

Kleiner Nachruf auf PISA – nebst einer Warnung vor den verheerenden Folgen dieses Unternehmens

Das schlechte Abschneiden der deutschen Schülerinnen und Schüler bei dem ersten Durchgang des Programme for International Student Assessment (PISA) schockierte die deutsche Öffentlichkeit 2001 derart, dass die einfachsten Fragen – etwa im Sinne eines Elchtest – unterblieben: Was war hier eigentlich in wessen Auftrag von wem zu welchem Zweck und mit welchen, wie zustande gekommenen Ergebnissen getestet worden?

Darauf kann man inzwischen recht gelassen antworten:

Wer? Der Auftraggeber war die OECD, die Auftragnehmer waren fünf Forschungseinrichtungen; bei vier davon handelt es sich um privaten Unternehmen, sogenannte *educational-assessment*-Formen, die PISA entwickelt und damals bereits an 58 Staaten verkauft hatten. Das Auftragsvolumen ist beträchtlich. Nach Angaben aus F.D.P.-Kreisen kostet ein internationaler PISA-Durchgang 300 Millionen US-Dollar. Die Markterweiterung der Testindustrie auf Deutschland war – und ist durch die Periodisierung von PISA-Durchgängen im 3-Jahres-Rhythmus auch dauerhaft – gelungen.

Was? In der Mathematik wurde und wird bei PISA ‚mathematische Grundbildung‘ (‚mathematical literacy‘) getestet, wobei offen zugegeben wird, dass man sich dabei um die deutsche Lehrpläne nicht schert: „Der Begriff Grundbildung wurde gewählt, um zu betonen, dass mathematische Kenntnisse und Fähigkeiten, wie sie im traditionellen Curriculum der Schulmathematik definiert werden, im Rahmen von OECD/PISA nicht im Vordergrund stehen,“ erläutern die deutschen PISA-Verantwortlichen. Wie, könnte man fragen, kann man die Leistungen von Schulen und Schülerinnen und Schülern an Zielen messen, die sie gar nicht verfolgen? „PISA“, analysiert der Soziologe Richard Münch, „bringt eine andere Lernkultur zur globalen Vorherrschaft, als es der deutschen Tradition von klassischer Bildung und Vermittlung von Fachwissen entspricht, ohne dass man ohne weiteres die generelle Überlegenheit der neuen Lernkultur über die alte deutsche Fachbildungstradition behaupten könnte.“ Der Erziehungswissenschaftler Volker Ladenthin sieht hierin eine grundsätzliche Funktionalisierung von Bildung: „Es geht erklärtermaßen der OECD (als Wirtschaftsorganisation) um die Durchsetzung eines zwar als "angelsächsisch" bezeichneten, in Wirklichkeit aber längst auch bei uns bekannten, schlicht funktionalen Bildungsbegriffs; es geht um Universalisierung (also ökonomische Globalisierung), d.h. die Auflösung föderaler und letztlich auch nationaler Bildungssysteme; es geht um die Transformation von Selbstbestimmung und Kulturidentität in funktionale Basiskompetenzen; es geht um Normierung und Vereinheitlichung statt um Individualisierung und Differenzierung. Als Grund und Legitimation für diesen Wandel werden (...) ausdrücklich und ausschließlich die von der ökonomischen Situation her verursachten und definierten "Qualitätsanforderungen" benannt.“

Wozu? Zum Zweck der Testungen liest man im Vorwort des Buches PISA `06: „Die Auftraggeber, das sind die OECD-Staaten und weitere Staaten, erwarten von dem internationalen Vergleich empirisch fundiertes Steuerungswissen. Die teilnehmenden Staaten möchten von PISA zum Beispiel erfahren, wo sie im internationalen Vergleich stehen, welche Bildungsergebnisse andere Staaten erreichen und unter welchen Bedingungen das geschieht.“ Der Erziehungswissenschaftler Klaus-Jürgen Tillmann, der in einem hochrangigen Forschungsprojekt „Ministerielle Steuerung und Leistungsvergleichstudien“ deren politische Folgen in vier Bundesländern detailliert untersucht hat, konstatiert ernüchternd: „Im Ergebnis zeigt sich, dass PISA nur selten neue Programme initiiert –

aber um so häufiger herangezogen wird, um die ohnehin verfolgte politische Linien zu legitimieren. Die Hoffnung, dass PISA ‚Steuerungswissen‘ produziert, lässt sich kaum einhalten.“ Um doch diesen Eindruck zumindest zu erwecken, definieren die deutschen PISA-Verantwortlichen das Steuerungswissen selbstreferentiell: Wenn die Ergebnisse der deutschen Schülerinnen und Schüler beim nächsten PISA-Durchgang besser sind, dann sind diese gebildeter. Dies erklärt auch, warum nach dem Durchgang 2006, dessen Ergebnisse von der Öffentlichkeit kaum noch zur Kenntnis genommen wurden, die Medien sich hauptsächlich nur für den Streit zwischen dem OECD-Koordinator von PISA und den deutschen Pisanern interessierten, ob eine geringfügige Erhöhung der erreichten Punktzahl gegenüber 2003 als Fortschritt zu werten sei. Auch wenn im Prinzip eine solche Zunahme zu erwarten war, weil sich die deutschen Schülerinnen und Schüler und ihre Lehrerinnen und Lehrer auf die Testungen einzustellen wissen, gaben die Ergebnisse das kaum her.

Wie? Der Genauigkeitsanspruch von PISA wird von neutralen Expertinnen und Experten angezweifelt. Der im Umgang mit großen, schlecht konditionierten Datenmengen vertraute Physiker Joachim Wuttke stellt fest: „Allein schon 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. Allein solche Faktoren können mehr als 10 Punkte ausmachen. Und damit ist noch gar nicht die Kernfrage angeschnitten: kann man Schülerleistungen überhaupt auf einer eindimensionalen Skala messen? In einer kultur- und sprachübergreifenden fairen Art und Weise? Man kann nicht.“ Wuttke kommt zu dem Schluss, dass PISA ein teurer Zufallszahlengenerator ist.

Das kümmert die deutsche Kultusministerkonferenz wenig; sie dokumentierte ihre Tatkraft vielmehr durch einen Beschluss zum ‚Bildungsmonitoring‘, in dem PISA bis 2018 festgeschrieben und zum internationalen Maßstab erhoben wird; darüber hinaus werden eine Vielzahl von bundes- und landesweiten Test, Lernstandserhebungen und Vergleichsarbeiten als Gradmesser der ‚Erträge‘ schulischer Bildung einführt. Deutschland ist zum Testland geworden. Das ist die fundamentale Wirkung von PISA, nicht die marginalen Änderungen in den Ergebnissen der einzelnen Durchläufe, die wenig besagen, sondern dass Schulbildung zur Test-Bildung geworden ist und der Unterricht auf diese und auf das Erzielen guter Testergebnisse ausgerichtet wird.

Die Auswirkungen von mehr als zwei Jahrzehnten testresultat-gesteuerter Bildungspolitik sind in den USA und Großbritannien wissenschaftlich untersucht worden. Der Mathematikdidaktiker Uwe Maier fasst sie in folgenden Stichpunkten zusammen:

- Schulen, die aufgrund schlechter Testwerte von Sanktionen bedroht sind, reduzieren ihr Curriculum auf testrelevante Fächer.
- Auch innerhalb der getesteten Fächer werden die Lerninhalte deutlich eingeschränkt; z.B. Reduktion des Mathematikunterrichts auf Basisfertigkeiten.
- Durch unterschiedliche Strategien bzw. „Schummeleien“ werden die Testergebnisse manipuliert.
- Schwache Schüler werden vor dem Test vom Unterricht ausgeschlossen bzw. nur Schüler mit „Steigerungspotential“ erhalten eine Lernförderung.
- Sinkende Motivation bei Lehrern und Schülern und Zunahme der Schulabbrecherquoten.

Die American Evaluation Association (AEA) warnt dringlich in einem Position Statement vor einer testresultat-gesteuerten Bildungspolitik. Auszüge (in meiner Übersetzung):

„Unser Verständnis der bisherigen Belege aus zwei Jahrzehnten zeigt, dass Tests nicht zu einer besseren Bildungspolitik und Praxis führen, auch wenn sie gelegentlich das Lehren und Lernen in einigen Klassenzimmern und Schulen verbessern. Die Konsequenzen, die uns am meisten besorgen, sind erhöhte Ausfallraten, die Deprofessionalisierung der Lehrer und der Schulverwaltung, der Verlust der Integrität des Curriculums, abnehmendes kulturelles Gespür, unverhältnismäßige Verteilung der Bildungsressourcen in Testprogramme statt in die Anwerbung qualifizierter Lehrer und die Bereitstellung vernünftiger Bildungsprogramme. (...) Schul- und Schülervergleiche auf Grund von Testergebnissen befördern ein teaching to the test insbesondere in einer Art und Weise, die nicht zu einer Verbesserung im Lehren und Lernen führt. Obwohl solche Testungen seit mehr als zwei Jahrzehnten durchgeführt werden, haben sie weder die Qualität der Schulen verbessert noch in Ungleichheiten in den Leistungen verringert weder was die Geschlechtsunterschiede noch was die Rassen- und Klassenunterschiede anbetrifft; noch haben sie das Land in moralischer, sozialer oder ökonomischer Sicht vorangebracht.“

In Deutschland werden durch die Periodisierung von PISA und die Vielzahl anderer Testungen die Testorientierung des Unterrichts und die Testfähigkeit und damit die erreichten Punktzahlen der Schülerinnen und Schüler zwar mäßig zunehmen (Die Sau wird vom Wiegen zwar nicht fetter, aber sie lernt sich schwer zu machen), aber über eine bessere Schulbildung sagt das nichts aus. Auch die Lehrerinnen und Lehrer werden in ihrer schweren Arbeit durch die Testungen nicht unterstützt, sondern unter Druck gesetzt und behindert. Die Mängel des deutschen Schulsystems sollen nicht bestritten werden, aber durch eine Testbildungs-Politik oder eine Test-Bildungspolitik werden sie eher verstärkt und durch neue ergänzt statt bearbeitet und möglicherweise beseitigt.

Der Autor ist Professor für Didaktik der Mathematik an der Universität Potsdam und Herausgeber und einer der Autoren des Buches: PISA&Co – Kritik eines Programms. Hrsg.: Th. Jahnke und W. Meyerhöfer. Zweite erweiterte Auflage. Franzbecker Verlag. Hildesheim 2007. ISBN 9783881204644