



ZENTRUM FÜR INTERNATIONALE  
BILDUNGSVERGLEICHSTUDIEN

Zentrum für internationale  
Bildungsvergleichsstudien (ZIB) e.V.  
An-Institut der Technischen  
Universität München

Prof. Dr. Manfred Prenzel  
Vorstandsvorsitzender

Postanschrift:  
Arcisstraße 21  
80333 München

Tel: +49.89.289.25242  
E-Mail: manfred.prenzel@tum.de

## **„Wir hatten leider Pech mit der Zusammensetzung unserer Schülerschaft!“**

### **PISA und ein Preis für Unstatistik**

*Ein Kommentar zur Pressemitteilung des Rheinisch-Westfälischen Instituts für  
Wirtschaftsforschung vom 17.12.2013, vgl. <http://www.unstatistik.de>*

Das Programme for International Student Assessment (PISA) vergleicht regelmäßig die Leistungen, die von fünfzehnjährigen Schülerinnen und Schülern in den Bereichen Mathematik, Naturwissenschaften und Lesen erreicht werden. Diese internationale Vergleichsstudie ist von der OECD als Vergleich von Staaten angelegt, der Aufschluss über die Stärken und Schwächen der jeweiligen Bildungssysteme geben soll.

Im ersten Vergleich (PISA 2000) lag Deutschland in allen Leistungsbereichen signifikant unter dem Durchschnitt der OECD-Staaten (Lesen 484 Punkte, Mathematik 490 Punkte, Naturwissenschaften 487 Punkte). Der letzte Vergleich (PISA 2012) zeigte für Deutschland in allen drei Bereichen Ergebnisse, die statistisch bedeutsam den OECD-Mittelwert übertrafen (Lesen 508 Punkte, Mathematik 514 Punkte, Naturwissenschaften 524 Punkte).

Nachdem die Bildungspolitik in Deutschland bei PISA 2000 heftig für die schwachen Ergebnisse gescholten und zum Handeln aufgefordert worden war, drückten die politischen Stellungnahmen bei PISA 2012 nun Freude über die deutlich verbesserten Leistungen der Schülerinnen und Schüler in Deutschland aus.

In einer Pressemitteilung des Rheinisch-Westfälischen Instituts für Wirtschaftsforschung aus Essen vom 17. Dezember 2013 wird vor einer Überinterpretation der Ergebnisse gewarnt und deshalb der Rückschluss auf Lernfortschritte in Deutschland als „Unstatistik des Monats“ gewürdigt. Konkret angesprochen wird der Leistungszuwachs von 11 Punkten zwischen 2003 und 2012 – den Erhebungsrounden, in denen Mathematik die Hauptdomäne war.

In der Pressemitteilung wird erstens behauptet, es sei nur „eine kleine Stichprobe deutscher Schüler getestet worden, bei einer anderen Stichprobe wäre etwas anderes herausgekommen“. Diese Aussage ist falsch: Bei PISA werden in allen Staaten repräsentative Zufallsstichproben gezogen. Von der Größe sind sie so angelegt, dass mit einer definierten, sehr geringen Fehlermarge von der untersuchten Stichprobe auf die Grundgesamtheit aller Fünfzehnjährigen geschlossen werden kann. Ebenso können Vergleiche zwischen den Staaten und Veränderungen zwischen den PISA-Messungen statistisch geprüft werden. Dies ist das in der Statistik übliche Verfahren: Bei einer anderen repräsentativen Zufallsstichprobe wären mit höchster Wahrscheinlichkeit vergleichbare Ergebnisse gemessen worden. Die Fehlermargen werden in allen PISA-Berichten dargestellt. Unterschiede werden dort nur interpretiert, wenn die Irrtumswahrscheinlichkeit sehr gering ist.

Das zweite Argument, das insbesondere die ‚unstatistischen‘ Schlussfolgerungen belegen soll, betrifft die Zusammensetzung der Stichprobe. Es sei außer Acht gelassen worden, „dass sich über die Jahre hinweg die Zusammensetzung der Stichprobe verändert hat.“ Bezogen auf den Vergleich der Messzeitpunkte 2003 und 2012 wird behauptet: Die veränderten Zusammensetzungen der Stichproben in Deutschland nach Alter, Geschlecht, sozioökonomischen Hintergrund und Migrationshintergrund könnten für die Verbesserungen in den Mathematikergebnissen verantwortlich sein.

Betrachten wir die Zusammensetzungen der Stichproben in Deutschland zu den beiden Erhebungszeitpunkten, dann ergibt sich das folgende Bild: Der Altersdurchschnitt betrug in PISA 2003 15,78 Jahre, in PISA 2012 15,83 Jahre. An PISA 2003 nahmen 49,66 Prozent Mädchen teil, bei PISA 2012 waren es 49,13 Prozent. Der Index für den sozioökonomischen und soziokulturellen Status (ESCS) differiert zwischen PISA 2003 und 2012 um 0,03 Punkte. Während in PISA 2003 20,9 Prozent der Fünfzehnjährigen einen Zuwanderungshintergrund aufwiesen, stieg dieser Anteil in PISA 2012 auf 25,8 Prozent. Der Anteil von Jugendlichen mit Zuwanderungshintergrund ist in Deutschland also seit 2003 gestiegen. Die Kennwerte für den sozioökonomischen Hintergrund haben sich zwischen 2003 und 2012 nicht signifikant verändert.

Die Pressemitteilung rekurriert in diesem Zusammenhang auf eine Abbildung des OECD-Berichts, in der die Verbesserung der Mathematikleistungen in Deutschland auf die Veränderungen in den entsprechenden Zusammensetzungen der Stichproben zurückgeführt wird. Nun zeigen die Befunde für PISA 2012 aber, dass sich in Deutschland besonders die Schülerinnen und Schüler mit Zuwanderungshintergrund (gestiegener Anteil) und niedrigen Sozialschichtkennwerten in ihren Mathematikleistungen verbessert haben! Und wo haben sie das gelernt? Mathematik lernt man nicht auf der Straße, sondern in der Schule.

Obwohl die Unstatistik-Jury gegenüber der PISA-Studie so kritisch ist, nimmt sie Modellrechnungen, welche die OECD selbst als spekulativ bezeichnet, wörtlich. In einem Gedankenexperiment hatten die OECD-Mitarbeiter errechnet, welche Kompetenzwerte für Deutschland gemessen worden wären, wenn die heutige Zusammensetzung der Schülerschaft schon bei den früheren PISA-Messungen gegolten hätte. Dieses Gedankenexperiment führt im aktuellen OECD-Bericht dazu, dass für den Zeitraum zwischen 2000 und 2009 ein fiktiver Rückgang der Lesekompetenz um einen Punkt ermittelt wird. Vor drei Jahren hatte die OECD dasselbe Gedankenexperiment schon einmal durchgeführt; damals wurde für denselben Zeitraum ein fiktiver Gewinn von 17 Punkten ermittelt. Solche Widersprüche zeigen, dass die Modellrechnungen der OECD keinen Erklärungswert haben – nicht zuletzt, weil sie die Leistungsentwicklung innerhalb der sozialen Gruppen nicht berücksichtigen.

In Deutschland werden seit PISA 2000 regelmäßig nationale PISA-Berichte vorgelegt, verfasst von Bildungsforschern, die sehr genau darauf achten, zwischen Spekulation und Fakten zu trennen. Die Kritiker aus Essen hätten gut daran getan, sich diese Berichte anzusehen, anstatt einem Gedankenexperiment auf den Leim zu gehen, das wissenschaftlich nicht aussagefähig ist.

Der Psychologe Gerd Gigerenzer, der Statistiker Walter Krämer und der RWI-Vizepräsident Thomas Bauer verleihen die „Unstatistik des Monats“. Welchen Preis hätten sie anlässlich PISA 2000 verliehen, wenn die Bildungspolitik die Befunde so interpretiert hätte: „Wir hatten leider Pech mit unserer Schülerzusammensetzung. Aber das ändert sich ja vielleicht wieder“. Die „Unstatistik des Monats“ geht mit Kompliment an die Jury zurück.

Für das Zentrum für internationale Vergleichsstudien (ZIB) e.V.

München, Frankfurt, Kiel, den 19. Dezember 2013

Prof. Dr. Manfred Prenzel

ZIB/

TUM School of Education

München

Prof. Dr. Eckhard Klieme

Deutsches Institut für  
Internationale Pädagogische  
Forschung (DIPF)

Frankfurt a. M.

Prof. Dr. Olaf Köller

Leibniz-Institut für die  
Pädagogik der Naturwissen-  
schaften und Mathematik (IPN)

Kiel