

PISA: ein teurer Zufallszahlengenerator

Rangverschiebungen sind bedeutungslos

Kaum sind die Ergebnisse von PISA 2006 heraus, streiten sich die Verantwortlichen um die Deutungshoheit. OECD-Koordinator Andreas Schleicher warnt davor, die Daten ernst zu nehmen; Deutschland-Chef Manfred Prenzel widerspricht umgehend. Ist die Verbesserung Deutschlands um 14 Punkte eine echte Verbesserung? Darf man das diesjährige Naturwissenschafts-Ranking mit dem aus 2000 und 2003 vergleichen?

Die Antwort ist ernüchternd: die Ungenauigkeit von PISA ist noch viel größer als plus oder minus 14 Punkte. Allein die unterschiedliche Schülermotivation kann, amerikanischen Studien zufolge, mit 50 oder mehr Punkten durchschlagen. In Seoul wird vor der Testung die Nationalhymne gesungen; in Hamburg geben die ersten Schüler nach fünf Minuten ab.

Dazu kommen Uneinheitlichkeiten bei der Stichprobenziehung, bei den Teilnahmequoten und beim Ausschluss behinderter Schüler; Finnland schließt sogar Legastheniker aus. Auf den Verdacht, dass Polen 2003 unvollständige Testhefte unterdrückt hat, hat die OECD bis heute nicht reagiert. Der hohe Anteil von "Risikoschülern" in Deutschland liegt vielleicht nur daran, dass besonders erschöpfende Stichproben gezogen werden.

Und damit ist noch gar nicht die Kernfrage angeschnitten: kann man Schülerleistung überhaupt auf einer eindimensionalen Skala messen? In einer kultur- und sprachübergreifend fairen Art und Weise?

Verzerrungen und Übersetzungsfehler

Man kann nicht. Einige Verzerrungen liegen auf der Hand: 2003 waren die Aufgabentexte im Deutschen um 16% länger als im englischen Original. Während Australier oder US-Amerikaner schon in der Grundschule lernen, dass bei Multiple-Choice-Aufgaben immer nur eine Antwort richtig sein kann, kreuzen über 10% der deutschen Testteilnehmer auch mal mehrere Antworten an.

Der Test ist nicht frei von Übersetzungsfehlern. Offensichtlich war der Übersetzer bei der verbalen Beschreibung eines Funktionsgraphen überfordert: die Aufgabe zur Windenergie ist im Deutschen missverständlich.

In einer Multiple-Choice-Aufgabe wird die Antwort "Die Evolution ist eine Theorie, die durch Forschung bewiesen worden ist" als falsch gewertet. Über diese Wertung könnte man stundenlang diskutieren – in einem wissenschaftstheoretischen Oberseminar. Dass Fünfzehnjährige mit einer solchen

Frage konfrontiert werden, haben sie allein dem Übersetzer zu verdanken: an der Stelle von "Forschung" steht im englischen Original nicht "research", sondern "scientific experiments".

Und wo "hemisphere" mit "Hemisphäre" übersetzt wird, während es in Schulbüchern "Erdhälfte" heißt, ist es kaum verwunderlich, wenn die Antworten in überdurchschnittlichem Maße vom soziokulturellen Hintergrund abhängen.

Kein Naturwissenschaftstest

Im übrigen ist der PISA-Naturwissenschaftstest keineswegs ein reiner Naturwissenschaftstest. Nur wenige Aufgaben fragen klar umrissenes Sachwissen ab, wie zum Beispiel: welcher Planet kann zwischen Erde und Sonne treten, Merkur, Mars, Jupiter oder Saturn?

Andere Aufgaben sind überaus komplex, mit umfangreichem Text-, Bild- und Tabellenmaterial. Aus den Antwortstatistiken kann man nicht ersehen, ob die Schüler eher am Lesen, an der Logik oder an fehlendem Spezialwissen scheitern. Hätten Sie gewusst, ob der Stichlingsmann die Stichlingsfrau am Gesicht oder am Bauch erkennt?

Bei etlichen Aufgaben ist die Kunst gefragt, einen Antwortsatz hinzuschreiben, der Information aus dem Text aufgreift, ohne sie wörtlich zu wiederholen. Manchmal ist schwer zu erraten, was die Prüfer hören wollen. Erst die Korrekturanweisungen, die nur in der englischen Aufgabenveröffentlichung vollständig wiedergegeben sind, zeigen, wie sehr das, was hier gemessen wird, von anglo-amerikanischen Prüfungsgewohnheiten geprägt ist.

Schleicher hat vollkommen recht, wenn er vor einer Überinterpretation der PISA-Daten warnt: Aus 14 Punkten Veränderung einen Trend abzulesen, ist unseriös.

Und deutsche Kultusminister haben vollkommen recht, wenn sie ihm den Geldhahn zudrehen wollen. Denn die dreijährigen Wiederholungen von PISA sind eingekauft worden, um Trends zu messen. Stattdessen bekommt man Zufallszahlen.

Joachim Wuttke

Der Autor ist habilitierter Physiker mit Schulerfahrung. Seine Kritik an der PISA-Statistik ist in zwei Sammelbänden erschienen: "PISA & Co - Kritik eines Programms", hrsg. v. Jahnke/Meyerhöfer, Franzbecker-Verlag und "PISA zufolge PISA", hrsg. v. Hopmann/Brinek/Retzl, LIT-Verlag.