

10.05.2017

## Anhang zu

# TIMSS-Bericht 2015 und die verdrehte Zahlensprechweise im Deutschen

Quelle:

## TIMSS-Bericht 2015

Wendt H, Bos W, Selter C, Köller O, Schwippert K, Kasper D (2016) TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. Waxmann, Münster und New York

<https://www.waxmann.com/fileadmin/media/zusatztexte/3566Volltext.pdf>

### 1. Definition der mathematischen Inhaltsbereiche

Drei Inhaltsbereiche: Arithmetik, Geometrie/Messen, Umgang mit Daten

S. 89, Tab. 3.4:

Der Inhaltsbereich „Arithmetik“ umfasst die Teilgebiete „Natürliche Zahlen“, „Brüche und Dezimalzahlen“, „Terme sowie einfache Gleichungen und Beziehungen“  
der Inhaltsbereich „Geometrie/Messen“ umfasst die Teilgebiete „Punkte, Geraden und Winkel“ sowie „Zwei- und dreidimensionale Figuren“;  
der Inhaltsbereich „Umgang mit Daten“ umfasst „Daten lesen, interpretieren und darstellen“.

Genauere Definition der drei Inhaltsbereiche in Tab. 3.4 auf S. 89:

„Der Inhaltsbereich Arithmetik

Natürliche Zahlen

- über Wissen zu Stellenwerten verfügen, einschließlich des Lesens und Schreibens von Zahlen und des Darstellens von Zahlen mit Hilfe von Worten, Zeichnungen und Symbolen“

S. 94: „Tabelle 3.8 zeigt für die drei Studienzyklen (2007, 2011, 2015) die Verteilung der Aufgaben auf die jeweiligen mathematischen Inhaltsbereiche. Der Tabelle ist zu entnehmen, dass in TIMSS 2015 die Arithmetik mit 53 Prozent der Aufgaben den größten Teilbereich stellt. Hingegen entstammen 33 Prozent der Testaufgaben dem Bereich Geometrie/Messen, während 14 Prozent dem Inhaltsbereich Umgang mit Daten zuzuordnen sind.“

S. 15: „Etwa 80 bis 90 Prozent der mathematischen und naturwissenschaftlichen Aufgaben können für Deutschland als curricular valide eingestuft werden; das heißt, dass die untersuchten Kompetenzen entsprechend der curricularen Vorgaben bis zum Ende der Grundschulzeit vermittelt werden sollen.“

S. 96: „Da sich im Vergleich zu TIMSS 2011 bei der TCMA zu TIMSS 2015 keine nennenswerten Unterschiede bei den in Tabelle 3.10 aufgeführten ‚kritischen‘ Aufgabenmerkmalen curricular nicht valider Aufgaben ergeben haben, verweisen wir auf deren detaillierte Erläuterung im Berichtsband zu TIMSS 2011 (Selter et al., 2012, S. 85).

Tabelle 3.11 zeigt die Verteilung curricular nicht valider Testaufgaben auf die drei mathematischen Inhaltsbereiche. Aus den Daten geht hervor, dass mehr als ein Viertel der Aufgaben aus Arithmetik den curricularen Vorgaben nicht genügen.

Bei Umgang mit Daten liegt der Anteil bei 17 Prozent, bei Geometrie/Messen hingegen bei lediglich 9 Prozent. Der vergleichsweise hohe Anteil curricular nicht valider Testaufgaben in der Arithmetik erklärt sich dadurch, dass das zu Arithmetik zählende Themenfeld Brüche und Dezimalzahlen Aufgaben enthält, die nicht zum Standard des Grundschul-Curriculums der Bundesrepublik Deutschland gehören. Es sind dies vor allem Aufgaben mit ‚reinen‘, also nicht auf Größen bezogenen Brüchen und Dezimalzahlen.“

## **2. Teilnehmer und Durchführung**

S. 37: „Der vorliegende Bericht stellt zentrale Ergebnisse aus TIMSS 2015 für Schülerinnen und Schüler am Ende der vierten Jahrgangsstufe vor.“

S. 104: „Insgesamt nahmen 48 Staaten sowie 7 sogenannte Benchmark-Teilnehmer, also einzelne Regionen (etwa von Teilnehmerstaaten) wie Ontario (Kanada), an TIMSS 2015 teil. In diesem Abschnitt werden auf der Gesamtskala Mathematik (siehe Abbildung 3.5) Ergebnisse aus 53 Staaten und Regionen bei TIMSS 2015 berichtet, darunter 22 Teilnehmer aus der EU beziehungsweise 27 Teilnehmer und drei Benchmark-Teilnehmer aus der OECD (zu Details der Teilnahme siehe Kapitel 2 in diesem Band). Im Vergleich zu TIMSS 2011 liegt eine ähnliche Beteiligung vor, dort nahmen 57 Staaten und Regionen teil sowie drei Staaten mit ihren sechsten Klassenstufen. In TIMSS 2007 waren 43 Staaten und Regionen beteiligt.“

S. 45: „In Deutschland zeigt sich im Vergleich der Erhebungszyklen von TIMSS 2007 bis TIMSS 2015, dass je nach Bundesland und je nach Erhebungsjahr unterschiedliche Verpflichtungsgrade zur Teilnahme bestehen. Einheitlich definiert ist seit TIMSS 2007, das heißt ebenso in TIMSS 2011 und in TIMSS 2015, dass die Teilnahme an dem Leistungstest für alle Schülerinnen und Schüler verpflichtend ist.“

S. 45: Es nahmen alle 16 Bundesländer teil (Siehe Tabelle 2.3).

S. 60: „Die Haupterhebung von TIMSS 2015 fand in Deutschland in der Zeit vom 4. Mai bis zum 29. Mai 2015 statt.“

S. 62: „Die endgültige TIMSS-Stichprobe in Deutschland umfasst demnach 204 Schulen. Darunter befinden sich fünf nachgerückte Schulen. An allen 204 getesteten Schulen konnte die Leistungsmessung planmäßig durchgeführt werden. Von den 4 202 Schülerinnen und Schülern der Stichprobe nahmen 3 948 Schülerinnen und Schüler tatsächlich am TIMSS-Test teil.“

S. 63: Ein Überblick zur Zahl der teilnehmenden Schulen und Schülerinnen/Schüler in TIMSS 2007 bis 2015 gibt Tabelle 2.6.

### **3. Vergleichsgruppen**

S. 42: „Wie auch in den vorangegangenen Studienzyklen werden bei der Ergebnisdarstellung Vergleichsgruppen (VG) gebildet, denen jene Staaten und Staatengruppen angehören, die sich für einen Vergleich Deutschlands – etwa aufgrund eines ähnlichen kulturellen Hintergrunds oder einer vergleichbaren wirtschaftlichen Situation – eignen. Es handelt sich hierbei um die an TIMSS 2015 teilnehmenden Mitglieder der Europäischen Union (EU) und der Organisation for Economic Cooperation and Development (OECD). Mittelwerte werden somit nicht nur für einzelne Teilnehmer berichtet, sondern auch für die beiden Vergleichsgruppen VG EU und VG OECD.“

S. 42: Siehe Tabelle 2.2 zum Aufbau der VG<sub>EU</sub> (n=22) und VG<sub>OECD</sub> (n=27). VG<sub>EU</sub> umfasst 21 Staaten inkl. Deutschland und die flämische Region Belgiens (Ausnahme: in Norwegen nehmen 5. Klassen teil). VG<sub>OECD</sub> umfasst 27 Staaten, inkl. Deutschlands.

### **4. Ergebnisse**

S. 69: „TIMSS ist als Trendstudie konzipiert, das heißt, dass die Erhebungen in regelmäßigen Abständen wiederholt werden, um Veränderungen in den mittleren Leistungsniveaus der Teilnehmerstaaten über die Zeit hinweg darstellen zu können. Um die Schülerleistungen über die verschiedenen Zyklen vergleichen zu können, wird für die Darstellung der Ergebnisse dieselbe Skala zugrunde gelegt. Die Vergleichbarkeit der Ergebnisse wird dadurch sichergestellt, dass in jedem Zyklus Aufgaben aus den vorangegangenen Studienzyklen erneut eingesetzt werden.“

S. 68: „Zur Darstellung der Leistungswerte wurde für TIMSS 1995 und damit für alle bislang nachfolgenden Studienzyklen von TIMSS ein Mittelwert M (für arithmetisches Mittel) von 500 Punkten und eine Standardabweichung SD (für Standard Deviation) von 100 Punkten festgelegt (Martin & Kelly, 1998; Mullis et al., 1998). Die Wahl der Einheiten für diese Skala basiert ausschließlich auf Konventionen. Werte, die nahe beim Mittelwert liegen, kommen häufiger vor als Extremwerte. Oft ergibt sich eine Normalverteilung, wie sie in Abbildung 2.6 dargestellt ist.“

S.69: „Veränderungen über die Zyklen hinweg können auch im Jahr 2015 mit dem Mittelwert in TIMSS 1995 von  $M = 500$  verglichen werden. Dieser Referenzwert wird in TIMSS als Skalenmittelwert bezeichnet.“

Auszüge aus Tabelle (Abb. 3.5) auf S.107 und Tabelle (Abb. 3.8) auf S. 119

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

Abbildung 3.8: Testleistung der Schülerinnen und Schüler in Mathematik im internationalen Vergleich in den Inhaltsbereichen *Arithmetik*, *Geometrie/Messen* und *Umgang mit Daten*

Leistungsmittelwerte (M) und Standardfehler der Mittelwerte (SE):

M (SE) für	Arithmetik	Geometrie/Messen	Umgang mit Daten	Gesamt
VG OECD	527 (0.5)	530 (0.5)	530 (0.6)	528 (0.5)
VG EU	526 (0.6)	529 (0.6)	525 (0.8)	527 (0.6)
Deutschland	515 (2.1)	531 (2.5)	535 (2.6)	522 (2.0)

S. 107, Abbildung 3.5: der Gesamtskalenwert für Deutschland (522) liegt statistisch signifikant unter dem Gesamtskalenwert für die EU (527) und statistisch signifikant unter dem Gesamtskalenwert für die OECD (528).

S.118, 2. Absatz:

„Zu den Ergebnissen: Die Viertklässlerinnen und Viertklässler in Deutschland konnten im Bereich Arithmetik 515 Punkte, im Bereich Geometrie/Messen 531

Punkte und im Bereich Umgang mit Daten 535 Punkte erzielen. Die maximale Differenz zwischen den Leistungswerten in diesen drei Inhaltsbereichen beträgt 20 Punkte. Die Verteilung der Schülerleistungen über die drei Inhaltsbereiche ist somit inhomogen. Die Leistung im Inhaltsbereich Arithmetik fällt signifikant schwächer aus als der Gesamtmittelwert von Deutschland in Mathematik (522). Die Leistungswerte in den Inhaltsbereichen Geometrie/Messen sowie Umgang mit Daten sind hingegen signifikant höher als der Gesamtmittelwert. Somit bleibt festzuhalten, dass die Grundschülerinnen und Grundschüler in Deutschland ihre relativen Stärken in den Testaufgaben zum Umgang mit Daten und zu Geometrie/Messen aufweisen, während sie in Arithmetik deutlich schwächer abschneiden.“

S. 118 und 120:

„Analysiert man die Testergebnisse separat nach Inhaltsbereichen, so ergibt sich, dass im Bereich Arithmetik 24 Teilnehmer, im Bereich Geometrie/Messen 14 und im Bereich Umgang mit Daten 11 Teilnehmer signifikant besser als Deutschland abschneiden. Demgegenüber erzielen in Arithmetik 8 Teilnehmer, in Geometrie/Messen 12 und in Umgang mit Daten 15 Teilnehmer signifikant niedrigere Leistungsmittelwerte als Deutschland.“

S. 120, 2. Absatz, siehe insbesondere den letzten Satz:

„Im Vergleich der Leistungsmittelwerte Deutschlands in den drei Inhaltsbereichen mit denen der VG EU und VG OECD ergibt sich folgendes Bild: Kein signifikanter Unterschied liegt für den Bereich *Geometrie/Messen* vor, wogegen der Leistungswert für *Umgang mit Daten* signifikant über dem Vergleichswert der teilnehmenden EU-Staaten liegt. Im Vergleich zum Leistungsmittelwert der OECD-Staaten liegt in diesem Bereich jedoch kein Unterschied vor. Im Inhaltsbereich *Arithmetik* ist der Leistungsmittelwert für Deutschland hingegen signifikant unter denen der Vergleichsgruppen verortet.“

S. 134: „Somit zeigen die Schülerinnen und Schüler in Deutschland ihre relativen Schwächen in dem anteilig mit 53 Prozent der Testaufgaben am häufigsten vertretenen Bereich Arithmetik.“

S. 268/270: „Für den Bereich Arithmetik lassen sich für 25 der 40 dargestellten Teilnehmerstaaten und Regionen signifikante Leistungsunterschiede feststellen. Mit

Ausnahme von Finnland, wo sich Leistungsvorsprünge der Mädchen beobachten lassen, fallen diese, sofern signifikant, zugunsten der Jungen aus. Die mittlere Differenz für die Vergleichsgruppe EU liegt bei 8 Punkten, die mittlere Differenz für die Vergleichsgruppe OECD bei 7 Punkten. In Deutschland erreichen Mädchen im Inhaltsbereich Arithmetik 510 Punkte, während der Mittelwert für Jungen in diesem Bereich bei 520 Punkten liegt. Die Differenz von 9 Punkten zugunsten der Jungen, die sich unter Berücksichtigung der Nachkommastellen ergibt, ist signifikant.“

S. 108, Auswertung allein mit curricular validen Aufgaben: „Im Folgenden wird überprüft, ob sich daraus Auswirkungen auf die Testergebnisse Deutschlands im internationalen Vergleich ergeben haben. Hierzu wird die Gesamtskala Mathematik in einer abgewandelten Form berechnet. Dabei werden die Leistungsmittelwerte aller teilnehmenden Staaten und Regionen erneut berechnet, wobei ausschließlich die 137 für Deutschland curricular validen Aufgaben ausgewertet werden.“

S. 109, Ergebnis dieser Auswertung: „Angesichts der beschriebenen Daten lässt sich festhalten, dass die Testergebnisse Deutschlands auf der TIMSS-Leistungsskala nicht auf das Testdesign zurückzuführen sind. Modifikationen würden vermutlich allenfalls marginale Unterschiede in der internationalen Rangreihe herbeiführen.“

## **5. Vergleich mit TIMSS 2011 und 2007**

Auszüge aus Tabelle (Abb. 3.6) auf S.111 und Tabelle (Abb. 3.9) auf S. 21

Abbildung 3.6: Vergleich der Testleistungen zwischen TIMSS 2007, 2011 und 2015 – Gesamtskala Mathematik

Abbildung 3.9: Vergleich der Testleistungen in Mathematik zwischen TIMSS 2007, 2011 und 2015 in den Inhaltsbereichen Arithmetik, Geometrie/Messen und Umgang mit Daten

Deutschland: Leistungsmittelwerte (M) und Standardfehler der Mittelwerte (SE):

M (SE) für	Arithmetik	Geometrie/Messen	Umgang mit Daten	Gesamt
TIMSS2007	524 (2.2)	520 (2.3)	515 (2.1)	525 (2.3)
TIMSS2011	527 (2.4)	536 (2.6)	531 (2.5)	528 (2.6)
TIMSS2015	515 (2.1)	531 (2.5)	535 (2.6)	522 (2.0)

S. 120:

„Aus den Abbildungen 3.9 und 3.10 geht für Deutschland hervor, dass im Bereich Arithmetik in TIMSS 2015 signifikant schwächere Testleistungen als in TIMSS 2007 erzielt wurden, wogegen in Geometrie/Messen sowie in Umgang mit Daten keine statistisch bedeutsamen Unterschiede der Punktwerte vorliegen.“

S. 120:

„Der Vergleich der Daten der Studienzyklen von TIMSS 2015 und 2011 zeigt für Deutschland lediglich im Bereich Umgang mit Daten signifikant schwächere Leistungswerte in TIMSS 2015. Für die Bereiche Arithmetik und Geometrie/Messen sind hingegen keine statistisch signifikanten Unterschiede festzuhalten.“

## **6. Österreich**

Österreich hat nicht an TIMSS 2015 teilgenommen, jedoch an TIMSS 2007 und 2011.

Zur Übersicht siehe <https://www.bmb.gv.at/ministerium/vp/2012/20121211a.html>:

„TIMSS (Mathematik): Leichte Verbesserung gegenüber TIMSS 2007 (508 gegenüber 505). Internationaler Vergleich: 23. Rang von 50 TIMSS-Staaten; 14. Rang von 21 EU-TIMSS-Staaten. Die österreichischen SchülerInnen erreichen im Jahr 2011 im Mittel 508 Punkte und liegen damit signifikant über dem Schnitt aller TIMSS-Staaten (491). Innerhalb der 50 Teilnehmerländer bedeutet dies Rang 23. Die 21 an TIMSS teilnehmenden EU-Länder erreichen einen Mittelwert von 519 Punkten. Der österreichische Mittelwert liegt im EU-Vergleich auf Rang 14 von 21. Der Abstand zu den besten asiatischen Ländern Singapur (606), Korea (605) und Hongkong (602) beträgt etwa 100 Punkte. Deutschland (528) liegt 20 Punkte vor Österreich.“

TIMSS 2011-Bericht zu Österreich: <https://www.bifie.at/buch/1742>

Auszüge aus Tabelle 2.1 auf S. 24 und Tabelle 2.5 auf S. 32:

Tab 2.1: Mathematik-Gesamtskala: Mittelwerte und Standardabweichungen der Teilnehmerländer (TIMSS 2011)

Tab 2.5: Mittelwerte der drei Inhaltsbereiche in Mathematik (TIMSS 2011)

Deutschland (D) und Österreich (Ö): Leistungsmittelwerte (M)

M für	Arithmetik	Geometrie/Messen	Umgang mit Daten	Gesamt
D-2011	527	536	531	528
Ö-2011	506	512	515	508

Wie in Deutschland zeigt der Bereich „Arithmetik“ also die niedrigsten Leistungspunkte.

Zahlen, S. 34:

„Ein Verständnis für Zahlen und das Arbeiten mit Zahlen bilden die Grundlage für Mathematik im Primarbereich. Dementsprechend wird dieser Inhaltsbereich im Vergleich zu den anderen stärker gewichtet: die Hälfte der Testzeit in Mathematik ist dem Thema Zahlen gewidmet. Dazu zählen der Umgang mit ganzen Zahlen, Brüchen und Dezimalzahlen, das Lösen einfacher Gleichungen und das Fortsetzen von Zahlenreihen. Schüler/innen aus Singapur, Korea und Hongkong sind mit Mittelwerten von über 600 Punkten bei diesen Aufgaben am besten. Der Schnitt der 14 Vergleichsländer beträgt 524 Punkte. Österreich liegt mit einem Mittelwert von 506 Punkten signifikant unter diesem Referenzwert. ... Im Vergleich zu TIMSS 2007 ist in Österreich keine Änderung zu verzeichnen.“

## **7. Dänemark und Niederlande**

TIMSS2015:

Auszüge aus Tabelle (Abb. 3.5) auf S.107 und Tabelle (Abb. 3.8) auf S. 119

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

Abbildung 3.8: Testleistung der Schülerinnen und Schüler in Mathematik im internationalen Vergleich in den Inhaltsbereichen *Arithmetik*, *Geometrie/Messen* und *Umgang mit Daten*

Leistungsmittelwerte (M) und Standardfehler der Mittelwerte (SE):

M (SE) für	Arithmetik	Geometrie/Messen	Umgang mit Daten	Gesamt
Dänemark	535 (2.7)	555 (3.2)	526 (3.5)	539 (2.7)
Niederlande	531 (2.2)	522 (1.9)	539 (3.4)	530 (1.7)
Deutschland	515 (2.1)	531 (2.5)	535 (2.6)	522 (2.0)