fähigkeiten[3] und in Modell III zudem die Mathematik- und Naturwissenschaftskompetenzen der Viertklässler:innen kontrolliert, um zu prüfen, ob ungleiche Chancen in ungleichen Leistungen begründet liegen. In Modell II weisen dieselben drei EGP-Klassen signifikante Unterschiede auf wie in Modell I, wobei die Faktoren in den beiden Dienstklassen etwas kleiner werden. Der Unterschied zwischen der EGP-Klasse VII und der Referenzgruppe wird hingegen nominell sogar etwas größer (für die Interpretation von *odds ratios* siehe Kuckartz et al., 2013). Die Erklärungskraft von Modell II ist akzeptabel (Nagelkerkes R^2: 0.23). Werden zusätzlich die Mathematik- und Naturwissenschaftskompetenzen berücksichtigt (Modell III), hat das Pseudo-R^2 nach Nagelkerke den Wert 0.41 und die Modellgüte ist damit gut (Kuckartz et al., 2013). In diesem Modell weist nur die obere Dienstklasse einen signifikanten Unterschied zur Referenzgruppe auf. Mehr als doppelt so groß ist die Chance auf eine Gymnasialpräferenz für Schüler:innen aus dieser Gruppe im Vergleich zu (Fach-)Arbeiterkindern.

Bei den Gymnasialpräferenzen der Erziehungsberechtigten weisen in Modell I alle Ausprägungen der unabhängigen Variable signifkante Effekte auf (Nagelkerkes R^2: 0.09; akzeptabel). In Modell II (Nagelkerkes R^2: 0.21; akzeptabel) sind es die beiden Dienstklassen sowie die Gruppe der un- und angelernten Arbeiter:innen und in Modell III (Nagelkerkes R^2: 0.33; akzeptabel) sind es die beiden Dienstklassen sowie die Gruppe der Selbstständigen. Sofern die Werte signifikant ausfallen, ist die Chance auf eine Gymnasialpräferenz (jeweils verglichen mit der Referenzgruppe) in der EGP-Klasse VII rund halb so groß und in den EGP-Klassen II, III und IV etwa anderthalb- bis zweieinhalbmal so groß. Am deutlichsten sind die Effekte in der oberen Dienstklasse. Ohne die Berücksichtigung weiterer Variablen haben Kinder aus dieser Gruppe eine viereinhalbmal so große Chance, auf ein Gymnasium zu wechseln, wie Kinder von (Fach-)Arbeiter:innen. Werden sowohl die kognitiven Fähigkeiten als auch die Mathematik- und Naturwissenschaftskompetenz kontrolliert, ist diese Chance immer noch rund dreimal so groß.

Für den Trend seit 2011 werden in Tabelle 12.6 die relativen Chancen (*odds ratios*) für die *service class* (EGP I und II) im Vergleich zur *working class* (EGP V, VI, VII) für eine Gymnasialpräferenz der Lehrkräfte und der Erziehungsberechtigten berichtet. Beispielsweise bedeutet der Wert 3.45 (2023, Lehrkräfte, Modell I) in dieser Tabelle, dass Kinder, deren Erziehungsberechtigte einer der Dienstklassen angehören, ohne Berücksichtigung weiterer Merkmale eine 3.45-mal so hohe Chance auf eine Gymnasialpräferenz ihrer Lehrkräfte haben wie Kinder, deren Erziehungsberechtigte zu den Arbeiter:innen gehören. Ohne

3 Die erhobenen Subskalen des *kognitiven Fähigkeitstests* (KFT) wurden seit TIMSS 2007 unter der Annahme der Messinvarianz skaliert.

Berücksichtigung weiterer Einflussfaktoren (Modell I) zeigen sich sowohl bei den Lehrkräften als auch bei den Erziehungsberechtigten 2011 die größten sozialen Disparitäten. Am geringsten ist dieser Effekt bei den Erziehungsberechtigten 2015 und bei den Lehrkräften 2023. Unter Kontrolle der kognitiven Fähigkeiten (Modell II) fallen alle Effekte geringer aus, sind aber weiterhin signifikant. Werden zusätzlich die mathematischen und naturwissenschaftlichen Kompetenzen kontrolliert (Modell III), lassen sich für 2011 weder für die Präferenzen der Lehrkräfte noch für die Präferenzen der Erziehungsberechtigten signifikante Effekte des sozialen Status feststellen. Ebenso verhielt es sich 2015 im Modell der Erziehungsberechtigten. Die übrigen fünf Koeffizienten sind signifikant, aber kleiner als die entsprechenden Werte aus Modell II.

In Tabelle 12.7 sind die relativen Chancen (*odds ratios*) für eine Gymnasialpräferenz der Lehrkräfte und der Erziehungsberechtigten in Abhängigkeit vom Migrationshintergrund der Familie dargestellt. Als Referenzgruppe dienen jeweils Kinder, deren Lehrkräfte bzw. Erziehungsberechtigte eine Präferenz für eine andere Schulform als das Gymnasium haben und deren Eltern beide im Ausland geboren wurden.

Tabelle 12.6: Relative Chancen (*odds ratios*) für eine Gymnasialpräferenz der Lehrkräfte bzw. der Erziehungsberechtigten für Kinder aus Familien der *service class* (EGP I und II) im Vergleich mit Kindern aus Familien der *working class* (EGP V, VI und VII) bei TIMSS 2011, 2015, 2019 und 2023

	Gymnasialpräferenz der Lehrkräfte			Gymnasialpräferenz der Erziehungsberechtigten		
	Modell I	Modell II	Modell III	Modell I	Modell II	Modell III
2023	3.45**	2.78**	2.01**	4.10**	3.37**	2.73**
2019	3.91**	3.37**	2.41*	3.86**	3.38**	2.53*
2015	3.67**	3.36**	2.65**	3.04**	2.65**	–ns
2011	4.48**	3.79*	–ns	4.61**	3.86*	–ns

Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (für nähere Informationen siehe Kapitel 7 in diesem Band).
Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p \leq .05$); ** = signifikant ($p \leq .01$)
Modell I: ohne Kontrolle von Kovariaten
Modell II: Kontrolle der kognitiven Fähigkeiten
Modell III: Kontrolle der kognitiven Fähigkeiten und der Mathematik- und Naturwissenschaftskompetenzen

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Tabelle 12.7: Relative Chancen (*odds ratios*) für eine Gymnasialpräferenz der Lehrkräfte bzw. der Erziehungsberechtigten nach Migrationshintergrund der Familie

im Ausland geboren	Gymnasialpräferenz der Lehrkräfte			Gymnasialpräferenz der Erziehungsberechtigten		
	Modell I	Modell II	Modell III	Modell I	Modell II	Modell III
kein Elternteil	2.52**	1.84**	–ns	1.29*	–ns	0.65**
ein Elternteil	2.77**	2.25**	1.87**	1.64**	–ns	–ns
beide Elternteile	Referenzgruppe (*odds ratio* = 1)					
Nagelkerkes R^2	0.05	0.23	0.43	0.00	0.14	0.31

Signifikanzniveau: ns = nicht signifikant; * = signifikant ($p \leq .05$); ** = signifikant ($p \leq .01$)
Modell I: ohne Kontrolle von Kovariaten
Modell II: Kontrolle von kognitiven Fähigkeiten
Modell III: Kontrolle von kognitiven Fähigkeiten und der Mathematik- und Naturwissenschaftskompetenzen

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Hinsichtlich der Präferenzen der Lehrkräfte zeigt sich sowohl in Modell I (ohne Kontrolle von Kovariaten) als auch in Modell II (Kontrolle der kognitiven Fähigkeiten), dass Kinder, deren Eltern beide in Deutschland geboren wurden, eine signifikant höhere Chance auf eine Gymnasialpräferenz haben, als Kinder, deren Eltern beide im Ausland geboren wurden. Werden zusätzlich die Mathematik- und Naturwissenschaftskompetenzen der Viertklässler:innen kontrolliert, gibt es keinen signifikanten Unterschied zwischen den beiden Gruppen. Werden Familien, in denen ein Elternteil im Ausland geboren wurde, mit Familien, in denen beide Elternteile im Ausland geboren wurden, verglichen, finden sich in allen drei Modellen signifikante Effekte zugunsten der erstgenannten Gruppe. Das Pseudo-R^2 nach Nagelkerke ist im ersten Modell sehr gering (0.05), im zweiten Modell akzeptabel (0.23) und im dritten Modell gut (0.43).

Im ersten Modell (ohne Kontrolle von Kovariaten) für die Erziehungsberechtigten ergeben sich geringe, aber signifikante Effekte zuungunsten der Referenzgruppe. Nagelkerkes R^2 hat allerdings den Wert 0.00, sodass das Modell nicht weiter interpretiert werden sollte. Im zweiten Modell (Kontrolle der kognitiven Fähigkeiten) ist die Erklärungskraft mit einem Pseudo-R^2 von 0.14 gering und es gibt keine signifkanten Unterschiede zwischen den Gruppen. Das dritte Modell (Kontrolle der kognitiven Fähigkeiten und der Mathematik- und Naturwissenschaftskompetenz) hat eine akzeptable Erklärungskraft (Nagelkerkes R^2: 0.31). Zwischen Familien, in denen ein Elternteil bzw. beide Elternteile im Ausland geboren wurden, gibt es keine signifikanten Unterschiede hinsichtlich der Gymnasialpräferenz der Erziehungsberechtigten. Kinder aus Familien, in denen beide Elternteile in Deutschland geboren wurden, haben hingegen nur eine rund halb so große Chance auf ein Gymnasium zu wechseln wie Kinder, deren Eltern beide im Ausland geboren wurden. Diese vergleichsweise hohen Bildungsaspirationen von Familien mit Migrationshintergrund sind aus unterschiedlichen Studien bekannt (Becker & Gresch, 2016) und wurden auch in vorangegangen Erhebungszyklen von TIMSS und IGLU festgestellt (Stubbe et al., 2020; Stubbe et al., 2023).

Eine alternative und anschauliche Möglichkeit, den Zusammenhang zwischen Hintergrundmerkmalen und Schullaufbahnpräferenzen darzustellen, ist die Berechnung der Kompetenzwerte, die notwendig sind, damit Viertklässler:innen mit einer Wahrscheinlichkeit von mindestens 50 Prozent eine Präferenz für das Gymnasium erhalten. Diese Kompetenzwerte wurden in den IGLU- und TIMSS-Zyklen bis einschließlich 2016 als ‚kritische Werte' bzw. ‚gruppenspezifische Standards' bezeichnet. Seit TIMSS 2019 wird ein geringfügig abweichendes Berechnungsverfahren gewählt und, um den Unterschied deutlich zu machen, entsprechend von Schwellenwerten gesprochen.

Tabelle 12.8 stellt diese Schwellenwerte getrennt für Lehrkräfte und Erziehungsberechtigte jeweils für Mathematik und Naturwissenschaften dar. Als Referenz werden zunächst die Schwellenwerte für die Gesamtpopulation berichtet. Dabei zeigt sich, dass die Viertklässler:innen für eine wahrscheinliche Gymnasialpräferenz ihrer Lehrkräfte in Mathematik eine Kompetenz benötigen, die mit 544 Punkten 20 Punkte über dem nationalen Mittelwert liegt, und in Naturwissenschaften eine Kompetenz, die mit 540 Punkten 25 Punkte über dem nationalen Mittelwert liegt. Die Schwellenwerte der Erziehungsberechtigten liegen in beiden Kompetenzen nur 1 Punkt über denen der Lehrkräfte. Es kann also wie in den vorangegangenen Zyklen nicht grundsätzlich festgestellt werden, dass die Erziehungsberechtigten geringere (oder höhere) Leistungen für eine Gymnasialpräferenz erwarten als Lehrkräfte.

Tabelle 12.8: Schwellenwerte der Mathematik- und Naturwissenschaftskompetenz für eine fünfzigprozentige Wahrscheinlichkeit für eine Gymnasialpräferenz der Lehrkräfte bzw. der Erziehungsberechtigten

	Schwellenwert für eine Gymnasialpräferenz der Lehrkräfte		Schwellenwert für eine Gymnasialpräferenz der Erziehungsberechtigten	
	Mathematik	**Naturwissenschaften**	**Mathematik**	**Naturwissenschaften**
Gesamt	544	540	545	541
Sozioökonomische Stellung der Familie (EGP-Klasse)				
Obere Dienstklasse (I)	512	491	508	494
Untere Dienstklasse (II)	537	530	534	528
Routinedienstleistungen (III)	554	552	540	535
Selbstständige (IV)	534	532	532	530
(Fach-)Arbeiter:innen (V, VI)	573	575	549	546
Un- und angelernte Arbeiter:innen (VII)	646	769	575	580
im Ausland geboren				
kein Elternteil	548	552	540	543
ein Elternteil	526	511	527	516
beide Elternteile	538	504	565	538
Geschlecht				
Mädchen	533	535	534	534
Jungen	556	546	548	548

Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (für nähere Informationen siehe Kapitel 7 in diesem Band).

IEA: Trends in International Mathematics and Science Study © TIMSS 2023

Werden die Werte getrennt nach EGP-Klasse betrachtet, zeigt sich, dass Kinder aus der oberen Dienstklasse eine Mathematikkompetenz von nur 512 Punkten (12 Punkte unter dem nationalen Mittelwert) benötigen, damit eine fünfzigprozentige Wahrscheinlichkeit für eine Gymnasialpräferenz der Lehrkräfte besteht. Im Bereich Naturwissenschaften liegt dieser Wert sogar bei nur 491 Punkten (24 Punkte unter dem nationalen Mittelwert). Kinder von (Fach-)Arbeiter:innen benötigen für eine entsprechende Präferenz in Mathematik eine um 61 Punkte höhere Kompetenz und in Naturwissenschaften sogar eine 84 Punkte höhere Kompetenz als Kinder aus der oberen Dienstklasse. Diese Differenzen liegen insbesondere im Bereich Naturwissenschaften sehr deutlich über dem, was Grundschulkinder durchschnittlich in der vierten Klasse dazulernen (ca. 38 bis 55 Punkte in Mathematik und 25 bis 30 Punkte in Naturwissenschaften pro Jahr; Wendt et al., 2017). Deutlich geringer fallen diese Differenzen aus, wenn die Schwellenwerte der Erziehungsberechtigten betrachtet werden: 41 Punkte im Bereich Mathematik und 52 Punkte in Naturwissenschaften.

Insgesamt lässt sich feststellen, dass Lehrkräfte und Erziehungsberechtigte sehr ähnliche Kompetenzen für eine Gymnasialpräferenz erwarten. Auch in den EGP-Klassen I bis IV gibt es nur sehr geringe Unterschiede zwischen den beiden Gruppen. Von Kindern aus den EGP-Klassen V bis VII erwarten Grundschullehrkräfte hingegen deutlich höhere Kompetenzen, als es die Erziehungsberechtigten tun. Dies führt dazu, dass die sozialen Disparitäten in den Präferenzen der Lehrkräfte wie beschrieben größer sind als in den Präferenzen der Erziehungs-

berechtigten. Allerdings ist anzumerken, dass die Schwellenwerte für die EGP-Klassen V bis VII nur eingeschränkt interpretiert werden sollten. Die Lehrkräfte scheinen hinsichtlich dieser Gruppen insgesamt starke Vorbehalte gegen eine Gymnasiallaufbahn zu haben.

Es bleibt somit abzuwarten, ob TIMSS 2023 in diesem Punkt ein Ausreißer ist und die Befunde in zukünftigen Zyklen wieder etwas geringere soziale Disparitäten in den Präferenzen der Lehrkräfte (im Vergleich zu den Präferenzen der Erziehungsberechtigten) ergeben, so wie es in vorangegangenen TIMSS- und IGLU-Zyklen der Fall war (Stubbe et al., 2020 für TIMSS 2019 und Stubbe et al., 2023 für IGLU 2021).

Auch hinsichtlich des Migrationshintergrundes zeigen sich im Rahmen von TIMSS 2023 einige Resultate bei den Schwellenwerten, die von den Befunden aus vorherigen Zyklen abweichen. Wird zur Interpretation ergänzend Tabelle 12.7 hinzugezogen, zeigt sich, dass es bei den Erziehungsberechtigten offensichtlich eine unterschiedliche Bewertung von Mathematik- bzw. Naturwissenschaftskompetenz gibt, die so in vorherigen Zyklen nicht erkennbar ist. Auch für die Lehrkräfte relativieren sich die im Folgenden dargestellen Ergebnisse, wenn Tabelle 12.7 ergänzend berücksichtigt wird.

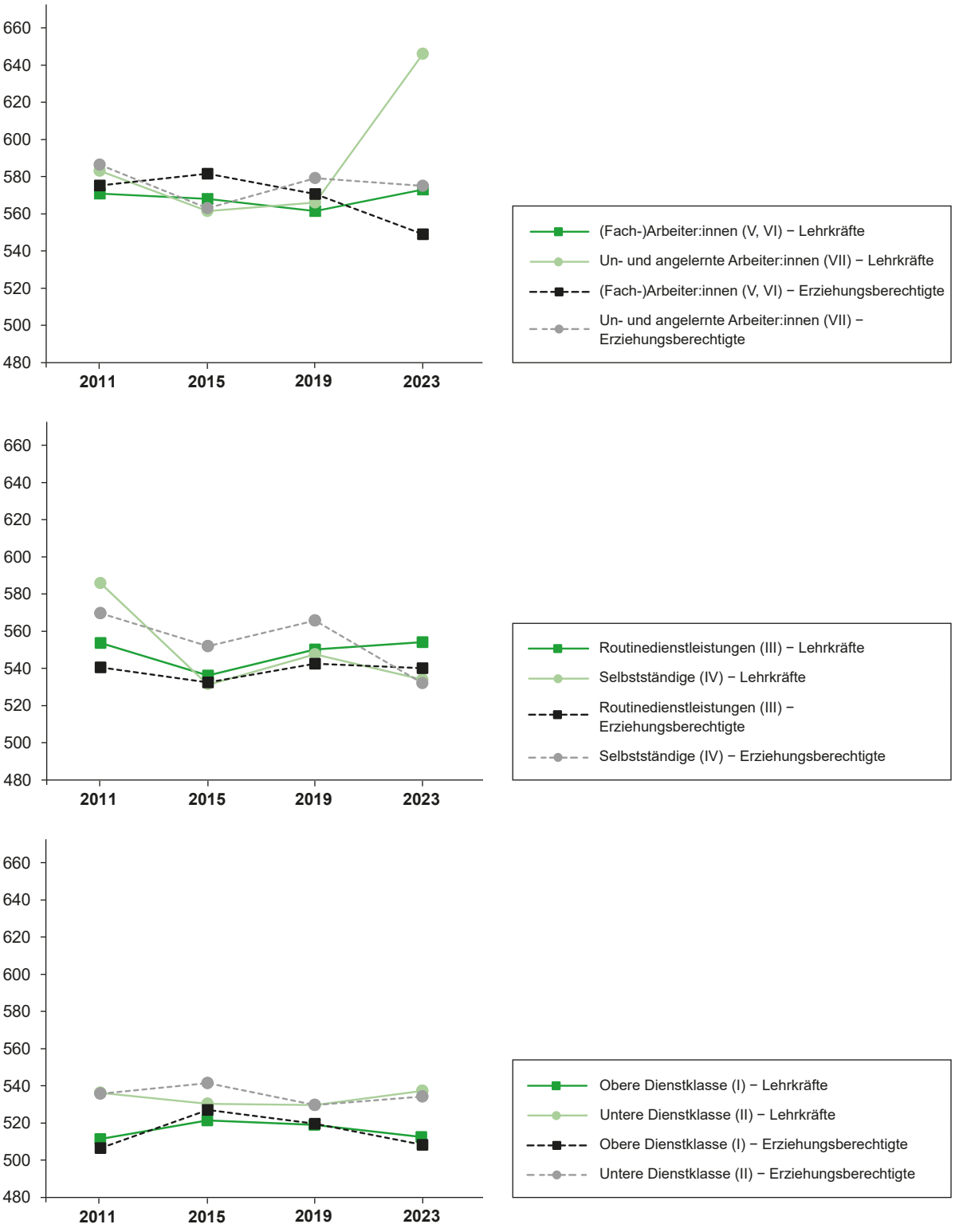
Bei TIMSS 2023 kann für den Bereich Mathematik festgestellt werden, dass sowohl Lehrkräfte als auch Erziehungsberechtigte den kleinsten Schwellenwert aufweisen, wenn ein Elternteil im Ausland geboren wurde. Werden die Gruppen *kein Elternteil im Ausland geboren* und *beide Elternteile im Ausland geboren* verglichen, zeigt sich sowohl bei den Lehrkräften als auch bei den Erziehunsberechtigten der gegenteilige Effekt im Vergleich zu vorherigen Zyklen. Lehrkräfte erwarten von Viertklässler:innen mit Migrationshintergrund eine 10 Punkte geringere Kompetenz für eine Gymnasialpräferenz als von Viertklässler:innen ohne Migrationshintergrund. Erziehungsberechtigte mit Migrationshintergrund erwarten von ihren Kindern eine 25 Punkte höhere Kompetenz für eine Gymasialpräferenz als Erziehungsberechtigte ohne Migrationshintergrund.

Im Bereich Naturwissenschaften waren die Befunde schon in vorherigen Zyklen weniger eindeutig. Im aktuellen TIMSS-Zyklus zeigt sich bei den Lehrkräften, dass der Schwellenwert für Kinder, deren Eltern beide in Deutschland geboren wurden, am höchsten ist. Wurde ein Elternteil im Ausland geboren, fällt der Wert 41 Punkte niedriger aus (TIMSS 2019: 15 Punkte). Wurden beide Elternteil im Ausland geboren, sind es 48 Punkte (TIMSS 2019: 11 Punkte). Bei den Erziehungsberechtigten gibt es zwischen den Gruppen *kein Elternteil im Ausland geboren* und *beide Elternteile im Ausland geboren* nur eine Differenz von 5 Punkten (TIMSS 2019: 51 Punkte), wobei die Anforderungen in Familien mit Migrationshintergrund etwas geringer sind. Der niedrigste Schwellenwert zeigt sich in der Gruppe *ein Elternteil im Ausland geboren*.

Hinsichtlich der Geschlechterunterschiede gibt es in der aktuellen Studie keine substanziellen Veränderungen zu den vorangegangenen Zyklen. Mädchen benötigen sowohl in Mathematik als auch in Naturwissenschaften eine geringere Kompetenz als Jungen, um eine Gymnasialpräferenz zu erhalten. Diese Differenzen liegen zwischen 11 Punkten (Lehrkräfte/Naturwissenschaften) und 23 Punkten (Lehrkräfte/Mathematik).

Für den Kompetenzbereich Mathematik sind in Abbildung 12.3 die Schwellenwerte der Lehrkräfte und der Erziehungsberechtigten für TIMSS 2011, 2015, 2019 und 2023 getrennt nach EGP-Klasse dargestellt. Für die EGP-Klassen V bis VII lassen sich hier sehr deutlich die beschriebenen Veränderungen der sozialen Disparitäten hinsichtlich der Präferenzen von Lehrkräften und Erziehungsberech-

Abbildung 12.3: Schwellenwerte der Mathematikkompetenz für eine fünfzigprozentige Wahrscheinlichkeit für eine Gymnasialpräferenz der Lehrkräfte bzw. der Erziehungsberechtigten bei TIMSS 2011, 2015, 2019 und 2023



Durch ein verändertes Erhebungsverfahren in TIMSS 2023 sind Analysen mit den EGP-Klassen nur eingeschränkt mit den vorangegangenen Zyklen vergleichbar (für nähere Informationen siehe Kapitel 7 in diesem Band).
Abweichungen zu den im Rahmen von IGLU/TIMSS 2011 und TIMSS 2015 berichteten Werten (Stubbe et al., 2012; Stubbe et al., 2016) ergeben sich aus dem veränderten Berechnungsverfahren und daraus, dass für 2015 mit imputierten Daten gearbeitet wurde.

 © TIMSS 2023

tigten erkennen – mit den bereits erwähnten Einschränkungen hinsichtlich der Interpretierbarkeit. Die Erwartungen der Lehrkräfte fallen 2023 höher aus als in den Vorjahren (insbesondere für die Gruppe der un- und angelernten Arbeiter:innen). Die Erwartungen der Erziehungsberechtigten dagegen fallen 2023 niedriger aus als in den Vorjahren, insbesondere für die Gruppe der (Fach-)Arbeiter:innen.

Für die EGP-Klassen I und II lässt sich feststellen, dass die Unterschiede zwischen den beiden Gruppen 2023 größer ausfallen als 2019 und 2015. Bei den EGP-Klassen III und IV ist am auffälligsten, dass der Schwellenwert der Erziehungsberechtigten aus der Gruppe der Selbstständigen 2023 deutlich geringer ist als in den Vorjahren.

12.6 Zusammenfassung

Die in diesem Kapitel vorgelegten Trendanalysen zum Übergang von der Grundschule in die Sekundarstufe bei TIMSS 2011, 2015, 2019 und 2023, die an die Ergebnisse der fünf bisherigen IGLU-Zyklen anschließen, zeigen ein insgesamt relativ konstantes Bild. Den Angaben der Erziehungsberechtigten zufolge wechselt 2023 knapp die Hälfte der Viertklässler:innen auf ein Gymnasium, jedes vierte Kind auf eine Realschule und jedes fünfte Kind auf eine Schule mit mehreren Bildungsgängen. Hinsichtlich der Schullaufbahnpräferenzen der Lehrkräfte befinden sich die Werte seit 2011 auf einem gleichbleibenden Niveau. In diesem Zusammenhang ist unverändert bedenklich, dass es den Ländern der Bundesrepublik Deutschland nicht gelingt, sich auf einheitliche Namen für die neuen Schulformen zu einigen. Es ist anzunehmen, dass diese Tatsache die Wahl einer weiterführenden Schulform erschwert – und zwar insbesondere in den bildungsferneren Haushalten.

Erwartungskonform zeigt sich, dass leistungsrelevante Merkmale der Schüler:innen entsprechend der bildungspolitischen Vorgaben einen relativ engen Zusammenhang mit der Schullaufbahnpräferenz der Lehrkräfte aufweisen. Schüler:innen mit höheren Testleistungen erhalten mit größerer Chance eine Gymnasialempfehlung, während Schüler:innen mit geringeren Leistungen eher eine Hauptschulempfehlung erhalten. Kinder mit mittleren Testleistungen weisen Präferenzen für alle drei Schulformen auf.

Bedenklich ist hingegen der auch unter Kontrolle der Kompetenzen hohe Zusammenhang zwischen den nicht leistungsrelevanten Hintergrundmerkmalen der Schüler:innen und den Schullaufbahnpräferenzen sowohl der Lehrkräfte als auch der Erziehungsberechtigten. Für TIMSS 2023 konnte dabei eine deutliche Veränderung der sozialen Disparitäten bei Lehrkräften festgestellt werden, deren Interpretation allerdings nicht unproblematisch ist. Für Kinder aus der EGP-Klasse (Fach-)Arbeiter:innen und noch deutlicher für Kinder aus der EGP-Klasse un- und angelernte Arbeiter:innen erwarten die Grundschullehrkräfte sehr hohe Kompetenzen in Mathematik und Naturwissenschaften, damit eine Gymnasialpräferenz hinreichend wahrscheinlich wird. Ohne weitere Daten lässt sich nicht beantworten, ob dies ein einmaliger Ausreißer ist (möglicherweise als Konsequenz der COVID-19-Pandemie) oder ob Lehrkräfte Kindern aus diesen Gruppen tatsächlich keine erfolgreiche Gymnasiallaufbahn zutrauen.

Eine weitere Abweichung von den bisherigen TIMSS-Zyklen zeigt sich bei den Schwellenwerten im Bereich Mathematik hinsichtlich des Migrationshintergrundes. Anders als in der Vergangenheit erwarten Lehrkräfte von Schüler:innen mit Migrationshintergrund eine geringere Kompetenz für eine Gymnasialpräfe-

renz als von Viertklässler:innen ohne Migrationshintergrund. Umgekehrt erwarten Erziehungsberechtigte mit Migrationshintergrund von ihren Kindern eine höhere Kompetenz für eine Gymasialpräferenz als Erziehungsberechtigte ohne Migrationshintergrund. Insgesamt bleibt abzuwarten, ob TIMSS 2023 hinsichtlich der letztgenannten Punkte ein Ausreißer ist oder ob die Befunde in zukünftigen Zyklen reproduziert werden.

Literatur

Albrecht, R., Neumann, M., Jansen, M., Becker, M., Maaz, K., & Baumert, J. (2018). Schulformwechsel im zweigliedrigen Schulsystem. Determinanten für das Nichtbestehen des Probejahres am Gymnasium in Berlin. *Zeitschrift für Erziehungswissenschaft*, *21*, 839–865. https://doi.org/10.1007/s11618-018-0810-6

Arnold, K.-H., Bos, W., Richert, P., & Stubbe, T. C. (2007). Schullaufbahnpräferenzen am Ende der vierten Klassenstufe. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 271–297). Waxmann.

Arnold, K.-H., Bos, W., Richert, P., & Stubbe, T. C. (2010). Der Übergang von der Grundschule in die Sekundarstufe: Schullaufbahnpräferenzen von Lehrkräften und Eltern im Ländervergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert, I. Tarelli & R. Valtin (Hrsg.), *IGLU 2006. Die Grundschule auf dem Prüfstand. Vertiefende Analysen zu Rahmenbedingungen schulischen Lernens* (S. 13–32). Waxmann.

Baumert, J. (2006). Was wissen wir über die Entwicklung von Schulleistungen? *Pädagogik, 58*, 40–46.

Becker, B., & Gresch, C. (2016). Bildungsaspirationen in Familien mit Migrationshintergrund. In C. Diehl, C. Hunkler & C. Kristen (Hrsg.), *Ethnische Ungleichheiten im Bildungsverlauf. Mechanismen, Befunde, Debatten* (S. 73–115). Springer. https://doi.org/10.1007/978-3-658-04322-3_3

Bos, W., Voss, A., Lankes, E.-M., Schwippert, K., Thiel, O., & Valtin, R. (2004). Schullaufbahnempfehlungen von Lehrkräften für Kinder am Ende der vierten Jahrgangsstufe. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich* (S. 191–228). Waxmann.

Cremer, W., & Tobisch, A. (2021). Übergänge in weiterführende Schulformen aus (verfassungs-)rechtlicher Perspektive. *Die Deutsche Schule*, *113*(4), 396–408. https://doi.org/10.31244/dds.2021.04.03

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*(4), 415–441. https://doi.org/10.2307/589632

Glock, S., Krolak-Schwerdt, S., & Pit-ten Cate, I. M. (2015). Are school placement recommendations accurate? The effect of students' ethnicity on teachers' judgments and recognition memory. *European Journal of Psychology of Education*, *30*(2), 169–188. https://doi.org/10.1007/s10212-014-0237-2

Graalmann, K., große Prues, P., Hollen, M., & Thiersch, S. (Hrsg,). (2023.). *Gesamtschule – Status quo und quo vadis?* Waxmann. https://doi.org/10.31244/9783830997467

Klapproth, F., Krolak-Schwerdt, S., Glock, S., Böhmer, M., & Martin, R. (2013). Die prognostische Validität der Sekundarschulempfehlung in Luxemburg: Eine Gegenüberstellung von Verbleibsquoten und Leistungstestdaten. *Schweizerische Zeitschrift für Bildungswissenschaften*, *35*(2), 319–345. https://doi.org/10.25656/01:10275

KMK (Ständige Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland). (2015). Übergang von der Grundschule in Schulen des Sekundarbereichs I und Förderung, Beobachtung und Orientierung in den Jahrgangsstufen 5 und 6 (sog. Orientierungsstufe). http://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2015/2015_02_19-Uebergang_Grundschule-SI-Orientierungsstufe.pdf

Kuckartz, U., Rädiker, S., Ebert, T., & Schehl, J. (2013). *Statistik. Eine verständliche Einführung* (2. Aufl.). Springer. https://doi.org/10.1007/978-3-531-19890-3

Lintorf, K. (2012). *Wie vorhersagbar sind Grundschulnoten? Prädiktionskraft individueller und kontextspezifischer Merkmale*. Springer. https://doi.org/10.1007/978-3-531-94339-8

Maaz, K. (2020). Der lange Arm der sozialen Ungleichheit wirkt bis ins Erwachsenenalter. Entwicklungen und Perspektiven sozialer Disparitäten im Bildungssystem. In C. Fischer, C. Fischer-Ontrup, F. Käpnick, N. Neuber, C. Solzbacher & P. Zwitserlood (Hrsg.), *Begabungsförderung, Leistungsentwicklung, Bildungsgerechtigkeit – für alle! Beiträge aus der Begabungsforschung* (S. 25–41). Waxmann. https://doi.org/10.31244/9783830990666

Maaz, K., & Dumont, H. (2019). Ungleichheiten des Bildungserwerbs nach sozialer Herkunft, Migrationshintergrund und Geschlecht. In O. Köller, M. Hasselhorn, F. Hesse, K. Maaz, J. Schrader, H. Solga, C. K. Spieß & K. Zimmer (Hrsg.), *Das Bildungswesen in Deutschland: Bestand und Potenziale* (S. 299–332). Klinkhardt / utb. https://doi.org/10.36198/9783838547855

Pfost, M., Rausch, T., Schiefer, I. M., & Artelt, C. (2018). Zur Entwicklung von Gymnasiastinnen und Gymnasiasten ohne Gymnasialempfehlung. *Zeitschrift für Erziehungswissenschaft*, *21*(3), 511–534. https://doi.org/10.1007/s11618-017-0787-6

Rabenstein, K., Horn, K.-P., Keßler, C. I., & Stubbe, T. C. (2020). Inklusion als Herausforderung für das Gymnasium. Zur Einleitung in das Thema und in den Band. In K. Rabenstein, T. C. Stubbe & K.-P. Horn (Hrsg.), *Inklusion und Gymnasium. Studien zu Perspektiven von Lehrkräften und Studierenden* (S. 7–34). Universitätsverlag Göttingen. https://doi.org/10.17875/gup2020-1340

Schuchart, C., & Schimke, B. (2019). Lohnt sich das Nachholen eines Schulabschlusses? Alternative Wege zur Hochschulreife und ihre Arbeitsmarkterträge. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *71*(2), 237–273. https://doi.org/10.1007/s11577-019-00621-x

Stubbe, T. C. (2009). *Bildungsentscheidungen und sekundäre Herkunftseffekte. Soziale Disparitäten bei Hamburger Schülerinnen und Schülern der Sekundarstufe I.* Waxmann.

Stubbe, T. C., Bos, W., & Euen, B. (2012). Der Übergang von der Primar- in die Sekundarstufe. In W. Bos, I. Tarelli, A. Bremerich-Vos & K. Schwippert (Hrsg.), *IGLU 2011. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 209–226). Waxmann.

Stubbe, T. C., Bos, W., & Schurig, M. (2017). Der Übergang von der Primar- in die Sekundarstufe. In A. Hußmann, H. Wendt, W. Bos, A. Bremerich-Vos, D. Kasper, E.-M. Lankes, N. McElvany, T. C. Stubbe & R. Valtin (Hrsg.), *IGLU 2016. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 235–250). Waxmann. https://doi.org/10.25656/01:15476

Stubbe, T. C., Kasper, D., & Jentsch, A. (2020). Schullaufbahnpräferenzen am Übergang in die Sekundarstufe und der Zusammenhang mit leistungsrelevanten und sozialen Merkmalen. In K. Schwippert, D. Kasper, O. Köller, N. McElvany, C. Selter, M. Steffensky & H. Wendt (Hrsg.), *TIMSS 2019. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 315–330). Waxmann. https://doi.org/10.31244/9783830993193

Stubbe, T. C., Lorenz, J., Bos, W., & Kasper, D. (2016). Der Übergang von der Primar- in die Sekundarstufe. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 351–365). Waxmann. https://doi.org/10.25656/01:14022

Stubbe, T. C., Schaufelberger, R., Krieg, M., Kleinkorres, R., & Schlitter, T. (2023). Schullaufbahnpräferenzen am Übergang in die Sekundarstufe und der Zusammenhang mit leistungsrelevanten und sozialen Merkmalen. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher & T. C. Stubbe (Hrsg.), *IGLU 2021. Lesekompetenzen von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (S. 231–248). Waxmann. https://doi.org/10.31244/9783830997009

Tiedemann, J., & Billmann-Mahecha, E. (2010). Wie erfolgreich sind Gymnasiasten ohne Gymnasialempfehlung? Die Kluft zwischen Schullaufbahnempfehlung und Schul-

formwahl der Eltern. *Zeitschrift für Erziehungswissenschaft*, *13*, 649–660. https://doi.org/10.1007/s11618-010-0146-3

Trautwein, U., Baeriswyl, F., Lüdtke, O., & Wandeler, C. (2008). Die Öffnung des Schulsystems: Fakt oder Fiktion? Empirische Befunde zum Zusammenhang von Grundschulübertritt und Übergang in die gymnasiale Oberstufe. *Zeitschrift für Erziehungswissenschaft, 11,* 648–665. https://doi.org/10.1007/s11618-008-0043-1

van Leest, A., Hornstra, L., van Tartwijk, J., & van de Pol, J. (2021). Test- or judgement-based school track recommendations: Equal opportunities for students with different socio-economic backgrounds? *British Journal of Educational Psychology*, *91*, 193–216. https://doi.org/10.1111/bjep.12356

Wendt, H., Bos, W., Kasper, D., Walzebug, A., Goy, M., & Jusufi, D. (2016). Ziele, Anlage und Durchführung der Trends in International Mathematics and Science Study (TIMSS 2015). In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 31–77). Waxmann. https://doi.org/10.25656/01:14022

Wendt, H., Kasper, D., Bos, W., Vennemann, M., & Goy, M. (2017). Wie viele Punkte auf der TIMSS-Metrik entsprechen einem Lernjahr? Leistungszuwächse in Mathematik und Naturwissenschaften am Ende der Grundschulzeit. In T. Eckert & B. Gniewosz (Hrsg.), *Bildungsgerechtigkeit* (S. 121–152). Springer. https://doi.org/10.1007/978-3-658-15003-7_8

Anhang

Tabelle 1: Weiterführende Informationen zu den Teilnehmerstaaten – TIMSS 2007

TIMSS 2007		Besonderheiten bzgl. der nationalen Zielpopulation				Besonderheiten bzgl. der nationalen Schul- und Schülerteilnahmequoten				
						Schulteilnahmequote in %		Schülerteilnahmequote in %	Gesamtteilnahmequote in %	
	Teilnehmer	Getestete Jahrgangsstufe	Durchschnittsalter**	Ausschöpfungsgrad***	Ausschlüsse****	*ohne* Ersatzschulen	*mit* Ersatzschulen		*ohne* Ersatzschulen	*mit* Ersatzschulen
	Algerien	4	10.2	100	2.1	99	99	97	97	97
	Armenien	4	10.6	100	3.4	93	100	96	90	96
	Australien	4	9.9	100	4.0	99	100	95	94	95
3	Dänemark	4	11.0	100	4.1	71	91	94	66	85
	Deutschland	4	10.4	100	1.3	96	100	97	93	96
	El Salvador	4	11.0	100	2.3	99	100	98	97	98
1 3	England[A]	5	10.2	100	2.1	83	90	93	77	84
2	Georgien	4	10.1	85	4.8	92	100	98	90	98
2 3	Hongkong	4	10.2	100	5.4	81	84	96	78	81
	Iran	4	10.2	100	3.0	100	100	99	99	99
2	Italien	4	9.8	100	5.3	91	100	97	88	97
	Japan	4	10.5	100	1.1	97	99	97	94	95
	Jemen	4	11.2	100	2.0	99	100	98	97	98
2	Kasachstan	4	10.6	94	5.3	99	100	100	99	100
	Katar	4	9.7	100	1.8	100	100	97	97	97
	Kolumbien	4	10.4	100	2.1	93	99	98	91	97
6	Kuwait	4	10.2	100	0.0	100	100	85	85	85
2	Lettland	4	11.0	72	4.6	93	97	95	89	92
2	Litauen	4	10.8	93	5.4	99	100	94	93	94
3	Marokko	4	10.6	100	1.4	81	81	96	77	77
1 2	Neuseeland[A]	4.5-5.5	10.0	100	5.4	97	100	96	93	96
3	Niederlande	4	10.2	100	4.8	48	95	97	46	91
2	Norwegen	4	9.8	100	5.1	88	97	95	83	92
	Österreich	4	10.3	100	5.0	98	99	98	96	97
	Russische Föderation	4	10.8	100	3.6	100	100	98	98	98
1 3	Schottland[A]	5	9.8	100	4.5	77	94	94	72	88
	Schweden	4	10.8	100	3.1	98	100	97	94	97
	Singapur	4	10.4	100	1.5	100	100	96	96	96
	Slowakei	4	10.4	100	3.3	98	100	97	95	97
	Slowenien	4	9.8	100	2.1	92	99	95	87	93
	Taiwan	4	10.2	100	2.8	100	100	100	100	100
	Tschechische Republik	4	10.3	100	4.9	89	98	94	83	92
	Tunesien	4	10.2	100	2.9	100	100	99	99	99
	Ukraine	4	10.3	100	0.6	96	96	97	93	93
	Ungarn	4	10.7	100	4.4	93	99	97	90	96
2 3	USA	4	10.3	100	9.2	70	89	95	66	84
	Benchmark-Teilnehmer									
2	Alberta, Kanada	4	9.8	100	7.6	99	99	96	94	94
2	Britisch-Kolumbien, Kanada	4	9.8	100	9.2	98	100	96	94	96
2 3 6	Dubai, VAE	4	10.0	100	5.4	75	75	91	67	67
2	Massachusetts, USA	4	10.3	100	10.4	92	96	96	88	92
2 3	Minnesota, USA	4	10.3	100	8.3	53	100	97	52	97
2	Ontario, Kanada	4	9.8	100	6.3	95	96	95	91	92
2	Québec, Kanada	4	10.1	100	6.4	97	98	86	83	84

Tabelle 2: Weiterführende Informationen zu den Teilnehmerstaaten – TIMSS 2019

TIMSS 2019 Teilnehmer		Teilnahme an TIMSS 2019	Teilnahme an Papier-TIMSS 2019	Teilnahme an TIMSS 2015*	Teilnahme an TIMSS 2011*	Teilnahme an TIMSS 2007*	Besonderheiten bzgl. der nationalen Zielpopulation: Getestete Jahrgangsstufe	Durchschnittsalter**	Ausschöpfungsgrad***	Ausschlüsse****	Besonderheiten bzgl. der nationalen Schul- und Schülerteilnahmequoten: Schulteilnahmequote in % ohne Ersatzschulen	Schulteilnahmequote in % mit Ersatzschulen	Schülerteilnahmequote in %	Gesamtteilnahmequote in % ohne Ersatzschulen	Gesamtteilnahmequote in % mit Ersatzschulen	Besonderheiten bzgl. der Leistungsmessung: Anteil der Schüler:innen ohne gültige Leistungswerte in % Naturwissenschaften	Anteil der Schüler:innen ohne gültige Leistungswerte in % Mathematik
7	Albanien		✓				4	10.0	100	4.2	99	99	99	98	98	3	2
8	Armenien		✓	✓	✓	(✓)	4	9.9	100	1.2	100	100	97	97	97	4	4
	Aserbaidschan		✓		✓		4	10.3	100	2.6	94	98	95	89	92	8	4
	Australien		✓	✓	✓	✓	4	10.1	100	4.8	99	100	94	93	94	2	4
	Bahrain		✓	✓	✓		4	9.8	100	0.8	100	100	98	98	98	4	9
3	Belgien (Fläm. Gem.)		✓	✓	✓		4	10.0	100	3.0	66	95	93	62	89	2	1
7	Bosnien und Herzegowina		✓				4	10.1	100	2.0	100	100	96	95	95	5	2
	Bulgarien		✓	✓			4	10.7	100	3.4	97	100	95	92	95	4	5
	Chile	✓		✓	✓		4	10.1	100	3.8	89	99	96	86	95	5	12
3 8	Dänemark	✓		✓	✓	✓	4	10.9	100	3.1	70	95	87	61	83	1	2
	Deutschland	✓		✓	✓	✓	4	10.4	100	3.9	97	100	97	94	97	2	2
1 2 8	England[A]	✓		✓	✓	✓	5	10.2	100	5.8	86	93	96	82	89	1	2
	Finnland	✓		✓	✓		4	10.8	100	3.3	99	100	97	96	97	1	2
	Frankreich	✓		✓			4	9.9	100	4.4	100	100	98	98	98	4	6
2	Georgien[B]	✓		✓	✓	✓	4	10.1	92	4.7	97	99	97	94	96	6	7
3 8	Hongkong	✓		✓	✓	✓	4	10.1	100	3.5	67	88	90	60	79	1	0
8	Iran		✓	✓	✓	✓	4	10.2	100	4.2	100	100	99	99	99	9	14
	Irland		✓	✓	✓		4	10.4	100	3.0	100	100	91	91	91	2	2
8	Italien	✓		✓	✓	✓	4	9.6	100	4.9	96	100	97	92	97	2	2
3 8	Japan		✓	✓	✓	✓	4	10.4	100	2.2	84	98	97	82	95	1	0
2	Kanada[B]	✓		✓			4	9.9	79	7.0	86	90	95	82	86	2	3
2 8	Kasachstan		✓	(✓)	✓	(✓)	4	10.4	100	5.8	100	100	99	99	99	2	2
8	Katar	✓		✓	✓	(✓)	4	9.9	100	2.2	100	100	97	97	97	8	12
2 7	Kosovo		✓				4	9.9	100	8.6	100	100	97	97	97	9	2
	Kroatien	✓		✓	✓		4	10.5	100	4.2	95	97	91	85	87	1	2
7 8	Kuwait		✓	✓	(✓)	(✓)	4	9.7	100	1.7	97	98	96	93	94	14	8
2 8	Lettland		✓			(✓)	4	10.8	100	6.9	92	99	94	87	93	0	1
2 8	Litauen	✓		✓	✓	✓	4	10.7	100	6.7	100	100	94	94	94	1	2
	Malta	✓			✓		4	9.8	100	4.5	100	100	96	96	96	3	4
5 7 8	Marokko		✓	✓	✓	(✓)	4	10.1	100	1.8	100	100	99	99	99	17	6
7	Montenegro		✓				4	9.8	100	4.6	100	100	98	98	98	8	3
1 2 8	Neuseeland[A]		✓	✓	✓	✓	4.5 - 5.5	10.0	100	6.9	87	99	94	81	93	3	6
3 8	Niederlande	✓		✓	✓	✓	4	10.1	100	3.5	46	75	97	45	73	2	1
3 8	Nordirland		✓	✓	✓		4	10.4	100	2.8	60	86	91	55	78	2	2
1 3 8	Norwegen	✓		✓	(✓)	(✓)	5	10.7	100	4.7	70	90	94	66	84	1	2
	Oman		✓	✓	✓		4	9.7	100	2.2	99	100	98	98	98	10	14
2 8	Österreich	✓			✓	✓	4	10.4	100	5.4	99	99	97	97	97	1	1
2 3 4 5 7	Pakistan		✓				4	10.6	100	7.5	77	99	98	75	96	31	16
2 4 5 7	Philippinen		✓				4	10.1	100	7.7	100	100	98	98	98	34	20
8	Polen		✓	✓	(✓)		4	10.3	100	3.1	96	100	93	89	93	1	3
2	Portugal	✓		✓	✓		4	10.0	100	7.8	87	100	94	81	94	2	2
	Republik Korea (Südkorea)	✓		✓	✓		4	10.5	100	2.3	99	99	98	97	97	0	0
7	Republik Nordmazedonien		✓				4	9.8	100	3.8	98	100	95	94	95	11	3
2 8	Russische Föderation	✓		✓	✓	✓	4	10.8	100	6.3	99	99	98	97	97	0	1
2 7	Saudi-Arabien		✓	✓	✓		4	9.9	100	10.5	98	99	99	97	98	12	5
2	Schweden	✓		✓	✓	✓	4	10.8	100	5.4	100	100	95	95	95	1	3
2	Serbien		✓	✓	✓		4	10.6	100	8.2	97	100	97	95	97	3	5
2 8	Singapur	✓		✓	✓	✓	4	10.4	100	12.8	100	100	97	97	97	1	0
2	Slowakei	✓		✓	✓	✓	4	10.4	100	5.5	97	99	97	93	96	2	5
2	Spanien	✓		✓	✓		4	9.9	100	5.4	97	99	95	92	95	1	4
1 5 7	Südafrika[A]		✓	✓			5	11.5	100	1.1	96	99	98	94	97	28	6
	Taiwan	✓		✓	✓	✓	4	10.2	100	2.0	95	99	99	94	98	0	0
	Tschechische Republik	✓		✓	✓	✓	4	10.4	100	4.7	99	100	96	95	96	1	2
1 2	Türkei	✓		(✓)	(✓)		5	10.6	100	7.0	99	100	99	98	99	3	5
	Ungarn	✓		✓	✓	✓	4	10.5	100	4.1	93	99	97	90	96	1	3
2 3	USA	✓		✓	✓	✓	4	10.2	100	7.2	76	88	96	73	84	2	3
	Vereinigte Arabische Emirate (VAE)	✓		✓	✓		4	9.7	100	3.2	100	100	96	96	96	7	8
	Zypern		✓	✓			4	9.8	100	4.6	99	100	97	96	97	2	2
	Benchmark-Teilnehmer																
	Abu Dhabi, VAE	✓		✓	✓		4	9.7	100	3.6	100	100	95	95	95	14	14
2 8	Dubai, VAE	✓		✓	✓	✓	4	9.9	100	5.6	100	100	97	97	97	1	2
	Madrid, Spanien	✓					4	9.9	100	3.5	100	100	96	96	96	1	2
	Moskau, Russische Föderation	✓					4	10.8	100	2.1	99	100	98	97	97	0	0
8	Ontario, Kanada	✓		✓	✓	✓	4	9.8	100	4.9	93	95	95	88	90	2	3
3 8	Québec, Kanada	✓		✓	✓	✓	4	10.1	100	4.0	82	86	96	79	83	1	1

Tabelle 3: Weiterführende Informationen zu den Teilnehmerstaaten – TIMSS 2023

TIMSS 2023 Teilnehmer		Teilnahme an TIMSS 2023*	Teilnahme an TIMSS 2019*	Teilnahme an TIMSS 2015*	Teilnahme an TIMSS 2011*	Teilnahme an TIMSS 2007*	Besonderheiten bzgl. der nationalen Zielpopulation: Getestete Jahrgangsstufe	Durchschnittsalter**	Ausschöpfungsgrad***	Ausschlüsse****	Besonderheiten bzgl. der nationalen Schul- und Schülerteilnahmequoten: Schulteilnahmequote in % ohne Ersatzschulen	Schulteilnahmequote in % mit Ersatzschulen	Schülerteilnahmequote in %	Gesamtteilnahmequote in % ohne Ersatzschulen	Gesamtteilnahmequote in % mit Ersatzschulen	Besonderheiten bzgl. der Leistungsmessung: Anteil der Schüler:innen ohne gültige Leistungswerte in % Naturwissenschaften	Anteil der Schüler:innen ohne gültige Leistungswerte in % Mathematik
2 3	Albanien	✓	✓				4	10.1	100	8.2	54	54	92	48	48	2	2
2 8	Armenien	✓	✓	✓	✓	(✓)	4	9.9	100	6.1	98	99	96	94	95	4	2
	Aserbaidschan	✓	✓		✓		4	10.0	100	2.9	95	99	93	88	92	6	5
	Australien[E]	✓	✓	✓	✓	✓	4	10.1	100	3.7	98	99	91	90	90	1	2
	Bahrain	✓	✓	✓	✓		4	10.0	100	1.1	100	100	98	98	98	5	5
2 3	Belgien (Fläm. Gem.)	✓	✓	✓	✓		4	10.0	100	5.7	74	92	91	67	84	1	1
2	Belgien (Frz. Gem.)	✓					4	9.9	100	8.9	95	99	95	90	94	2	2
2	Bosnien und Herzegowina	(✓)	✓				4	10.2	62	2.5	99	99	91	87	87	4	7
2 3 4	Brasilien	✓					4	10.0	100	6.5	83	95	95	79	90	9	16
	Bulgarien[C]	✓	✓	✓			4	10.7	100	3.7	96	99	95	92	94	2	4
2 3	Chile	✓	✓	✓	✓		4	10.2	100	5.6	79	98	90	71	89	3	7
2 3 8	Dänemark	✓	✓	✓	✓	✓	4	10.9	100	5.2	76	97	94	71	91	1	1
	Deutschland	✓	✓	✓	✓	✓	4	10.4	100	4.4	97	100	95	93	95	2	2
1 2 3 8	England[A E]	✓	✓	✓	✓	✓	5	10.3	100	5.5	84	86	94	79	81	1	1
	Finnland	✓	✓	✓	✓		4	10.8	100	2.4	100	100	95	95	95	1	1
2	Frankreich	✓	✓	✓			4	9.9	100	5.8	99	99	98	97	97	2	3
2	Georgien	✓	✓	✓	✓	✓	4	10.2	93	3.5	97	100	98	95	98	3	3
3 8	Hongkong	✓	✓	✓	✓	✓	4	10.1	100	1.6	76	89	90	68	80	1	1
	Irak[D]	✓					4	10.3	–	–	–	–	–	–	–	–	–
4 8	Iran[C]	✓	✓	✓	✓	✓	4	10.2	100	0.7	100	100	97	97	97	12	21
	Irland	✓	✓	✓	✓		4	10.5	100	3.7	99	100	93	93	93	1	1
2 8	Italien	✓	✓	✓	✓	✓	4	9.8	100	7.0	96	100	96	92	96	1	2
8	Japan	✓	✓	✓	✓	✓	4	10.5	100	2.5	89	94	95	84	90	0	0
	Jordanien	✓		✓			4	9.9	100	0.3	100	100	97	97	97	8	10
2 3	Kanada	✓	✓	✓			4	9.9	79	14.3	81	85	95	77	81	1	2
2 8	Kasachstan	✓	✓	(✓)	✓	(✓)	4	10.2	100	5.8	98	100	98	96	98	3	3
8	Katar	✓	✓	✓	✓	(✓)	4	10.0	100	3.3	100	100	96	96	96	5	5
2	Kosovo	✓	✓				4	10.0	100	8.1	100	100	95	95	95	6	4
5 8	Kuwait[C]	✓	✓	✓	(✓)	(✓)	4	9.8	100	0.7	93	93	96	89	89	18	9
8	Lettland	✓	✓			(✓)	4	10.9	100	4.7	96	99	92	88	91	1	1
2 8	Litauen	✓	✓	✓	✓	✓	4	10.8	100	8.6	99	100	92	91	92	0	0
	Macau	✓					4	9.9	100	3.4	100	100	94	94	94	1	0
8	Marokko[C]	✓	✓	✓	✓	(✓)	4	10.1	100	1.6	99	99	98	98	98	14	6
2	Montenegro	✓	✓				4	9.9	100	8.7	100	100	95	95	95	4	3
1 2 3 8	Neuseeland[A E]	✓	✓	✓	✓	✓	4.5 - 5.5	10.0	100	8.5	63	88	89	56	78	2	3
3 8	Niederlande[E]	✓	✓	✓	✓	✓	4	10.0	100	4.4	54	81	95	51	77	1	1
	Nordmazedonien	✓	✓				4	9.9	100	3.5	99	99	93	92	92	6	5
1 2 8	Norwegen	✓	✓	✓	(✓)	(✓)	5	10.8	100	6.1	95	96	92	87	88	1	1
	Oman	✓	✓	✓	✓		4	9.8	100	2.4	99	100	97	96	97	7	10
2 8	Polen	✓	✓	✓	(✓)		4	10.9	100	6.5	97	100	89	87	89	0	1
	Portugal	✓	✓	✓	✓		4	10.1	100	4.3	97	100	97	93	96	2	1
	Republik Korea (Südkorea)	✓	✓	✓	✓		4	10.4	100	3.4	93	100	96	89	96	0	0
2 3	Rumänien	✓			✓		4	11.1	100	6.5	69	71	87	59	60	2	2
2	Saudi-Arabien	✓	✓	✓	✓		4	9.9	100	11.2	97	100	98	95	98	6	9
2	Schweden	✓	✓	✓	✓	✓	4	10.8	100	6.2	100	100	95	94	95	1	1
2	Serbien	✓	✓	✓	✓		4	10.7	100	10.2	99	99	94	92	93	1	1
2 8	Singapur	✓	✓	✓	✓	✓	4	10.4	100	13.1	100	100	95	95	95	0	0
	Slowakei	✓	✓	✓	✓	✓	4	10.5	100	2.5	99	100	96	95	96	2	2
2	Slowenien	✓		✓	✓	✓	4	10.0	100	5.1	96	99	92	88	91	1	2
2	Spanien	✓	✓	✓	✓		4	10.0	100	6.0	100	100	93	92	92	1	2
1 5	Südafrika[A C]	✓	✓	✓			5	11.3	100	0.2	96	96	87	83	83	27	10
	Taiwan	✓	✓	✓	✓	✓	4	10.3	100	2.6	97	99	96	93	95	0	0
2	Tschechien	✓	✓	✓	✓	✓	4	10.4	100	7.3	100	100	96	96	96	1	1
1 2 3	Türkei	(✓)	✓	(✓)	(✓)		5	10.9	100	21.9	77	100	98	76	98	1	1
2	Ungarn	✓	✓	✓	✓	✓	4	10.6	100	5.1	100	100	96	96	96	2	2
2 3	USA[E]	✓	✓	✓	✓	✓	4	10.2	100	5.5	62	82	93	57	75	2	3
2	Usbekistan	✓					4	10.6	100	5.1	100	100	99	99	99	6	5
	Vereinigte Arabische Emirate (VAE)	✓	✓	✓	✓		4	10.0	100	2.8	98	98	96	94	94	5	4
2	Zypern[C]	✓	✓	✓			4	9.9	100	7.0	99	99	96	95	95	4	4
	Benchmark-Teilnehmer																
	Abu Dhabi, VAE	✓	✓	✓	✓		4	9.9	100	3.5	97	97	97	93	93	8	6
8	Dubai, VAE	✓	✓	✓	✓	✓	4	10.1	100	3.6	99	99	95	95	95	1	1
	Kurdistan (Irakische Gebiete)[D]	✓					4	10.1	–	–	–	–	–	–	–	–	–
2 8	Ontario, Kanada[E]	✓	✓	✓	✓	✓	4	9.8	100	13.2	95	95	95	90	90	1	2
2 3 8	Québec, Kanada	✓	✓	✓	✓	✓	4	10.1	100	8.8	79	85	96	75	81	1	1
	Schardscha, VAE	✓					4	10.0	100	1.4	99	99	95	94	94	3	3

Als Abweichung von der internationalen Vorgabe werden folgende Fälle gekennzeichnet: Getestete Jahrgangsstufe entspricht nicht der vierten Jahrgangsstufe (vgl. Fußnote A); Ausschöpfungsgrad der nationalen Zielpopulation kleiner 100 %; Ausschlüsse von der nationalen Zielpopulation (Gesamtquote) größer 5 %, aber kleiner oder gleich 10 %; Schulteilnahmequote kleiner 85 %, aber größer oder gleich 75 %; Gesamtteilnahmequote kleiner 75 %, aber größer oder gleich 60 %; Anteil der Schüler:innen ohne gültige Leistungswerte in Mathematik bzw. Naturwissenschaften größer 15 %, aber kleiner oder gleich 25 %

Als starke Abweichung von der internationalen Vorgabe werden folgende Fälle gekennzeichnet: Ausschlüsse von der nationalen Zielpopulation (Gesamtquote) größer 10 %; Schulteilnahmequote kleiner 75 %; Gesamtteilnahmequote kleiner 60 %; Anteil der Schüler:innen ohne gültige Leistungswerte in Mathematik bzw. Naturwissenschaften größer 25 %

 = Teilnahme an eTIMSS in TIMSS 2019

 = Teilnahme an Papier-TIMSS in TIMSS 2019

 = Teilnahme an *Less Difficult Mathematics* in TIMSS 2019

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

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

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

4 = Sehr hoher Anteil (mind. 15 %) an Schüler:innen mit nicht skalierbaren Leistungswerten in Mathematik

5 = Sehr hoher Anteil (mind. 15 %) an Schüler:innen mit nicht skalierbaren Leistungswerten in Naturwissenschaften

6 = Abweichender Testzeitpunkt

7 = Teilnehmer haben in TIMSS 2019 an TIMSS *Less Difficult Mathematics* teilgenommen (Mullis et al., 2020)

8 = Bei Trendvergleichen über die Studienzyklen werden einzelne Teilnehmer aufgrund von Änderungen in den Übersetzungen der Testinstrumente oder in den Erhebungsbedingungen nicht berücksichtigt. Darüber hinaus sind bei Trendvergleichen über mehrere Zyklen hinweg folgende Besonderheiten hinsichtlich der Stichproben und Testsprachen zu beachten: Berücksichtigung unterschiedlicher Unterrichtssprachen, Veränderungen hinsichtlich der Testsprachen oder deren Kombination, Änderungen in den Ausschluss- und Teilnahmequoten sowie für 2015 die Teilnahme an TIMSS *Numeracy* bzw. in 2019 an TIMSS *Less Difficult Mathematics* (Cotter et al., 2020; Ebbs et al., 2020; Johansone & Malak, 2008; Martin et al., 2016; Mullis et al., 2016; Mullis et al., 2020; von Davier, Fishbein et al., 2024).

* = Symbole in Klammern indizieren eine Teilnahme am Studienzyklus unter veränderten Testbedingungen, die eine Vergleichbarkeit der Studienzyklen beeinträchtigen und Trendvergleiche ausschließen (Martin et al., 2008; Martin et al., 2016; Martin et al., 2012; Mullis et al., 2008; Mullis et al., 2012; Mullis et al., 2016; Mullis et al., 2020; von Davier, Kennedy et al., 2024).

** = Durchschnittsalter (mit Dezimalstellen) berechnet als: Anzahl der Jahre + Anzahl der Monate/12 (z. B. 9 Jahre, 3 Monate = 9 + 3/12 = 9.25 Jahre)

*** = Ausschöpfungsgrad der nationalen Zielpopulation in Prozent bezogen auf die internationale Vorgabe (100 %)

**** = Ausschlüsse von der nationalen Zielpopulation (Gesamtquote) in Prozent

A = Da das Durchschnittsalter der Viertklässler:innen in diesen Staaten bedeutsam unter der internationalen Vorgabe (9.5 Jahre) lag, wurde gemäß der Kombination der Kriterien ‚formale Beschulungszeit' und ‚Durchschnittsalter des Schülerjahrgangs' eine höhere Jahrgangsstufe als äquivalente nationale Zielpopulation festgelegt (Mullis et al., 2008; Mullis et al., 2020; von Davier, Kennedy et al., 2024).

B = Besonderheiten hinsichtlich Teilnahme und Testung; nähere Informationen finden sich in den internationalen Berichten der jeweiligen Zyklen (Mullis et al., 2008; Mullis et al., 2020; von Davier, Kennedy et al., 2024)

C = Papierbasierte Kompetenztestung

D = Aufgrund international nicht vergleichbarer Daten werden die Teilnehmer von der Berichtslegung ausgeschlossen (von Davier, Kennedy et al., 2024).

E = Für Australien, England, die Niederlande und die USA liegen für den Elternfragebogen keine Ergebnisse vor. Die Ergebnisse von Neuseeland und Ontario, Kanada werden aufgrund der geringen Rücklaufquoten des Elternfragebogens nicht berichtet (von Davier, Kennedy et al., 2024).

Die in diesem Anhang berichteten Informationen basieren auf dem Stand vom 06.09.2024.

Literatur

Cotter, K. E., Centurino, V. A. S., & Mullis, I. V. S. (2020). Developing the TIMSS 2019 mathematics and science achievement instruments. In M. O. Martin, M. von Davier & I. V. S. Mullis (Hrsg.), *Methods and procedures: TIMSS 2019 technical report* (S. 1.1–1.36). TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2019/methods/chapter-1.html

Ebbs, D., Wagner J.-P., & Netten, A. (2020). Instrument translation and layout verification for TIMSS 2019. In M. O. Martin, M. von Davier & I. V. S. Mullis (Hrsg.), *Methods and procedures: TIMSS 2019 technical report* (S. 5.1–5.23). TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2019/methods/chapter-5.html

Johansone, I., & Malak, B. (2008). Translation and national adaptations of the TIMSS 2007 assessment and questionnaires. In J. F. Olson, M. O. Martin & I. V. S. Mullis (Hrsg.), *TIMSS 2007 technical report* (S. 63–75). TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/TIMSS2007/techreport.html

Martin, M. O., Mullis, I. V. S., & Foy, P. (2008). *TIMSS 2007 international science report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grad*es. TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/TIMSS2007/PDF/TIMSS2007_InternationalScienceReport.pdf

Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in science*. TIMSS & PIRLS International Study Center. http://timssandpirls.bc.edu/timss2015/international-results/

Martin, M. O., Mullis, I. V. S., Foy, P., & Stanco, G. M. (2012). *TIMSS 2011 international results in science*. TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2011/downloads/T11_IR_Science_FullBook.pdf

Mullis, I. V. S., Martin, M. O., & Foy, P. (2008). *TIMSS 2007 international mathematics report. Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. TIMSS & PIRLS International Study Center. https://timss.bc.edu/TIMSS2007/mathreport.html

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2011/downloads/T11_IR_Mathematics_FullBook.pdf

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in mathematics*. TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/publications/timss/2015-methods.html

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science.* TIMSS & PIRLS International Study Center. http://timssandpirls.bc.edu/timss2019/international-results/

von Davier, M., Fishbein, B., & Kennedy, A. (Hrsg.). (2024). *TIMSS 2023 technical report (methods and procedures)*. TIMSS & PIRLS International Study Center.

von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U., & Yin, L. (2024). *TIMSS 2023 international results in mathematics and science*. TIMSS & PIRLS International Study Center.

Zum Zeitpunkt der Veröffentlichung des vorliegenden Bandes waren der Technical Report und der finale internationale Bericht zu TIMSS 2023 noch nicht publiziert.

Weiterführende Informationen sind auf der internationalen Projekt-Website zu finden: https://timssandpirls.bc.edu/timss2023/