

Rainer Block

Modelle zur Berechnung der Leistungseffektivität von Schulen

Bericht zur aktuellen OECD-Expertise „Measuring Improvements in Learning Outcomes. Best Practices to Assess the Value-Added of Schools“

Value-Added Modelling to Estimate School Effects

Report on the Topical OECD Publication “Measuring Improvements in Learning Outcomes. Best Practices to Assess the Value-Added of Schools”

Um es direkt zu Anfang in aller Klarheit zu sagen: Die Zielgruppen der jüngsten OECD-Publikation sind – entgegen anders lautender Behauptungen im Klappentext – die politischen Entscheidungsträger und die Bildungsadministration der OECD-Mitgliedsländer. Gerade deswegen aber lohnt sich die Lektüre, zeigt sie doch auf, welcher Diskurs – aller Wahrscheinlichkeit nach – in naher Zukunft verstärkt Einzug in die schulpolitische Diskussion auch der Bundesrepublik Deutschland halten wird.

Auf 220 Seiten werden den „policy makers and administrators“ die Vor- und Nachteile verschiedener (statistischer) Modelle und Systeme zur Berechnung der Leistungseffektivität von Schulen (so genanntes Value-Added-Modelling) handbuchartig aufbereitet, um schließlich konkrete politische, institutionelle und technische Handreichungen zur Implementierung derartiger Messmodelle anzubieten. Warum hat die Bildungsadministration der Länder – aus Sicht der OECD – eine derartige Nachhilfe überhaupt nötig? Der Grund ist einfach: Das in vielen OECD-Ländern mittlerweile installierte outputorientierte Steuerungsmodell im Schulwesen, häufig verbunden mit einem öffentlichen Ranking der Leistungseffektivität der einzelnen Schulen, ist auf aussagekräftige und valide Indikatoren zur Messung und Bewertung eben jener Leistungsfähigkeit der Schulen angewiesen, damit es seine intendierten Steuerungswirkungen entfalten kann. Wie der OECD-Bericht lehrt, sind aber nur in wenigen OECD-Ländern Mess- und Bewertungsverfahren implementiert, die den wissenschaftlichen und akademischen Ansprüchen an einen ‚fairen‘ Schulvergleich genügen. Verbreitet sind diese Modelle –

nach Ansicht der Autoren – hingegen insbesondere in den USA und im Vereinigten Königreich.

Der OECD-Bericht geht auf eine Arbeitsgruppe im Rahmen des „Education Committee of the OECD, Directorate for Education“ aus dem Jahre 2005 zurück, geleitet von Ben Jensen, unter Beteiligung externer Wissenschaftler (allen voran Henry Braun). An der Arbeitsgruppe haben – ohne Deutschland – folgende 13 Staaten teilgenommen: Australien, Belgien (Flämischer Teil), Dänemark, Frankreich, Niederlande, Norwegen, Polen, Portugal, Schweden, Slowenien, Spanien, Tschechische Republik und Vereinigtes Königreich. Der Bericht basiert auf vier zweitägigen Treffen der Arbeitsgruppe. Er fokussiert insbesondere die technischen, statistischen Aspekte des Value-Added-Modelling. ‚Value Added‘ meint dabei den Mehrwert, also den Zugewinn an Leistung, der unmittelbar auf die Schule und konkret das Unterrichtshandeln zurückzuführen ist. Value-Added-Modelling wird wie folgt definiert: „A class of statistical models that estimate the contributions of schools to student progress in stated or prescribed education objectives (e.g. cognitive achievement) measured at at least two points in time“ (S.17). Dies ist eine erstaunliche Engführung gegenüber den sonst üblichen Definitionen im Kontext der Schuleffektivitätsforschung (vgl. Hill 2001; Scheerens/Bosker 1997), werden so doch Querschnittuntersuchungen zur Erfassung der Schuleffektivität explizit aus der Gruppe des Value-Added-Modelling ausgeschlossen. Zur Erinnerung: In der Bundesrepublik Deutschland basieren die Schuleffektivitätsmessungen im Rahmen der Lernstandserhebungen beispielsweise ausschließlich auf Querschnittuntersuchungen.

Im weiteren Verlauf werden die Modelle dann im Detail vorgestellt: Linear regression value-added models, Variance component und random effect models, Fixed-effects value-added models, Dallas model, Multivariate random effect response models, Growth curve analysis. Diese Auflistung hat man so schon anders und zum Teil auch differenzierter gesehen (z.B. Lissitz 2005). Den Bericht zeichnet aber aus, dass er kein Modell a priori präferiert, die Vor- und Nachteile der einzelnen Ansätze thematisiert, auf die hohe Kongruenz der Ergebnisse über die einzelnen Modelle hinweg verweist und die Entscheidung für oder gegen ein spezifisches Modell nicht allein aus statistischen Erwägungen ableitet, sondern wesentlich vom konkreten administrativen Verwendungskontext des Messmodells abhängig macht.

Damit wären wir beim zweiten zentralen Teil des Bandes, der Implementierung von Value-Added-Ansätzen. Die Autoren schlagen den politischen Entscheidungsträgern und der Bildungsadministration ein Acht-Phasen-Entscheidungsmodell vor: Phase 1: Setting policy objectives and school performance measures, Phase 2: Presentation and use of value-added information, Phase 3: Data quality, Phase 4: Choosing an appropriate value-added model, Phase 5: Communication and stakeholder engagement strategies, Phase 6: Training, Phase 7: Pilot programme, Phase 8: Ongoing development. In diesem Zusammenhang erachten die Autoren zwei zu entscheidende und miteinander verbundene Schlüsselfragen als zentral. Erstens: Sollen die modellierten Schuleffektivitätswerte

veröffentlicht werden oder nicht? Zweitens: Produziert das gewählte statistische Modell so kleine Standardfehler, dass auch für kleinere Schulen aussagekräftige Effektivitätswerte berechnet werden können?

Der Bericht bietet insgesamt einen kompaktalen Überblick über die in der wissenschaftlichen Schuleffektivitätsforschung derzeit diskutierten und in der Bildungspraxis angewandten Value-Added-Modelle. Darüber hinaus kann er als Handreichung zur politischen und administrativen Implementierung derartiger Modelle herangezogen werden.

Insbesondere zwei Aspekte an diesem Bericht sind kritisch zu bewerten:

1) Obwohl der Bericht sowohl die Vor- als auch die Nachteile der einzelnen Value-Added-Modelle thematisiert, fällt die methodenkritische Bewertung der Ansätze insgesamt zu positiv aus. Die Probleme – insbesondere bei der Adjustierung von Leistungswerten, gerade wenn es darum geht, nicht nur die Schuleffektivität, sondern auch die Lehrereffektivität zu messen – sind Legion und in anderen Kontexten schon deutlicher formuliert worden (z.B. Lissitz 2005; Molnar 2002).

2) Der Bericht ist weitgehend blind hinsichtlich der kontraproduktiven Effekte für die Schulentwicklung, die mit der Einführung einer von öffentlichen Schulrankings begleiteten Schuleffektivitätsmessung einher gehen können. Nur an zwei Stellen in dem 220 Seiten langen Text und sehr cursorisch werden die nicht erwünschten, möglichen Nebeneffekte des Value-Added-Modelling in der Praxis thematisiert (u.a. unter dem Oberbegriff „incentives and sub-optimal outcomes“). Demgegenüber haben Nichols und Berliner (2007) in ihrer empirischen Untersuchung mit dem Titel „Collateral Damage: How High-Stakes Testing Corrupts America's Schools“ die verheerenden Effekte eines derartigen Steuerungsmodells für das Schulwesen in den USA umfangreich dokumentiert, wie z.B. die tendenzielle Verengung des Schulcurriculums, die Konzentration auf abfragbares Test-Wissen, Testmanipulationen durch Lehrkräfte, ein manipulativer Ausschluss schwacher Schüler/innen von Leistungsüberprüfungen, die Konzentration von Lehrerbemühungen auf Grenzfälle im Leistungsniveau (Vernachlässigung schwacher Schüler/innen), den Anstieg vorzeitiger Schulabgänger ohne Abschluss.

Derlei Problemstellungen bewegen sich jedoch jenseits des statistischen Horizontes, der den Schwerpunkt des OECD-Berichtes bildet. Böse Zungen könnten deshalb auch mutmaßen, die OECD-Handreichung sei ein Versuch, den Bildungsbereich der OECD-Staaten weiter für die kommerzielle Testindustrie zu öffnen.

Literatur

- Hill, P. (2001): Perspectives on Education. Teaching and School Effectiveness. Melbourne: Department of Education, Employment and Training.
- Lissitz, R.W. (Hrsg.) (2005): Value Added Models in Education: Theory and Applications. Maryland: Jam Press.
- Molnar, A. (2002): School Reform Proposals – The Research Evidence. Tempe: Arizona State University. Education Policy Research Unit. URL: http://epsl.asu.edu/epru/epru_2002_Research_Writing.htm; Zugriffsdatum: 16.03.2009.
- Nichols, S.L./ Berliner, D.C. (2007): Collateral Damage: How High-Stakes Testing Corrupts America's Schools. Harvard: Harvard Education Press.
- OECD (2008): Measuring Improvements in Learning Outcomes. Best Practices to Assess the Value-Added of Schools. Paris: OECD Publications. URL: <http://masetto.source.oecd.org/vl=2592460/cl=37/nw=1/rpsv/cw/vhosts/oecdthemes/99980029/v2008n20/contp1-1.htm>; Zugriffsdatum: 17.03.2009.
- Scheerens, J./Bosker, R.J. (1997): The Foundations of Educational Effectiveness. Oxford: Elsevier Science Ltd.

Rainer Block, Dr. phil., geb. 1961, wissenschaftlicher Mitarbeiter im Arbeitsbereich Bildungsforschung an der Johannes Gutenberg-Universität Mainz.

Anschrift: Johannes Gutenberg-Universität Mainz, Institut für Erziehungswissenschaft, AG Schulforschung/Schulpädagogik, Colonel-Kleinmann-Weg 2, 55128 Mainz
E-Mail: rainer.block@arcor.de