

Mathematik im Fokus

Christine Sälzer

Studienbuch Schulleistungs- studien

Das Rasch-Modell in der Praxis



Springer Spektrum

Mathematik im Fokus

Kristina Reiss

TU München, School of Education, München, Germany

Ralf Korn

TU Kaiserslautern, Fachbereich Mathematik, Kaiserslautern, Germany

Weitere Bände in dieser Reihe:

<http://www.springer.com/series/11578>

Die Buchreihe Mathematik im Fokus veröffentlicht zu aktuellen mathematikorientierten Themen gut verständliche Einführungen und prägnante Zusammenfassungen. Das inhaltliche Spektrum umfasst dabei Themen aus Lehre, Forschung, Berufs- und Unterrichtspraxis. Der Umfang eines Buches beträgt in der Regel 80 bis 120 Seiten. Kurzdarstellungen der folgenden Art sind möglich:

- State-of-the-Art Berichte aus aktuellen Teilgebieten der theoretischen und angewandten Mathematik
- Fallstudien oder exemplarische Darstellungen eines Themas
- Mathematische Verfahren mit Anwendung in Natur-, Ingenieur- oder Wirtschaftswissenschaften
- Darstellung der grundlegenden Konzepte oder Kompetenzen in einem Gebiet

Mathematik im Fokus ist ein zeitnahes Spiegelbild aktueller Themen, die aus der Sicht der Mathematik kompakt dargestellt und kommentiert werden.

Christine Sälzer

Studienbuch Schulleistungsstudien

Das Rasch-Modell in der Praxis



Springer Spektrum

Christine Sälzer
TUM School of Education
ZIB - Zentrum für Internationale Vergleichsstudien
München, Deutschland

Mathematik im Fokus
ISBN 978-3-662-45764-1
DOI 10.1007/978-3-662-45765-8

ISBN 978-3-662-45765-8 (eBook)

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

Springer Spektrum

© Springer-Verlag Berlin Heidelberg 2016

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen.

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier.

Springer Spektrum ist eine Marke von Springer DE.
Springer DE ist Teil der Fachverlagsgruppe Springer Science+Business Media
www.springer-spektrum.de

Inhaltsverzeichnis

1	Einleitung	1
2	Die großen Schulleistungsstudien in Deutschland	5
2.1	PISA: Das <i>Programme for International Student Assessment</i>	7
2.1.1	Kompetenzbegriff in PISA: Literacy	9
2.1.2	Design und Testkonzeption von PISA	12
2.2	TIMSS: Die Trends in International Mathematics and Science Study	15
2.2.1	Kompetenzbegriff in TIMSS	17
2.3	IGLU: <i>Internationale Grundschul-Lese-Untersuchung</i> (engl. PIRLS)	20
2.3.1	Kompetenzbegriff in IGLU	22
2.3.2	Design und Testkonzeption von TIMSS und IGLU	23
2.4	IQB-Ländervergleich zur Überprüfung der Bildungsstandards	26
2.4.1	Kompetenzbegriff im IQB-Ländervergleich	28
2.4.2	Design und Testkonzeption des IQB-Ländervergleichs	30
2.5	VERA: Vergleichsarbeiten in der Schule	33
2.6	NEPS: Das Nationale Bildungspanel – <i>National Educational Panel Study</i>	35
2.6.1	Kompetenzbegriff in NEPS	36
2.6.2	Design und Testkonzeption von NEPS	37
2.7	Zusammenfassender Überblick: Schulleistungsstudien in Deutschland	39
3	Grundzüge des Rasch-Modells	43
3.1	Erfassung von nicht beobachtbaren Eigenschaften	44
3.2	Datenmatrix: Antworten auf Testaufgaben	45
3.3	Modellgleichung	47
3.4	Personen- und itemcharakteristische Kurven	48

4	Schätzung von Kompetenzen in Studien wie PISA	53
4.1	Vergleichbarkeit von Schülerleistungen	54
4.2	Skalierung der Daten	55
4.3	Schätzung von Parametern im Rasch-Modell.....	57
4.4	Mehrdimensionale Erweiterungen des Rasch-Modells	60
5	Schulleistungsstudien lesen und verstehen	63
5.1	Schätzung von Kompetenzen mit einer Auswahl von Aufgaben	64
5.2	Individuelle Rückmeldung an Schulen und Schüler	65
5.3	Ergebnisse in Abhängigkeit der gezogenen Schüler	65
5.4	Vergleichbarkeit der Kompetenzen über Staaten und Bildungssysteme hinweg	67
5.5	<i>Teaching to the Test</i> : Vorteile durch gezieltes Üben?.....	71
5.6	Richtige, teilweise richtige und falsche Antworten auf Testfragen ..	73
5.7	Schüler als nachwachsender Rohstoff: Ist PISA eine Studie für die Wirtschaft?	74
6	Bilanz: Schulleistungsstudien und Kompetenzmessung	77
6.1	Mehrere Schulleistungsstudien in Deutschland	77
6.2	Kompetenzmessung durch Tests und Schätzverfahren	78
6.3	Weiterführende Fragen	79
	Literaturverzeichnis	81

Kapitel 1

Einleitung

Die PISA-Studie (*Programme for International Student Assessment*) steht in Deutschland mittlerweile stellvertretend für eine ganze Reihe von Schulleistungstudien. Seit dem Jahr 2000 wird PISA alle drei Jahre in den OECD-Staaten und einer wachsenden Zahl von Partnerstaaten durchgeführt und bringt als Ergebnisse eine Sammlung von Indikatoren hervor, anhand derer sich relative Stärken und Schwächen der beteiligten Bildungssysteme ableiten lassen. PISA 2000 war nicht die erste international vergleichende Schulleistungstudie, an der Deutschland teilnahm, aber nach längerer Zeit eine der ersten relativ umfassenden Untersuchungen, die im Dezember 2001 unerwartet schwache Ergebnisse hervorbrachte (Baumert et al. 2001). In allen drei untersuchten Kompetenzdomänen, Lesen, Mathematik und Naturwissenschaften, lagen die Leistungen der Schülerinnen und Schüler in Deutschland unterhalb des OECD-Durchschnitts – damals der Auslöser des sogenannten PISA-Schocks. Dass heute nach wie vor vom PISA-Schock gesprochen wird, ist bemerkenswert, denn bereits einige Jahre vor der Veröffentlichung der ersten PISA-Runde hatten die Ergebnisse der ersten TIMSS-Studie (*Trends in International Mathematics and Science Study*) Schülerinnen und Schülern in Deutschland vergleichsweise schwache Leistungen in Mathematik und den Naturwissenschaften bescheinigt (Baumert et al. 1997; Beaton et al. 1996).

Im Zusammenhang mit PISA und dem PISA-Schock wird noch immer Bezug auf die anfangs überraschend mittelmäßigen Ergebnisse genommen. Allerdings hat sich seither manches verändert, wenn man die Ergebnisse weiterer PISA-Erhebungsrunden oder auch anderer Schulleistungsvergleiche betrachtet. Zum einen erzielen die Schülerinnen und Schüler in Deutschland nach mittlerweile fünf abgeschlossenen Erhebungsrunden Leistungen, die deutlich über dem Durchschnitt der OECD-Staaten liegen (vgl. etwa Hohn et al. 2013; Sälzer, Reiss et al. 2013; Schiepe-Tiska et al. 2013). Aufgeholt haben insbesondere Jugendliche an nicht gymnasialen Schulformen, Jugendliche mit Zuwanderungshintergrund sowie eher leistungsschwache Schülerinnen und Schüler. Zum anderen widmet sich PISA inzwischen ausschließlich dem internationalen Vergleich der Schülerkompetenzen, nachdem die Studie ursprünglich in Deutschland in zwei Teilstudien aufgliedert war. Neben dem internationalen Vergleich (PISA-I für *PISA-International*) unternahm PISA in

den ersten Jahren auch einen Vergleich der 16 deutschen Bundesländer (PISA-E für *PISA-Erweiterung*). Seit 2009¹ sind beide Studien aufgrund ihres großen Umfangs institutionell getrennt und werden separat durchgeführt.

Dieses Buch ist aus der Koordination der PISA-Studie in Deutschland heraus entstanden, die seit 2010 am Zentrum für internationale Vergleichsstudien (ZIB e. V.) an der TU München ihren Sitz hat. Im Laufe der Arbeit an und mit PISA tauchten in Gesprächen mit Studierenden, Lehrkräften, Schulleitungen oder weiteren Interessierten immer wieder Fragen zu PISA selbst auf, aber auch zur Mess- und Vergleichbarkeit von Kompetenzen und Leistungen. Die wohl kniffligste unter diesen Fragen ist diejenige, wie und warum Schülerkompetenzen durch eine begrenzte Anzahl von Aufgaben pro Schüler gemessen werden können, die sich noch nicht einmal auf den jeweiligen Lehrplan dieser Schülerinnen und Schüler beziehen. Wenn sogar innerhalb Deutschlands unterschiedliche Lehrpläne vorliegen, wie soll dann ein internationaler Vergleich möglich und sinnvoll sein? Diese Frage verknüpft zwei Aspekte: Einerseits die Messung von Kompetenzen und andererseits die sogenannte curriculare Validität, also den Lehrplanbezug der Testaufgaben, anhand derer die Messung erfolgt. Um zu verstehen, wie die Messung von Kompetenzen in zahlreichen Schulleistungsstudien funktioniert, sollte man beide Aspekte nachvollziehen. Was bedeutet eigentlich „Messen“ im Zusammenhang mit nicht beobachtbaren Eigenschaften wie Kompetenzen? Welche Annahmen stehen dahinter? Und kann man sich nicht auch sehr täuschen, wenn man aus den Antworten von Schülerinnen und Schülern auf einige Testfragen auf deren Kompetenz schließt? Noch dazu, wenn die Jugendlichen die Themen der Aufgaben teilweise noch gar nicht im Unterricht hatten? Und sind angesichts der unterschiedlichen Bildungssysteme und Voraussetzungen der Schülerinnen und Schüler deren Leistungen überhaupt international miteinander vergleichbar? Fragen wie diese leiteten die Arbeit am vorliegenden Buch. Es richtet sich vorwiegend an Studierende in den Bereichen Lehramt oder Bildungswissenschaften. Die einzelnen Kapitel sind so geschrieben, dass sie auch jedes für sich gelesen und verstanden werden können. Es ist als einführendes Studienbuch zur Auseinandersetzung mit den Ergebnissen aus Untersuchungen wie PISA gedacht.

Das Buch gibt zunächst einen Überblick über die großen Studien, an denen Deutschland derzeit teilnimmt (Kapitel 2). All diese Studien haben gemeinsam, dass sie bestimmte Kompetenzen messen. Sie unterscheiden sich jedoch darin, was unter Kompetenz im Detail verstanden wird, sowie bezüglich ihrer Zielgruppen und Schwerpunkte. Das Design einer Schulleistungsstudie sowie die zu Grunde gelegte Testkonzeption sind zwei wesentliche Elemente, welche die Ergebnisse der Studie sowie deren Interpretation bestimmen. Unter dem Design einer Studie versteht man deren Anlage, Aufbau und letztlich das planvolle Vorgehen bei der Datenerhebung und -auswertung. Mit der Testkonzeption einer Schulleistungsstudie ist die theoretische Grundlage für die Aufgabenentwicklung sowie für die Organisation des gemessenen Konstrukts gemeint (im Kontext von Schulleistungsstudien sind dies Kompetenzen bei Schülerinnen und Schülern). Daher werden in Kapitel 2 für jede der dort aufgeführten Studien sowohl der Kompetenzbegriff als auch das Design

¹ Bereits 2008 erfolgte die Erhebung für den Ländervergleich der Bildungsstandards für das Fach Französisch als erste Fremdsprache in Anknüpfung an die entsprechende Normierungsstudie.

und die Testkonzeption in aller Kürze beschrieben. Diese sehr verdichtete Darstellung soll einen Einstieg und Überblick über die Studienlandschaft in Deutschland geben. Detaillierte Beschreibungen der genannten Studien sind in der jeweils zitierten Literatur zu finden, die größtenteils online verfügbar und kostenlos zugänglich ist.

Um Kompetenzen von Schülerinnen und Schülern messen, abbilden und miteinander vergleichen zu können, werden häufig Modelle verwendet, die bestimmte Annahmen zu Grunde legen. In Studien wie PISA wird dazu häufig das nach seinem Erfinder Georg Rasch benannte „Rasch-Modell“ eingesetzt. In Kapitel 3 wird das Rasch-Modell in Grundzügen vorgestellt und anschließend seine Bedeutung für Schulleistungsstudien beschrieben und diskutiert. Kapitel 4 verknüpft die Themen der beiden vorangehenden Kapitel und geht auf die konkrete Anwendung des Rasch-Modells und seiner Annahmen in Schulleistungsstudien ein. Den Abschluss bilden Kapitel 5 und 6 mit einer zusammenfassenden Verdichtung von Lesehilfen für die beschriebenen Schulleistungsstudien, in der die zentralen Aspekte der Messung von Schülerkompetenzen anhand einer Auswahl von Testaufgaben nochmals aufgegriffen und durch Literaturempfehlungen ergänzt werden.