



Schiefer als der schiefe Turm?

An der PISA-Studie gibt es viel Kritik,
aber auch manche Vorurteile.

Der Kieler Bildungsforscher Olaf Köller hält
den internationalen Vergleich für zentral.

Der PISA-Kritiker Rainer Bölling
sieht fragwürdige Aussagen im OECD-Bericht

In den vergangenen Wochen wurde viel über die PISA-Studie gestritten. Der Deutsche Philologenverband forderte die Aussetzung der Studie, die Aufregung war groß, nachdem der aus Deutschland stammende Direktor für Bildung bei der OECD, Andreas Schleicher, in mehreren Zeitungsgesprächen scharfe Kritik am deutschen Bildungssystem und den Lehrern übte, die er als Befehlsempfänger bezeichnete. Professor Dr. Olaf Köller ist Geschäftsführender Direktor des Leibniz-Instituts für die Pädagogik der Naturwissenschaften und Mathematik (IPN) in Kiel und Mitglied des Vorstands des Zentrums für Internationale Vergleichsstudien (ZIB) in München, das in Deutschland für den deutschen PISA-Bericht zuständig ist. Dr. Rainer Bölling ist von Hause aus Lehrer, Bildungshistoriker, verfolgt die Datenerhebungen der PISA-Studien seit vielen Jahren und hat dazu kritische Texte publiziert.

Herr Köller, zu den gängigen Vorurteilen gehört, dass alle anderen Länder mehr Schüler vom Test ausschließen als Deutschland. Stimmt das? Im Länderbericht sprechen Sie von Zielpopulation und Teilnahmequote.

KÖLLER: Zur Zielpopulation in einem Land gehören alle Jugendlichen, die zu einem Stichtag zwischen 15 Jahren und drei Monaten und 16 Jahren und zwei Monaten alt sind. Die Zielpopulation wird in PISA aber dahingehend eingeschränkt, dass diese Gruppe auch in die Schule gehen und mindestens in der siebten Klasse sein muss. In Deutschland sind das alle 15 Jahre alten Jugendlichen. Darüber hinaus hat jedes Land die Möglichkeit, weitere Jugendliche auszuschließen.

Die Vorgabe der OECD ist, dass von allen 15-Jährigen nicht mehr als 5 Prozent ausgeschlossen werden sollten. Die Gründe für Ausschlüsse können sein: Die Schülerinnen und Schüler sprechen nicht die Testsprache oder werden in einer anderen Sprache beschult, sie sind nicht lang genug im Land oder physisch oder psychisch nicht in der Lage, den Test zu bearbeiten. In Deutschland lag diese Exklusionsquote im Jahre 2022 bei 2,5 Prozent, sodass 97,5 Prozent der in der Zielpopulation waren.

Aber es haben sich nicht alle Länder an die Vorgabe von höchstens 5 Prozent Ausschluss gehalten.

KÖLLER: Tatsächlich haben diese Fünfprozenthürde im Rahmen von PISA 2022 sechzehn Länder gerissen, Dänemark deutlich, das bei einer Ausschlussquote von 11,6 Prozent lag, und eine ganze Reihe von Ländern ganz knapp wie etwa Estland (5,9 Prozent).

Herr Bölling, Sie haben in früheren Veröffentlichungen darauf hingewiesen, dass die Ausschlussquote entscheidend für die Vergleichbarkeit der Ergebnisse ist, und auf ein besonders krasses Beispiel hingewiesen.

BÖLLING: Schanghai, das erstmals 2009 bei PISA auftauchte und wie auch 2012 mit großem Abstand an der Spitze lag, wies 2012 nur eine Erfassungsquote von 79 Prozent auf, weil die Wanderarbeiter- und Migrantenkinder gar keine höhere Schule besuchen durften. Die Ergebnisse von anderen chinesischen Provinzen, die teilgenommen hatten, wurden gar nicht veröffentlicht.

KÖLLER: Ein weiteres Beispiel ist Kolumbien, dort wurden von über 800.000 Fünfzehnjährigen 120.000 gar nicht beschult, zumindest nicht in Klasse sieben

oder höher. Die sind natürlich vornherein ausgeschlossen. Aber unter den Industrienationen, die uns interessieren, sind es wenige Länder, die den OECD-Standard von 5 Prozent Ausschlussquote substantziell reißen.

Wie wirkt sich die Höhe der Ausschlussquote auf das Ergebnis aus?

KÖLLER: Die OECD geht davon aus, dass 5 Prozent Exklusion ungefähr fünf PISA-Punkte ausmachen. Aber es gibt eine zweite Messlatte, die die OECD legt. Wenn die Zielpopulation feststeht, dann müssen sich 85 Prozent der in der Stichprobe gezogenen Schüler an der Testung beteiligen. In Deutschland hat das bei PISA 2022 dazu geführt, dass von den gezogenen Schülern 88 Prozent teilgenommen haben. Daraus kann man den Erfassungsgrad berechnen, der in Deutschland bei 92 Prozent lag. Wir haben etwa 20 OECD-Länder, die einen ähnlichen oder noch besseren Erfassungsgrad haben, und wir haben zwischen zehn und 15 Länder, die einen schlechteren haben.

Warum ist im deutschen Bericht dann nicht auch von Erfassungsgrad die Rede, wenn im OECD-Bericht nur der genannt ist?

KÖLLER: Wir halten das im technischen Teil immer sehr kurz. Wenn wir aber Aussagen über Spitzenländer wie Estland, Korea, Japan machen, dürfen das nur Länder sein, die alle Kriterien weitgehend erfüllen.

Herr Bölling, Sie haben in Ihren Veröffentlichungen immer vom Erfassungsgrad gesprochen.

BÖLLING: Es gibt ja mehrere Erfassungsgrade, zwei beziehen sich auf die nationale Schülerpopulation und drei auf die nationale Population der 15-Jährigen. In Län-



dern wie Deutschland sind beide praktisch identisch, in anderen Ländern aber nicht. Zum Beispiel ist Vietnam 2015 ziemlich weit oben im Ranking gelandet, aber der Erfassungsgrad drei der nationalen Population lag bei der Hälfte, ist später etwas korrigiert worden. Eigentlich hätte Vietnam also nicht in die Wertung einbezogen werden dürfen. Der Fall wurde schon angesprochen; er hat gerade in den USA kritischen Widerhall gefunden.

Wie viele Schüler, Herr Bölling, hat Deutschland denn noch auf der Ebene der Schulen ausgeschlossen?

BÖLLING: Deutschland hat dieses Mal bei einer anvisierten Teilnehmerquote von gut 6000 insgesamt 59 Schüler ausgeschlossen, Schweden 473 und Dänemark 902. Das sind schon ganz andere Größenordnungen, die sich auch auf das Ergebnis auswirken.

KÖLLER: Jein. Das sind die a priori Ausgeschlossenen. Für mich ist der entscheidende Index der Erfassungsgrad drei, der in Deutschland bei 92 Prozent liegt.

Zu den gängigen Feststellungen gehört ja, dass der Zusammenhang zwischen Bildungserfolg und sozialer Herkunft in Deutschland stärker ausgeprägt ist als anderswo. Herr Bölling, gibt es aus Ihrer Sicht dafür irgendwelche Belege in PISA?

BÖLLING: Diese These beruht auf dem PISA-Index des ökonomischen, sozialen und kulturellen Status (ESCS). Die Werte für Deutschland lagen bei den ersten beiden PISA-Erhebungen über 20 Prozent, das heißt: Mehr als 20 Prozent der Leistungsunterschiede waren auf die soziale Herkunft zurückzuführen. Allerdings hat sich dieser Wert für Deutschland im Laufe der Zeit verbessert. 2015 war er bis auf 15,7 Prozent gesunken und damit besser als der von PISA-Sieger Singapur. Dieser Wert ist allerdings jetzt wieder etwas angestiegen, was sicher auch mit den Flüchtlingen zusammenhängt, die seit 2015 ins Land gekommen sind. Das Problem ist aus meiner Sicht, dass immer so getan wird, als hätten die Werte sich nie geändert. Es gibt eine Reihe Länder, die auch heute über dem Wert Deutschlands liegen. Dazu gehören etwa Frankreich, Österreich, die Schweiz, Luxemburg und Ungarn.

KÖLLER: Die Sache mit den sozialen Disparitäten ist nicht trivial. Es liegt daran, dass in PISA zwei unterschiedliche Größen genutzt werden. Zum einen der sogenannte ISEI, (International Socio-Economic Index of Occupational Status), also das rein ökonomische Kapital. Findet nur

er Berücksichtigung, liegt Deutschland erheblich über dem OECD-Mittelwert. Wenn wir uns den von Herrn Bölling genannten Index der wirtschaftlichen Situation der Familien in Verbindung mit den kulturellen Angeboten anschauen, den

ESCS (Economic, Social and Cultural Status), ist Deutschland tatsächlich unauffällig. Es ist also eine Frage der Betrachtungsweise. Die GEW sucht sich dann den ISEI aus, und andere, die das nicht gut finden, nehmen den ESCS.

Je höher der Anteil von Migranten, desto schwächer sind die Ergebnisse, so lautet die Auffassung. Herr Bölling, es gibt aber auch Länder, in denen die Migranten bessere Ergebnisse erzielen.

BÖLLING: So steht es tatsächlich im aktuellen PISA-Band der OECD: „Länder mit einem höheren Anteil an Schüler*innen mit Migrationshintergrund verzeichneten auch höhere Durchschnittsleistungen in Mathematik.“ Dann werden Länder wie Hongkong, Kanada und die Schweiz genannt. Das ist eine Aussage, die wirklich in die Irre führt, denn es gibt ja ganz unterschiedliche Migranten. Kanada hat in der Tat ein gutes Drittel Migranten, verfolgt aber seit Langem eine sehr selektive Einwanderungspolitik, was bedeutet, dass nur qualifizierte Arbeitskräfte mit entsprechender Bildung und Sprachkenntnissen ins Land kommen.

Und die Schweiz?

BÖLLING: Dort hat man bisher darauf gesetzt, qualifizierte Akademiker nicht unbedingt im Lande auszubilden, sondern sie aus dem benachbarten Ausland zu holen. Und wer sind die größten Migrantengruppen in der Schweiz? Die kommen – in dieser Reihenfolge – aus Deutschland, Italien, Portugal und Frankreich. Sie sind deshalb Migranten, weil die Schweiz sehr zurückhaltend einbürgert; woanders wären sie wahrscheinlich längst Staatsbürger. Das ist eben ein völlig anderes Migrationsmuster, als wir es in Deutschland und anderen Ländern der EU kennen, wo wir als Migranten viele Asylsuchende haben, auch aus Ländern mit einem niedrigeren Bildungsniveau, aus arabischen Ländern, zu denen es kaum Zahlen gibt. Die nehmen ja kaum bei PISA teil.

Genannt werden ja auch die Arabischen Emirate, Herr Bölling.

BÖLLING: Die Migranten, die dort bei PISA die Mehrheit bilden, sind die qualifizierten ausländischen Arbeitskräfte, deren Kinder wahrscheinlich auf eine internationale Schule gehen. Und die schneiden natürlich besser ab als die einfachen Bewohner der Emirate. Da hat die OECD eine enorme Nebelkerze geworfen, die nicht akzeptabel ist.

KÖLLER: Kanada ist wiederholt als Land beschrieben worden, das qua Einwanderungsgesetz eine sehr positiv selektierte Zuwanderung hat. Und wir sehen, dass in den Ländern, die ähnlich wie wir 2015/16 von erheblicher Zuwanderung durch Flucht betroffen waren, die sogenannte

erste Generation in ihrer sozioökonomischen Situation ganz ähnlich ist. Das gilt für Belgien, die Niederlande, Österreich oder Frankreich. Das sind Geflüchtete und keine Arbeitsmigranten. Auffallend in der gesamteuropäischen Betrachtung ist aber, dass die erste Generation in keinem anderen Land so schlecht abschneidet wie in Deutschland. Wir müssen darüber nachdenken, was wir für diese Gruppe besser machen können.

Worin sehen Sie die Gründe dafür, Herr Köller?

KÖLLER: Wir halten es immer noch für wichtiger, die Zuwanderer frühzeitig in regulären Klassen zu integrieren als ihnen die deutsche Sprache hinreichend beizubringen. Wir lassen zu früh bei der Sprachförderung nach. Die Hälfte der ersten Generation ist 2015/16 nach Deutschland gekommen, zu dem Zeitpunkt waren sie in der zweiten Klasse. Wenn sie bis 2022 kein Deutsch gelernt haben, dann hat das viel mit unzureichender Sprachförderung zu tun. Wir sollten sie erst dann vollständig integrieren, wenn sie so gut Deutsch können, dass sie in allen Fächern anschlussfähig lernen können.

BÖLLING: Ich weiß nicht, ob Deutschland sich da so besonders schertut. Frankreich hat zum Beispiel keine besseren Werte, eher schlechtere. Der Ansatz, die Geflüchteten möglichst bald in den regulären Unterricht zu bringen, ist sicher verfehlt. Von Sprachkurslehrern weiß ich aber auch, dass die Arbeit vor Ort sehr frustrierend sein kann, dass manche Schüler sich beklagen über unser Land, das so eine blöde Sprache habe.

Herr Köller, zu den beliebten Schlussfolgerungen aus PISA gehört die Forderung, bis zur neunten Klasse gemeinsam zu lernen. Auch die wurde von Andreas Schleicher wieder erhoben, aber auch von Kollegen aus der Bildungsforschung. Gibt es dafür eigentlich in PISA irgendwelche Belege, die genau diese Schlussfolgerung zulassen?

KÖLLER: Nein.

Herr Bölling, sehen Sie welche?

BÖLLING: Nein.

Herr Köller, Sie haben mehrfach darauf hingewiesen, dass die Gymnasien in Deutschland einen besonderen Leistungsabfall von 50 PISA-Punkten in Mathe zu verzeichnen haben. Wie kommen Sie zu dieser Umrechnung?

KÖLLER: Ich habe mir die Leistungsentwicklung seit 2009 angeschaut und festgestellt, dass das Gymnasium auf der PISA-Metrik 50 Punkte verloren hat. Wir gehen seit über zwanzig Jahren der Frage nach, wie viele Punkte man in einem Schuljahr gutmachen kann. Und wir haben festge-



stellt, dass man 30 bis 40 Punkte in Mathematik, in Deutsch oder Englisch pro Jahr gutmachen kann. In den Naturwissenschaften sind es nur 20 bis 25 Punkte. Wenn die Gymnasiasten in Mathematik seit 2009 insgesamt 50 Punkte verloren haben, dann bräuchten sie gut anderthalb Schuljahre, um auf das Niveau zu kommen, auf dem sie 2009 waren.

Herr Bölling, Sie haben die Umrechnung in Schuljahre kritisiert. Warum?

BÖLLING: Zum einen haben diese Werte ja gewechselt. 2009 und 2012 galten 125 Punkte mal als Äquivalent für drei Schuljahre, 2015 waren es noch 100 für drei Jahre, und jetzt sollen schon 20 Punkte dem Lernfortschritt eines Schuljahres entsprechen. Das ist ein ganz neuer Wert, der aber laut OECD „nicht dazu verwendet werden sollte, Punktzahldifferenzen in Äquivalente in Schuljahren umzurechnen“. Hier gibt es also großen Klärungsbedarf.

Herr Köller, Sie haben mehrfach darauf hingewiesen, dass PISA bestenfalls Korrelationen, aber keine kausalen Erklärungen liefert. Das Problem ist, dass dann gerade OECD-Vertreter bestimmte Ergebnisse auf Kausalitäten zurückführen. Herr Bölling, Sie hatten ein Beispiel, zum Beispiel die frühkindliche Bildung. Vielleicht könnten Sie das noch mal kurz erläutern.

BÖLLING: Zu den OECD-Erzählungen gehört es ja, dass nach der ersten PISA-Studie vor allem die frühkindliche Erziehung ausgebaut wurde und das zu einer Verbesserung Deutschlands bei PISA geführt habe. Die besten PISA-Ergebnisse hatte

Deutschland 2012, die damals getesteten Fünfzehnjährigen wurden aber spätestens 2003 eingeschult, also vor dem Ausbau der frühkindlichen Bildung in Deutschland. Den falschen Zusammenhang hat seinerzeit sogar der stellvertretende Generalsekretär der OECD hergestellt. So etwas geht einfach nicht.

Herr Köller, sehen Sie das auch so?

KÖLLER: Wir haben in PISA korrelative Daten. Daraus kann man Hypothesen ableiten, wie sie zustande gekommen sind und was man tun muss, damit es besser wird. Der Dortmunder Bildungsforscher Wilfried Bos konnte bei den Daten der IGLU-Studie zeigen, dass die Kinder, die im Kindergarten waren, quer durch alle sozialen Schichten im Lesen besser abgeschnitten haben als Kinder, die nicht im Kindergarten waren. Wir haben viele Analysen, die die Bedeutung frühkindlicher Bildung zeigen. Aber es ist natürlich leichtfertig, aus PISA-Ergebnissen zu schlussfolgern, dass diese die Bedeutung frühkindlicher Bildung kausal belegen. Aber auch ich leite Hypothesen aus PISA ab. Auch ich behaupte ja, dass die sinkenden Gymnasialleistungen etwas mit der Unterrichtsqualität zu tun haben. Aber kausal lässt sich das natürlich nicht belegen.

Inwiefern kann PISA die Entwicklung des deutschen Schulsystems voranbringen?

KÖLLER: Wir brauchen Monitoring-Studien, die uns zeigen, wie sich das System im internationalen Vergleich entwickelt. Das halte ich für ganz zentral. Gerade bei dem trägen deutschen System müssen wir

hinschauen, was international gemacht wird. Die Tests haben sich seit PISA 2015 enorm weiterentwickelt, während die Test-Items beim IQB-Bildungstrend seit 2009 die gleichen sind. PISA zeigt den Innovationsgrad in Schulsystemen und weist darauf hin, wo sich etwas verändert – in Estland zum Beispiel.

Herr Bölling, worin sehen Sie die Leistung von PISA?

BÖLLING: Ich sehe viele Probleme, die nicht offen angesprochen werden, wo die Vergleichbarkeit infrage steht, und ich sehe auch eine Verengung der Auswertung in der Öffentlichkeit. Dafür steht besonders Andreas Schleicher, aber auch andere, die Behauptungen aufstellen, die allenfalls Vermutungen sind. Und es werden immer nur bestimmte Aspekte hervorgehoben. Man könnte ja auch mal fragen, warum Finnland so abgestürzt ist? Finnland hatte seine besten Ergebnisse 2006.

KÖLLER: Man muss damit leben, dass sich jeder das herauspickt, was zur eigenen Agenda passt. Aber wir sollten nicht versuchen, die Ergebnisse schönzureden und das System mit dem Hinweis zu schützen, PISA liefere nichts. Das ist genau das, was Trump in den USA mit der Klimaforschung gemacht hat.

Was kostet PISA im Jahr?

KÖLLER: Rund zwei Millionen im Jahr, darin sind die 5,5 Mitarbeiterstellen sowie die Erhebungskosten eingeschlossen.

Das Gespräch führte **Heike Schmall**.