

# **Schiefer als der schiefe Turm**

## **Die PISA-Studie weist schwerwiegende Mängel auf**

### **Von Rainer Bölling**

Anfang Dezember 2019 kommt eine neue Ausgabe der PISA-Studie heraus. Wird das deutsche Schulwesen im internationalen Vergleich wieder als leistungsschwach und sozial ungerecht dastehen? Schon im Januar 2018 wies der Präsident des Deutschen Lehrerverbandes, Heinz-Peter Meidinger, warnend darauf hin, dass seit 2015 durch die Zuwanderung von Flüchtlingen mit über 200.000 schulpflichtigen Kindern die Zahl der Familien zunehme, in denen zu Hause kein Deutsch gesprochen werde.<sup>1</sup> Sprachkenntnisse aber sind für das Abschneiden im PISA-Test von zentraler Bedeutung. Doch auch methodische Probleme der Studie können sich stark auf das nationale PISA-Ergebnis auswirken.

Das *Program for International Student Assessment* der OECD untersucht seit der Jahrtausendwende alltags- und berufsrelevante Kompetenzen fünfzehnjähriger Schülerinnen und Schüler in den Bereichen Leseverständnis, Mathematik und Naturwissenschaften. Dabei arbeitet es mit Stichproben, die manchmal weniger als ein Prozent der Zielgruppe umfassen.<sup>2</sup> Das ist in der Statistik üblich, birgt aber gleichwohl seine Tücken. Denn wenn eine Stichprobe nicht korrekt gezogen wird, kommt es zu falschen Ergebnissen. So etwa in Österreich, wo der PISA-Test 2003 deutlich schlechter ausfiel als in der ersten Runde drei Jahre zuvor. Nachdem für diesen vermeintlichen Niedergang die neue Regierung verantwortlich gemacht worden war, förderte eine Überprüfung durch versierte Statistiker Ungereimtheiten bei der Stichprobenziehung und Datenauswertung des Jahres 2000 zutage. Daher mussten die früheren Werte nach unten korrigiert werden. Auch die Stichprobe der USA von 2009 wurde von amerikanischen Forschern als unzuverlässig kritisiert. Nach ihren korrigierten Punktzahlen hätten die USA im Ranking um zehn und mehr Plätze besser abgeschnitten.<sup>3</sup> Derartige Fehler sind auch für andere Länder und Testrunden nicht auszuschließen.

Ein weiteres Problem besteht darin, dass nicht alle für den Test vorgesehenen Schüler an der Erhebung teilnehmen, wobei die Ausschlussgründe variieren. Zur Begrenzung des dadurch entstehenden Schätzfehlers gilt bei PISA eine Mindestbeteiligungsquote von 80 Prozent. Als Großbritannien diese Quote 2003 knapp verfehlte, wurde es in internationalen Vergleichen nicht berücksichtigt.<sup>4</sup> Doch als Schanghai 2012 ebenfalls diese Hürde riss, setzte sich die OECD über ihre eigenen Regeln hinweg und präsentierte Schanghai als Sieger und Vorbild für die Welt.<sup>5</sup> Das war umso fragwürdiger, als dort die Kinder von Millionen Wanderarbeitern

schon vorab außen vor blieben, weil sie gar nicht die höheren Schulen besuchen durften, an denen der PISA-Test stattfand.<sup>6</sup>

Ungeachtet dessen erklärte PISA-Koordinator Schleicher im Dezember 2013, in Shanghai gebe es in Mathematik keinen leistungsschwachen Schüler, und folgerte daraus: „Mathematik ist etwas, das jeder Schüler aus jeder Schicht lernen kann.“<sup>7</sup> Hätte er die einschlägigen PISA-Daten zur Kenntnis genommen, dann wäre ihm vielleicht aufgefallen, dass selbst in dieser unvollständigen Stichprobe 3,8 Prozent der Getesteten die Kompetenzstufe 2 verfehlt hatten, also als sogenannte Risikoschüler gelten.<sup>8</sup> Schleichers Aussage zeigt, dass seine bildungspolitischen Wunschvorstellungen und Ratschläge oft nur vorgeblich auf den von der OECD publizierten Daten beruhen.

Bei PISA 2015 erscheint Shanghai nicht mehr allein, sondern zusammen mit Peking, Jiangsu und Guangdong als Teil einer neuen Einheit (P-S-J-G) mit 223 Millionen Einwohnern, die China repräsentiert. Sie wird auf Platz 10 des offiziellen Rankings geführt, obwohl ihre Erfassungsquote nur 64 Prozent beträgt. Dennoch lag der Anteil der Risikoschüler in Mathematik hier gut viermal so hoch wie 2012 in Shanghai.<sup>9</sup> Noch krasser ist der Fall Vietnam, das 2015 im Ranking wegen seiner 525 Punkte in Naturwissenschaften auf Platz 8 steht, obwohl es mit 48,5 Prozent die Mindestbeteiligungsquote von 80 Prozent bei weitem verfehlt hat.

Die OECD räumt ein: „Wenn man annimmt, dass aus der PISA-Stichprobe ausgeschlossene Schülerinnen und Schüler wahrscheinlich schlechtere Ergebnisse erzielt hätten als die in den Stichproben vertretenen Schülerinnen und Schüler, dürften Länder mit niedrigen Erfassungsraten in Vergleichen favorisiert sein.“<sup>10</sup> Eben deshalb aber hätten beide Länder vom internationalen Vergleich ausgeschlossen werden müssen. Man fragt sich, warum die OECD bei bestimmten asiatischen Ländern selbst ihre eigenen Regeln über Bord wirft.

Kritik an diesem Vorgehen tut die OECD auf ihrer Website als Mythen ab: „Alle PISA-Proben sind repräsentativ für die 15-Jährigen, die an Schulen angemeldet sind“, heißt es dort. „Sie werden auf einer wissenschaftlich fundierten Basis ausgewählt, um sicherzustellen, dass alle Schülerinnen und Schüler gleichmäßig repräsentiert werden. Dabei werden 89 Prozent der 15-Jährigen von PISA abgedeckt.“<sup>11</sup> Doch das ist wohlgermerkt der OECD-Durchschnitt, der eben klare Erfassungsdefizite einzelner Teilnehmerländer verschleiert. Übrigens erreichen auch Länder wie Großbritannien, die USA oder Kanada nur Erfassungsquoten unter 85 Prozent, während Deutschland mit 96 Prozent in der Spitzengruppe liegt. Solche Unterschiede wirken sich durchaus auf die nationalen Durchschnittswerte aus.

Wenn die OECD Meldungen über die Defizite schwächerer Schülerinnen und Schüler veranschaulichen will, rechnet sie gern PISA-Punkte in Schuljahre um. So war im Oktober 2018 in der Studie „Equity in Education“ zu lesen, der Bildungsrückstand sozial benachteiligter Schüler betrage im OECD-Durchschnitt 88 Punkte beziehungsweise etwa drei Schuljahre.<sup>12</sup> Das wirkt noch dramatischer als zuvor. 2009 und 2012 galten bei PISA ca. 125 Punkte als Äquivalent für drei Schuljahre, 2015 waren es noch 100.<sup>13</sup> Wie unsinnig all diese Umrechnungen sind, wird dann deutlich, wenn man sie auf die Unterschiede *innerhalb* der Länder anwendet. Zwischen den leistungsstärksten und den leistungsschwächsten Schülern eines OECD-Landes (obere und untere 5 Prozent) lagen 2015 im Durchschnitt der drei Kompetenzbereiche fast überall 300 und mehr Punkte.<sup>14</sup> Nimmt man das mittlere Maß von 2015 als Umrechnungsfaktor, so würde das einen Bildungsrückstand der Leistungsschwächsten von neun bis zehn Schuljahren bedeuten. Demnach wären beispielsweise die schwächsten finnischen Fünfzehnjährigen nach acht Jahren Schulbesuch schon neun Jahre in Rückstand geraten. Offensichtlich dient dieses unbrauchbare Umrechnungsverfahren nur der Skandalisierung des Unspektakulären und gehört daher auf den Müllhaufen der empirischen Bildungsforschung.

Unabhängig von der Größe des Umrechnungsfehlers zeigt diese Betrachtung, dass es auch PISA-Musterländern nicht annähernd gelingt, „kein Kind zurückzulassen“ und alle zu guten Leistungen zu führen. Besonders groß ist der Abstand zwischen den besten und den schwächsten Schülern in den chinesischen Provinzen, aber auch in Singapur. Dennoch behauptet Schleicher, seine zahlreichen Reisen nach Asien hätten ihm gezeigt, „dass man praktisch alle Schüler zu wirklich guten Leistungen motivieren kann“.<sup>15</sup> Die PISA-Daten sprechen für das Gegenteil.

Im Übrigen liegt die Gruppe der besten Schüler (obere 5 Prozent) im internationalen Vergleich auf einem ähnlichen Niveau. Abgesehen von asiatischen Ländern mit teils fragwürdigen oder sehr speziellen Voraussetzungen gibt es eine Spitzengruppe mit durchschnittlichen Punktwerten zwischen 667 und 652. Dazu gehören (in dieser Reihenfolge) Kanada, Finnland, Deutschland (661), die Niederlande, Schweiz, Belgien und Frankreich. Wenn die Durchschnittswerte dieser Ländergruppe stärker voneinander abweichen, so liegt das nicht zuletzt am unterschiedlichen Anteil geringer qualifizierter Zuwanderer in der jeweiligen Stichprobe. Es wäre daher geboten, die sprachlichen, kulturellen und sozialen Voraussetzungen von Migrantenkindern genauer in den Blick zu nehmen.

Den PISA-Machern ist durchaus bekannt, dass Länder wie Kanada, Australien und Neuseeland eine selektive Einwanderungspolitik betreiben, so dass es sich bei den

Zuwanderern überwiegend um Hochqualifizierte mit Berufserfahrung handelt, die zudem die Landessprache beherrschen. Dagegen weisen etwa Deutschland, Österreich, Belgien und Frankreich einen hohen Anteil an dauerhaft ansässigen geringqualifizierten Zuwanderern auf. Zwar heißt es bei PISA, dass die im Ländervergleich bestehenden Unterschiede „auch unter dem Aspekt der Selektivität der Zuwanderungspolitik der Aufnahmeländer sowie der relativen kulturellen und sprachlichen Verwandtschaft der Herkunfts- und Aufnahmeländer betrachtet werden“ müssen. Doch eine solche Differenzierung findet tatsächlich nicht statt. Man erfährt lediglich, „dass der durchschnittliche Leistungsabstand zwischen Schülerinnen und Schülern mit und ohne Migrationshintergrund und gleichem sozioökonomischem Profil 31 Punkte beträgt“.<sup>16</sup> Wie sich Anteil und Qualifikationsniveau von Migrantenkindern konkret auf die nationalen PISA-Ergebnisse auswirken, wird jedoch nicht thematisiert.

Offensichtlich handelt es sich hier um ein Tabu, wie die Reaktion der deutschen PISA-Koordinatorin Kristina Reiss in einem Interview mit der ZEIT im Dezember 2016 zeigt. Als sie mit der These konfrontiert wurde, ohne Einwandererkinder wäre Deutschland auf dem Niveau von Europas Schulprimus Finnland, entgegnete sie: „Diese Argumentation ist gefährlich. Zum einen stammt ein gutes Viertel der Risikoschüler aus deutschstämmigen Familien. Zum anderen schaffen es Länder wie Kanada durchaus, Kinder aus Einwandererfamilien zu guten Leistungen zu führen.“<sup>17</sup> Doch Kanada verdankt seine guten PISA-Ergebnisse ja nicht zuletzt seiner selektiven Einwanderungspolitik. Dass dies hartnäckig unterschlagen wird, verstößt eklatant gegen wissenschaftliche Regeln.

Wenn hingegen die PISA-Ergebnisse Deutschlands Fortschritte erkennen lassen, ist die OECD schnell dabei, diese für sich zu reklamieren. So führte sie das gute Abschneiden 2012 auch auf den von ihr befürworteten Ausbau der frühkindlichen Förderung zurück. Ihr stellvertretender Generalsekretär behauptete sogar, man könne „in den PISA-Zahlen ganz deutlich messen, dass bereits ein Jahr frühkindliche Bildung nachher die Ergebnisse im PISA-Test deutlich verbessert“.<sup>18</sup> Doch die PISA-Zahlen geben nichts dergleichen her, denn die 2012 getesteten Fünfzehnjährigen wurden ja spätestens 2003 eingeschult, also vor dem Ausbau der frühkindlichen Bildung in Deutschland. Sie kann also nicht der Grund für die Verbesserung sein.

Die hier aufgezeigten methodischen Probleme und irreführenden Interpretationen der PISA-Studie sollten Grund genug sein, der Präsentation der neuen Ausgabe nicht blind zu vertrauen, sondern ihr kritisch zu begegnen. Auf eine konstruktive Diskussion mit den PISA-Machern zu hoffen erscheint allerdings unrealistisch, müssten diese dann doch ihren latenten Unfehlbarkeitsanspruch aufgeben.

- <sup>1</sup> <https://www.news4teachers.de/2018/01/pisa-zu-bildungsgerechtigkeit-vbe-und-lehrerverband-registrieren-fortschritte-aber> (Zugriff: 30.11.2019).
- <sup>2</sup> Beispielsweise wurden 2015 in Deutschland 6522 Schülerinnen und Schüler repräsentativ für 774.149 Fünfzehnjährige getestet. OECD: PISA 2015 Ergebnisse (Band I): Exzellenz und Chancengerechtigkeit in der Bildung. 2016, S. 315.
- <sup>3</sup> Rainer Bölling, Zur Fragwürdigkeit des PISA-Rankings, 2017. <https://bildung-wissen.eu/fachbeitraege/zur-fragwuerdigkeit-des-pisa-rankings.html> (Zugriff: 30.11.2019).
- <sup>4</sup> OECD: Lernen für die Welt von morgen. Erste Ergebnisse von PISA 2003, 2004, S. 371 u. 375-377.
- <sup>5</sup> OECD, PISA 2012 Ergebnisse: Was Schülerinnen und Schüler wissen und können. Band I (überarbeitete Ausgabe, Februar 2014): Schülerleistungen in Lesekompetenz, Mathematik und Naturwissenschaften, 2014, S. 290 f.
- <sup>6</sup> Tom Loveless, Attention OECD-PISA: Your-Silence-on-China-is-Wrong, 12.12.2013 <https://www.educationnext.org/attention-oecd-pisa-your-silence-on-china-is-wrong> (Zugriff: 30.11.2019).; Valerie Strauss, Did Shanghai cheat on PISA? 15.12.2013 <http://www.washingtonpost.com/blogs/answer-sheet/wp/2013/12/15/did-shanghai-cheat-on-pisa> (letzter Zugriff 2017).
- <sup>7</sup> Sabine Damaschke/Janine Albrecht, Ausbruch aus der Schockstarre, 03.12.2013 <https://www.dw.com/de/ausbruch-aus-der-schockstarre/a-17267586> (Zugriff: 30.11.2019)
- <sup>8</sup> PISA 2012 (Fn. 5), S. 20 u. 322.
- <sup>9</sup> PISA 2015 (Fn. 2), S. 315 u. 414.
- <sup>10</sup> Ebd., S. 229.
- <sup>11</sup> <https://www.oecd.org/berlin/themen/pisa-studie> (Zugriff: 30.11.2019).
- <sup>12</sup> OECD: Equity in Education. Breaking Down Barriers to Social Mobility. Paris 2018, S. 13.
- <sup>13</sup> OECD, PISA 2009 Ergebnisse: Was Schülerinnen und Schüler wissen und können. Schülerleistungen in Lesekompetenz, Mathematik und Naturwissenschaften, Band I, 2010, S. 13 u. ö.; PISA 2012 (Fn. 5), S. 4 u. 272; PISA 2015 (Fn. 2), S. 41 u. 162.
- <sup>14</sup> Berechnet nach PISA 2015 (Fn. 2), S. 351, 404 u, 417.
- <sup>15</sup> Wie Fn 7. - „Dass Schülerinnen und Schüler unabhängig von ihrem sozioökonomischen und soziokulturellen Status sowie unabhängig von Hautfarbe und Geschlecht gleiche Lernstände erreichen“, galt auch dem geistesverwandten Chancenspiegel der Bertelsmann-Stiftung als Inbegriff von Bildungsgerechtigkeit (Bertelsmann Stiftung: Chancenspiegel. Zur Chancengerechtigkeit und Leistungsfähigkeit der deutschen Schulsysteme, Gütersloh 2012, S. 24).
- <sup>16</sup> PISA 2015 (Fn. 2), S. 264 f.
- <sup>17</sup> Thomas Kerstan/Martin Spiewack, Wir klagen auf hohem Niveau. <https://www.zeit.de/2016/51/pisa-studi-bildungspolitik-kristina-reiss-olaf-koeller/komplettansicht> (Zugriff: 30.11.2019).
- <sup>18</sup> Christiane Habermalz, Große Fortschritte bei der frühkindlichen Bildung, 24.11.2015 [https://www.deutschlandfunkkultur.de/oecd-bildungsstudie-grosse-fortschritte-bei-der.2165.de.html?dram:article\\_id=337814](https://www.deutschlandfunkkultur.de/oecd-bildungsstudie-grosse-fortschritte-bei-der.2165.de.html?dram:article_id=337814) (Zugriff: 30.11.2019).