

PISA 2000

Ein differenzierter Blick auf die Länder
der Bundesrepublik Deutschland

Zusammenfassung zentraler Befunde



OECD
P I S A
Deutschland

Baumert, Artelt, Klieme, Neubrand, Prenzel, Schiefele,
Schneider, Schümer, Stanat, Tillmann, Weiß (Hrsg.)

Berlin: Max-Planck-Institut für Bildungsforschung, 2003

Zusammenfassung zentraler Befunde

aus

Jürgen Baumert, Cordula Artelt, Eckhard Klieme, Michael Neubrand, Manfred Prenzel, Ulrich Schiefele, Wolfgang Schneider, Klaus-Jürgen Tillmann, Manfred Weiß (Hrsg.), PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland. Opladen: Leske + Budrich.

1	Untersuchungsgegenstand und Durchführung der Studie	3
2	Literacy oder Lehrplanvalidität? Ländervergleiche auf der Basis lehrplanoptimierte PISA-Tests	9
3	Naturwissenschaftliche Teilkompetenzen im Ländervergleich	17
4	Selbstreguliertes Lernen: Motivation und Strategien in den Ländern der Bundesrepublik Deutschland	23
5	Soziale Lernziele im Ländervergleich	29
6	Computernutzung	35
7	Kompetenzerwerb, Bildungsbeteiligung und Schullaufbahn von Mädchen und Jungen im Ländervergleich	41
8	Schulleistungen von Jugendlichen mit Migrationshintergrund: Differenzierung deskriptiver Befunde aus PISA und PISA-E	51
9	Schulumwelten – institutionelle Bedingungen des Lehrens und Lernens	59
10	Unterrichtsqualität aus Schülerperspektive: Kulturspezifische Profile, regionale Unterschiede und Zusammenhänge mit Effekten von Unterricht	73
11	Familienstrukturen, Bildungslaufbahnen und Kompetenzerwerb	77
12	Möglichkeiten und Grenzen der Nutzung von Schulrückmeldungen im Rahmen von Schulleistungsuntersuchungen: Das Disseminationskonzept von PISA-2000	85
13	Literatur	89

2 **Literacy oder Lehrplanvalidität? – Ländervergleiche auf der Basis lehrplanoptimierter PISA-Tests**

Die Frage, ob sich anhand der internationalen PISA-Tests ein realistisches Bild des Leistungsstandes von Schülerinnen und Schülern in den Ländern der Bundesrepublik zeichnen lässt, hat zu Beginn des PISA-Projektes z.T. zu leidenschaftlichen Debatten geführt. Die Rahmenkonzeption des internationalen Tests bot hierzu Anlass; statt von einem lehrplanbasierten Modell der Wissensüberprüfung und Fähigkeitsdefinition auszugehen, wurde dem internationalen Test ein normatives Modell zu Grunde gelegt. Ausgangspunkt dieser normativen Definition von Basiskompetenzen in den drei in PISA untersuchten Kompetenzbereichen (Lesekompetenz, mathematische und naturwissenschaftliche Grundbildung) war die Beschreibung derjenigen Fähigkeiten, über die Schülerinnen und Schüler am Ende der regulären Schulzeit verfügen können sollten, um auf die Herausforderung des Berufslebens und der mündigen Teilhabe am gesellschaftlichen Leben vorbereitet zu sein.

Mit der Konstruktion der internationalen PISA-Tests für die drei Leistungsbereiche war also die Absicht verbunden, abzubilden, in welchem Maße Schülerinnen und Schüler am Ende der regulären Schulzeit über Basiskompetenzen im Bereich Lesen, Mathematik und Naturwissenschaften verfügen. Die Testentwickler beabsichtigten jedoch nicht, die Curricula der betroffenen Unterrichtsfächer in den Teilnehmerstaaten von PISA umfassend abzubilden. Insofern ist es nicht erstaunlich, dass einige der im Test verwendeten Aufgaben hinsichtlich der gestellten Anforderungen bzw. der Art der Aufgabenstellung von dem abweichen, was in den Lehrplänen der Länder der Bundesrepublik explizit thematisiert wird. Auch ist es nicht verwunderlich, dass einzelne Themen und Inhaltsbereiche, die in Lehrplänen vorgesehen werden, in PISA nicht oder mit einem anderen Schwerpunkt berücksichtigt werden. Zu einem solchen Eindruck kommt man vor allem dann, wenn man versucht, die Inhalte des PISA-Tests allein mit einem Unterrichtsfach in Beziehung zu setzen. Während dies im Bereich des mathematischen Lernens am ehesten zu rechtfertigen ist, scheitert dieses Vorhaben spätestens dann, wenn es um die Lesekompetenz geht. Das Vermittlungsmonopol für den verstehenden Umgang mit Textmaterial liegt in der Sekundarstufe I nicht allein beim Deutschunterricht. Mit wenigen Ausnahmen ist das Interpretieren von Texten, die Reflexion über und das Bewerten von Texten und Textteilen sowie die Entnahme von Informationen aus Texten eine zentrale Voraussetzung für gelingenden Unterricht, egal welcher Art und fachlichen Ausrichtung. Insbesondere die Tatsache, dass Texte in PISA auch als Tabel-

	Lesekompetenz (N = 129)	Mathematische Grundbildung (N = 31)	Naturwissenschaft- liche Grundbildung* (N = 34)	Insgesamt (N = 194)
Baden-Württemberg	76 (59 %)	21 (68 %)	12 (35 %)	109 (56 %)
Bayern	81 (63 %)	11 (35 %)	14 (41 %)	106 (55 %)
Berlin	43 (33 %)	8 (26 %)	11 (32 %)	62 (32 %)
Brandenburg	94 (73 %)	23 (74 %)	23 (68 %)	140 (72 %)
Bremen	81 (63 %)	11 (35 %)	22 (65 %)	114 (59 %)
Hamburg	74 (57 %)	19 (61 %)	17 (50 %)	110 (57 %)
Hessen	70 (54 %)	30 (97 %)	17 (50 %)	117 (60 %)
Mecklenburg-Vorp.	91 (71 %)	25 (81 %)	18 (53 %)	134 (69 %)
Niedersachsen	65 (50 %)	24 (77 %)	3 (9 %)	92 (47 %)
Nordrhein-Westfalen	82 (64 %)	21 (68 %)	12 (35 %)	115 (59 %)
Rheinland-Pfalz	86 (67 %)	15 (48 %)	16 (47 %)	117 (60 %)
Saarland	58 (45 %)	11 (35 %)	15 (44 %)	84 (43 %)
Sachsen	108 (84 %)	23 (74 %)	22 (65 %)	153 (79 %)
Sachsen-Anhalt	99 (77 %)	29 (94 %)	16 (47 %)	144 (74 %)
Schleswig-Holstein	87 (67 %)	11 (35 %)	15 (44 %)	113 (58 %)
Thüringen	119 (92 %)	26 (84 %)	14 (41 %)	159 (82 %)

* Die Bewertung wurde für Biologie, Physik, Chemie und Erdkunde getrennt vorgenommen.

Tabelle 1: Anzahl und Anteil der Aufgaben aus den internationalen Tests, die von den Lehrplanexperten der Länder als lehrplanvalide eingestuft wurden

len, Graphiken und schematische Zeichnungen aufgefasst werden, kennzeichnet eine Überlappung mit den Inhalten und gängigen Textsorten anderer Unterrichtsfächer (vgl. OECD, 1999).

Die Frage, inwiefern der internationale PISA-Test mit den in deutschen Lehrplänen gestellten Anforderungen vereinbar ist, wurde bereits mehrfach thematisiert. Es wäre genauso verkürzt, dem gesamten PISA-Test pauschal Lehrplanpassung zu unterstellen, wie ihm diese generell abzusprechen. In Abhängigkeit vom Fach ergeben sich durchaus differenzierte Passungsbilder (s. Artelt u.a., 2001; Klieme, Neubrand, & Lüdtke, 2001; Prenzel u.a., 2001). Um die unterschiedlich weit und allgemein gehaltenen Formulierungen in Lehrplänen mit den konkreten PISA-Aufgabenanforderungen in Beziehung zu setzen und letztlich einen lehrplanoptimierten Test pro Fach und Land zu erstellen, wurden im Rahmen von PISA 2000 alle PISA-Aufgaben von Lehrplanexpertinnen und -experten der 16 Länder hinsichtlich ihrer Passung mit den Lehrplänen der Sekundarstufe I beurteilt. Dabei wurde für jedes Land und jede Aufgabe ermittelt, ob diese Aufgabe mit den Lehrplänen des jeweiligen Fachs (Deutsch, Mathematik, Naturwissenschaften [Biologie, Chemie, Physik und Erdkunde]) für die jeweiligen Schulformen des Landes übereinstimmt. Eine Aufgabe wurde dann als lehrplanvalide angesehen, wenn alle Experten pro Land¹ der Meinung waren, dass die Lösung der Aufgabe vor der 9. Klassenstufe

erwartet werden kann. Tabelle 1 gibt einen Überblick über die Anzahl und den Anteil der Aufgaben, die jeweils als lehrplanvalide angesehen wurden. Zwischen den Ländern – besonders zwischen den alten und den neuen Ländern – bestehen erhebliche Unterschiede in der Anzahl der Aufgaben, die als mit dem Lehrplan des Landes vereinbar angesehen werden².

Das Ziel der Konstruktion von lehrplanoptimierten Ländertests bestand nicht darin, den Ländern curricular valide Tests zur Verfügung zu stellen, sondern einen Einblick in die Aussagekraft der internationalen PISA-Tests vor dem Hintergrund der Lehrpläne der Länder zu bekommen.

Auf der Basis dieser konsensuell ermittelten Aufgabenmenge pro Fach wurde für jedes Land ein individueller lehrplanoptimierter Ländertest erstellt. Grundlage für die Auswertung war die Erhebung im Jahre 2000. In die Rasch-Skalierung der Leistungsdaten der Schüler jedes Landes gingen jedoch immer nur die Aufgaben ein, die von den Experten als valide eingeschätzt wurden.

Durch die Skalierung der optimierten Ländertests analog zum internationalen Vorgehen (s.a. OECD, 2002b) ist es möglich, das Abschneiden 15-jähriger Schülerinnen und Schüler international zu verorten. In einem zweiten Schritt wurde auch die Gruppe der Neuntklässler separat analysiert. Die Ergebnisse dieser Schülergruppe sind aufgrund der anderen Zusammensetzung nicht mehr mit denen aus den anderen Teilnehmerstaaten vergleichbar. Da das Validitätskriterium für die PISA-Aufgaben jedoch für die Jahrgangsstufe 9 besser geeignet ist als für die 15-Jährigen, die sich auf unterschiedliche Klassenstufen verteilen, wurden die Auswertungen der länderspezifischen Tests für die Neuntklässler wiederholt. Beim Vergleich der durchschnittlichen Leistungen der Schülerinnen und Schüler im internationalen PISA-Lesetest mit denen in den verkürzten lehrplanoptimierten Ländertests ergaben sich jedoch zwischen der Gruppe der 15-Jährigen und der der Neuntklässler keine nennenswerten Unterschiede. Bei der Darstellung der wesentlichen Befunde konzentrieren wir uns daher auf die Leistungsvergleiche für die 15-Jährigen.

Wie verändern sich die Länderergebnisse, wenn optimierte Ländertests zu Grunde gelegt werden?

Abbildung 1 stellt die Mittelwerte der beiden Testversionen (internationaler PISA-Lesetest vs. lehrplanoptimierter Ländertest) pro Land nebeneinander. Die Höhe der Balken gibt jeweils den Mittelwert an. Das Vertrauensintervall um den Mittelwert³ ist durch die zusätzlich eingezeichnete Linie gekennzeichnet. Zwar stimmen die absoluten Werte der beiden Testversionen pro Land nicht immer überein – in Brandenburg, wo die größten Unterschiede existieren, ergibt sich eine Mittelwertdifferenz von 5 Punkten –, die Vertrauensintervalle um die beiden Mittelwerte pro Land überlappen sich jedoch deutlich, sodass die geringen Differenzen nicht statistisch abgesichert werden können. Auch

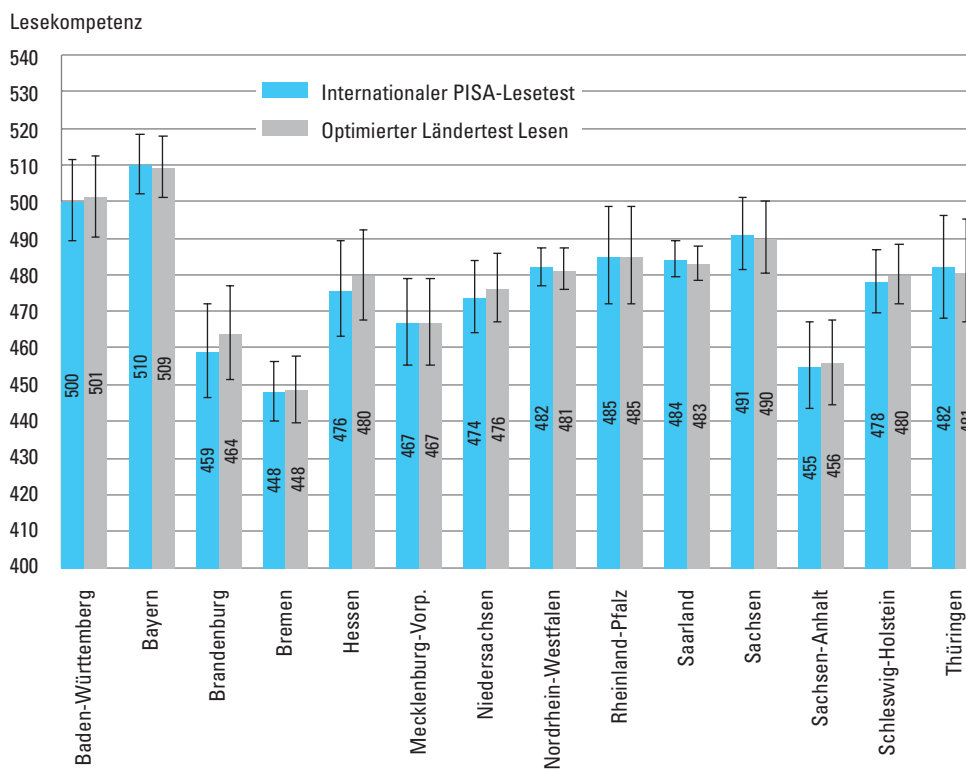


Abbildung 1: Lesekompetenz: Mittelwerte im internationalen PISA-Lesetest und im lehrplanoptimierten Ländertest Lesen

die Rangfolge der Länder verändert sich durch die Berücksichtigung von optimierten Ländertests nicht.

Insgesamt gesehen steigt der Mittelwert für die gesamte Bundesrepublik beim lehrplanoptimierten Lesetest um zwei Punkte im Vergleich zum vollständigen internationalen PISA-Lesetest. Während er bei der Originalskalierung bei 484 und damit signifikant unterhalb des OECD-Durchschnitts lag, ergibt der lehrplanoptimierte Test einen Mittelwert von 486, was immer noch signifikant unter dem OECD-Durchschnitt von 500 liegt und auch zu keiner Verschiebung der Rangposition Deutschlands beim internationalen Vergleich führt (vgl. Artelt u.a., 2001). – Auch für die Verteilung der Schülerinnen und Schüler auf die einzelnen Kompetenzstufen im Lesen ergeben sich keine nennenswerten Unterschiede. In keinem Land ändern sich die Ergebnisse des curricular validen Ländertests pro Kompetenzstufe um mehr als einen Prozentpunkt.

Im Bereich mathematische Grundbildung kommt es aufgrund der Aufgabenauswahl zu größeren Veränderungen im mittleren Leistungsniveau, als dies im Bereich Lesekompetenz der Fall ist (s. Abb. 2). Besonders deutlich zu sehen ist dies in Bayern und in Rheinland-Pfalz, wo ein Mittelwertunterschied von je 11 Punkten zwischen beiden Testversionen besteht. Allerdings führt die Auswahl lehrplanvalider Mathematikaufgaben in

Mathematische Grundbildung

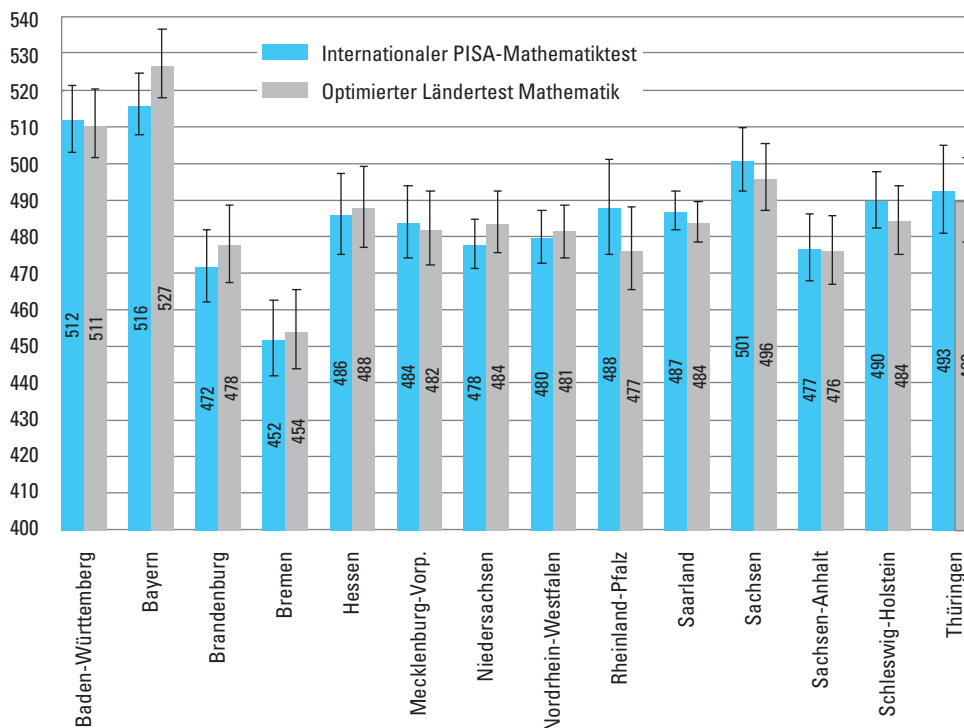


Abbildung 2: Mathematische Grundbildung: Mittelwerte im internationalen PISA-Mathematiktest und im optimierten Ländertest Mathematik

Bayern zu einer „Verbesserung“, in Rheinland-Pfalz hingegen zu einer „Verschlechterung“ des mittleren Leistungsniveaus. Auch wenn ein Unterschied von 11 Punkten bedeutsam erscheinen mag, handelt es sich doch jeweils um eine statistisch nicht nachweisbare Differenz. Wie aus Abbildung 2 zu entnehmen ist, überlappen sich die Vertrauensintervalle um die Mittelwerte des vollständigen internationalen PISA-Mathematiktests und des lehrplanoptimierten Ländertests in allen Ländern, auch in Bayern und Rheinland-Pfalz. Die Abweichungen in den Ländermittelwerten sind daher nicht als bedeutsamer Unterschied zu interpretieren, sondern sie könnten aufgrund der mit Unsicherheit (ausgedrückt durch den Standardfehler) behafteten Schätzung der Leistungsmittelwerte auch allein auf Zufallsschwankungen zurückzuführen sein.

Bezüglich der Verteilung der Schüler auf die Kompetenzstufen, und damit auch der Größe der Risikogruppe und der Schülergruppe, die das vorab definierte Standardniveau mathematischer Grundbildung erreicht (vgl. Klieme u.a., 2001), ergeben sich Unterschiede von maximal 4 Prozentpunkten zwischen den Ergebnissen des internationalen PISA-Mathematiktests und denen der optimierten Ländertests. Abweichungen kommen aber in beiderlei Richtungen vor, d.h. es gibt auch Länder, bei denen die Größe der Risi-

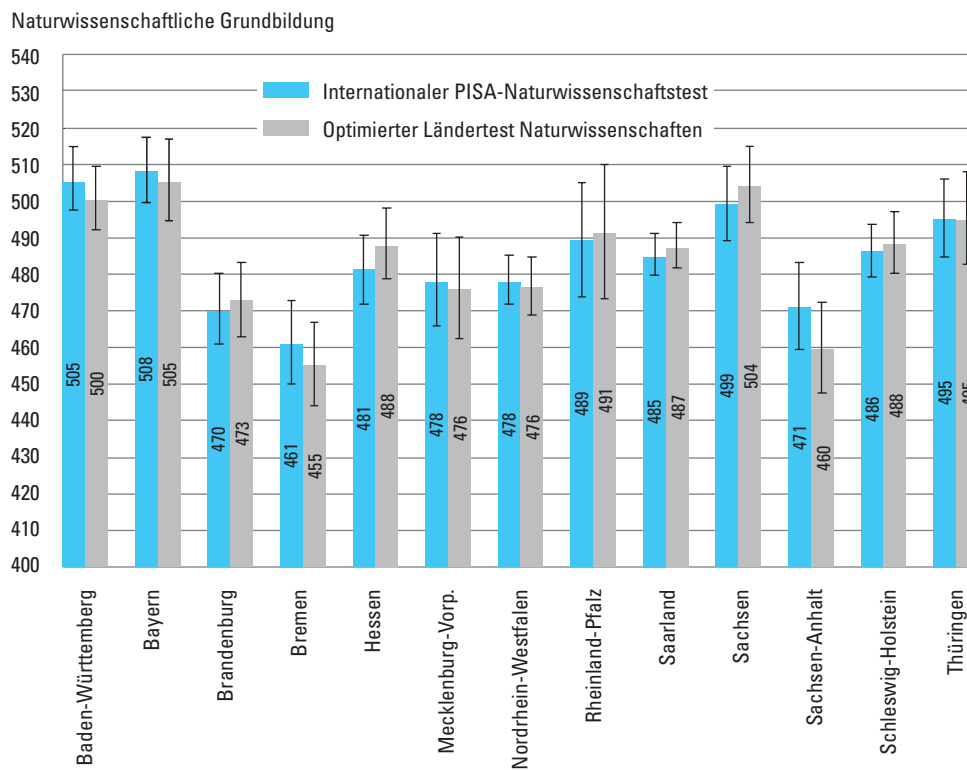


Abbildung 3: Naturwissenschaftliche Grundbildung: Mittelwerte im internationalen PISA-Naturwissenschaftstest und im optimierten Ländertest Naturwissenschaften

kogruppe bei Auswertung der lehrplanoptimierten Ländertests „größer“ wird. Wie auch schon bei den Mittelwerten bewegen sich diese „Veränderungen“ jedoch in einem Wertebereich, der auch per Zufall erwartet werden kann. Insgesamt kommt es also auch im Bereich Mathematik zu keinen statistisch absicherbaren Verschiebungen.

Abbildung 3 weist die Landesmittelwerte für beide Naturwissenschaftstestversionen aus. Niedersachsen ist in dieser und den folgenden Ergebnistabellen nicht enthalten, da in Niedersachsen lediglich drei Aufgaben von allen Experten konsistent als lehrplanvalide eingeschätzt wurden (vgl. Tab. 1). Diese geringe Aufgabenmenge reicht nicht aus, um zu aussagekräftigen Schätzungen des Leistungsniveaus der Schüler in den Naturwissenschaften zu gelangen.

Aufgrund überlappender Vertrauensintervalle ist der Mittelwertunterschied zwischen dem internationalen PISA-Naturwissenschaftstest und dem lehrplanoptimierten Ländertest in keinem der 13 Länder signifikant. Weder die „Verbesserung“ um 7 Punkte in Hessen bzw. 5 Punkte in Sachsen noch die „Verslechterung“ um 11 Punkte in Sachsen-Anhalt lässt sich zufallskritisch nachweisen. Auch die Veränderungen der Verteilungen der Schülerinnen und Schüler auf die Kompetenzstufen sind – ähnlich wie im Bereich Lesen und Mathematik – marginal.

Fazit

Insgesamt führt die Konzentration auf die lehrplanvaliden Aufgaben aus dem internationalen PISA-Test weder bei den 15-Jährigen noch bei den Neuntklässlern zu einer Verbesserung des Leistungsniveaus der Schülerinnen und Schülern in den drei untersuchten Kompetenzbereichen. Die Unterschiede liegen insgesamt im Bereich dessen, was aufgrund von Zufallsschwankungen erwartet werden kann. Diese Befunde bestärken die Annahme, dass die im PISA-Test gestellten Anforderungen nicht im Widerspruch zu dem stehen, was in deutschen Lehrplänen der Sekundarstufe I erwartet wird. Die Tatsache, dass die Auswahl von lehrplanvaliden Aufgaben nicht zu einer nachweisbaren Verbesserung des Leistungsniveaus führt, bedeutet im Umkehrschluss, dass die angeblich nicht lehrplanvaliden Aufgaben ähnlich gut gelöst werden.

Die PISA-Literacy-Tests überprüfen Basiskompetenzen, kein Spezialwissen in bestimmten Inhaltsbereichen. Auch wenn der Inhalt des Stoffes nicht explizit Gegenstand des Unterrichts ist und wenn die Bewältigung anhand des vorgegebenen Rahmenplans nicht notwendigerweise erwartet werden kann, zeigt sich bei den untersuchten Schülerpopulationen kein Leistungsunterschied.

Anmerkungen

- ¹ Pro Fach und Bildungsgang wurde jeweils ein Experte bzw. eine Expertin befragt.
- ² Vertiefende Analysen machen deutlich, dass die Einschätzungen der Lehrplanexperten mit Vorsicht zu behandeln sind. Die z.T. erwartungswidrigen Antwortmuster bei der Einschätzung der nationalen und internationalen Aufgaben zu Mathematik und Naturwissenschaften bei gleichzeitig z.T. gegenläufigen Mittelwertunterschieden in den Schülerleistungen machen deutlich, dass Aussagen der Lehrplanexperten im Rahmen der PISA-Befragung mit gewissen Interpretationsproblemen verbunden sind bzw. sein können. Vor dem Hintergrund der Tatsache, dass kein Training und keine Reliabilitätsüberprüfung vorgenommen wurde, ist ein solches Ergebnis allerdings auch nicht verwunderlich.
- ³ Das Vertrauensintervall um den Mittelwert kennzeichnet jenen Bereich, in dem der Mittelwert der gesamten Population 15-jähriger Schülerinnen und Schüler des jeweiligen Landes mit 95-prozentiger Wahrscheinlichkeit tatsächlich liegt. Das Vertrauensintervall (+/-2 Standardfehler) ist ein Maß für die Unsicherheit, die mit der stichprobenbasierten Schätzung von Populationskennwerten verbunden ist.