

Das kurze Gedächtnis der großen Reformen

Bildungsstandards können nicht als „Output“ kontrolliert werden

Kritische Anmerkungen von Hans Brügelmann zu:

Böttcher, W. (2003): Bildung, Standards, Kerncurricula. Ein Versuch, einige Missverständnisse auszuräumen. In: Die Deutsche Schule, 95. Jg., H. 2 (im Druck; Verweise beziehen sich auf das Ms.).

Böttcher ist ein begnadeter Rhetoriker – und Polemiker. Er neigt zu Vereinfachungen, unangekränkt von auch nur den leisesten Selbstzweifeln. Das stimmt misstrauisch.

Böttcher benennt „wunde Punkte“ im Schulsystem, die nicht verdrängt werden dürfen. Aber er suggeriert zu Unrecht,

- er habe mit seinem Konzept von „Standards“ und dem „Kerncurriculum“ Lösungen für diese Probleme;
- diese Ideen seien neu und bewährt;
- der Grundschulverband verdränge die Probleme und verweigere sich Versuchen, sie zu lösen.

Die von ihm benannten Probleme –

- eine oft fehlende Fokussierung des Unterrichts auf zentrale Erfahrungen;
- eine mangelnde Förderung von sozial benachteiligten Kindern;
- unzureichend entwickelte Rechenschaftsformen im Schulwesen – sind ernst zu nehmen.

Insofern besteht auch Übereinstimmung mit zentralen Forderungen des Grundschulverbands (vgl. die parallele Replik von Bartnitzky und die dort genannten Publikationen des Grundschulverbands zur „Leistung“, zur „Evaluation“ und zu den „Bildungsansprüchen der Kinder“):

- gegen die Beliebigkeit von Unterricht
- gegen eine „Glockenkurvenmentalität“ der Leistungsbeurteilung
- gegen Schule als rechenschaftsfreien Raum.

Aber die Frage ist, wie man diesen Forderungen am besten gerecht werden kann – ohne sich neue Probleme einzuhandeln:

- Schule ist nicht „im Kern“ der Ort, an dem die ältere Generation bestimmt (5¹);
- Standards sind im Bildungsbereich nicht (zumindest nicht allein) „output“-bezogen, sondern vorrangig als Prozesskriterien für Lernsituationen zu definieren;
- Ziele sind nicht als verbindliche Niveaus (4), sondern als fokussierte Perspektiven für die Entwicklung von Kindern zu bestimmen;
- repräsentative Leistungsdaten sind für LehrerInnen als „Spiegel“ und Bezugsrahmen für die Einordnung der eigenen Klasse wichtig, aber nicht als Vorgabe des „Outputs“ sinnvoll.

¹ Ich beziehe mich auf die Seiten des Druck-Manuskripts

Erstaunlich ist, wie a-historisch jemand, der für *Bildungsstandards* plädiert, ein so zentrales Thema angeht. So beschreibt Böttcher (2) die „standards based reform“ als Modell, das seit gut zehn Jahren in den USA betrieben werde². Das klingt erstens nach einer neuen Erfindung und zweitens nach einer Erfolgsgeschichte – vor allem, wenn man an die zu erreichenden Ziele „Anreize“ (6) binde³.

Demgegenüber ist festzuhalten:

- Bereits vor 150 Jahren in Zeiten knapper öffentlicher Kassen (Krim-Krieg!) bekamen SchülerInnen und Schulen im viktorianischen England öffentliche Zuschüsse nur beim Nachweis bestimmter Fachleistungen. Dieses System des „payment by results“ hat dank entsprechend definierter Anforderungen zu erheblichen Einsparungen im Staatshaushalt geführt, pädagogisch kann es aber so erfolgreich nicht gewesen sein, sonst hätte es sich wohl länger erhalten (House 1975, 73; s. a. Ascher 1996).
- Um 1970 waren es die VertreterInnen des programmierten Unterrichts und des „lernzielorientierten Ansatzes“, die glaubten, über detailliert ausgewiesene Curricula vorher festgelegte Lerneffekte garantieren zu können. War es nur die Uneinsichtigkeit der KritikerInnen (z. B. des Deutschen Bildungsrats 1974), die dazu führte, dass nach wenigen Jahren kaum mehr etwas von diesen Versprechungen zu hören war?
- Vor über 30 Jahren begannen die USA ein regelmäßiges „National Assessment of Educational Progress“ in den zentralen Fächern und parallel führten einzelne Bundesstaaten landesweite Testprogramme ein. Wie die Daten des NAEP zeigen, haben sich die Leistungen, z. B. im Lesen, seit 1970 kaum verändert, sind z. B. im Lesen bei den 17-Jährigen seit 1971 nur von 285 auf 288 Punkte „gestiegen“. Staaten wie New York und Kalifornien haben bereits in den 1970er Jahren die Leistungen der einzelnen Schulen auf Erwartungswerte bezogen, die aus den Lernvoraussetzungen der SchülerInnen und sozioökonomischen Indikatoren ihres Einzugsbereichs errechnet wurden, um erwartungswidrig niedrige Leistungen zum Anlass schulspezifischer Untersuchungen und Maßnahmen zu nehmen (Shepard 1979). Über 50% der Schulbezirke in Kalifornien behaupteten damals, ihre Programme aufgrund der gewonnenen Daten verändert zu haben (Carlson 1979) – mehr als 25 Jahre später sind die damaligen Probleme mit grundlegenden Leistungen unverändert Anlass für analoge Maßnahmen.
- Von 115 Schulbezirken in den USA, die sich 1978 darauf eingelassen hatten, ihre LehrerInnen nach den erreichten Schülerleistungen zu bezahlen, hatten schon fünf Jahre später weit über die Hälfte (68) diesen „merit pay plan“ wieder aufgegeben (Terhart 1997, 73, mit Hinweis u. a. auf empirische Belege von Murnane/ Cohen 1986 und Jacobsen 1989).
- Vor über 30 Jahren haben Firmen in den USA einzelnen Schulbezirken angeboten, die Leistungen der Schülerschaft gegen Honorar auf vereinbarte Werte zu steigern. Wegen fehlender Erfolge haben die Schulbezirke dieses „performance contracting“ bald wieder aufgegeben (House 1975, 73-74). Ende der 1990er Jahre starteten Privatfirmen wie „Edison“ einen erneuten Anlauf. Die Zwischenbilanz nach sieben Jahren (O'Reilly 2002):

² Einen ähnlichen Eindruck von „Modernität“ vermittelt das – ansonsten reflektiertere und gegenüber dem Import von US-Praktiken kritischere Gutachten zu „Bildungsstandards“ von Klieme u. a. (2003), das sich primär auf Literatur nach der Jahrtausendwende stützt und im Übrigen nur Titel zurück bis 1990 berücksichtigt.

³ Differenzierter: Böttcher, W. (2001): Ist ein Dialog zwischen Pädagogik und Ökonomie möglich? Eine Diskussion am Beispiel von Anreizsystemen und Schulentwicklung. In: Zeitschrift für Pädagogik, 47. Jg., Nr. 6, 893-911.

- die Leistungen der SchülerInnen haben sich entgegen dem Versprechen hoher Zuwächse im Durchschnitt nicht verbessert;
 - unzufriedene Vertragspartner haben die Verträge vorzeitig gekündigt;
 - die Firma macht keine Gewinne, sondern gerät immer tiefer in die Verlustzone;
 - die Aktien sind nach einem Höchststand von 37 \$ inzwischen auf einen Wert von 2 \$ gesunken.
- US-Bundesstaaten, die in den 90er Jahren die Vergabe von Abschlusszeugnissen, Gehaltszuschüsse für LehrerInnen und die finanzielle Ausstattung der Schulen an die Ergebnisse in ihren jährlichen Tests gebunden haben, berichten immer wieder über einen Anstieg der Leistungen auf diesen Tests. Inzwischen haben sich 28 Staaten, also mehr als die Hälfte diesem Konzept verpflichtet. Eine Studie von Amrein/ Berliner (2002) zeigt aber, dass der berichtete Leistungsanstieg in der Regel nur für den engen Bereich dieser Tests gilt; in unabhängigen Tests, z. B. für die Zulassung zu den Hochschulen, zeigt sich für 2/3 der 28 Staaten eine *Abnahme* der Testleistungen, zusätzlich werden wachsende Dropout-Raten berichtet, d. h. dass leistungsschwächere SchülerInnen aus dem System ganz herausfallen (vgl. auch die Zusammenfassung und den Kommentar von Winter in der New York Times v. 28.12.2002).
 - Das englische Department of Education and Skills musste zur Jahreswende 2002/2003 eingestehen, dass das Reformziel verfehlt und die versprochenen Leistungszuwächse nicht erreicht wurden: über drei Jahre hinweg waren es nur 65% der SchülerInnen, die das definierte Schwellenniveau erreichten (Guardian v. 5.1.2003). Vergessen ist, dass bereits vor über 20 Jahren eine „Assessment of Performance Unit“ eingerichtet worden war (Simon 1979; Kay 1979), die „firmer evidence about the standards of attainment nationally“ erheben sollte (CERI 1978, 12) – offensichtlich ohne größere Wirkung auf die Entwicklung der Qualität von Unterricht, wie die aktuellen Bemühungen zeigen.
 - Bereits Mitte der 70er Jahre gab es in Deutschland eine breite Kontroverse über einheitliche Prüfungsanforderungen im Abitur (die sog. „Normen-Bücher“, vgl. Flitner/ Lenzen 1977), in der die Schwierigkeiten, sich auf einheitliche Standards zu verständigen, und zudem, diese produktiv und verlässlich umzusetzen⁴, differenziert diskutiert wurden.
 - Auch die OECD selbst hat bereits vor über 20 Jahren ein „Basic Education Policies“-Projekt durchgeführt, in dem Standards und deren Evaluation eine zentrale Rolle spielten. Schon damals wurden Rechenschaftsformen aus verschiedenen OECD-Ländern dokumentiert und ihr jeweiliges Potenzial kritisch gesichtet (vgl. Brügelmann 1980).
 - Und sogar mehr als 30 Jahre ist es her, dass eine intensive Debatte über Annahmen und Wirkungen von „Input“- und „Output“-Modellen der Unterrichtsplanung und der Systemsteuerung geführt wurden (vgl. zur internationalen Diskussion u. a. Stenhouse 1975, für die BRD die Hefte der Zeitschrift „Thema Curriculum“ 1972-1974).
 - Auch die empirischen Befunde, auf die Böttcher selbst verweist, sind durchaus gemischt (vgl. dazu Becker 2001; 2002).

Liest man diese kurze Liste, so fallen mehrere Dinge auf (ausführlicher: Brügelmann u. a. 1999):

- Was heute als neues Rezept verkauft wird, hat eine lange Geschichte.

⁴ Ein Problem, das Böttcher selbst leider nur in einem Satz anspricht – dabei zeigt die von seinem Kollegen Becker (2001; 2002) referierte US-amerikanische Studie der John Hopkins University, dass hier eine sehr ernst zu nehmende Klippe liegt.

- Effekte, die für die Umsetzung dieser Rezepte versprochen werden, sind in der Vergangenheit nicht eingetreten.
- Einwände, die schon damals geltend gemacht wurden, werden heute (zum Teil: wieder...) nicht zur Kenntnis genommen.

Für manche scheint die Geschichte grundlegender „System-Reformen“ erst Mitte der 90er Jahre zu beginnen. Dass sie weder ihre Vorgänger noch deren Kritiker zur Kenntnis nehmen, ließe sich vielleicht noch hinnehmen. Aber dass Protagonisten einer empirisch zu fundierenden Evaluation Erfahrungen und Befunde zur Effektivität ihrer eigenen Vorschläge übersehen, erstaunt schon.

Amrein, A. L./ Berliner, D. C. (2002): High-stakes testing, uncertainty, and student learning. In: Education Policy Analysis Archives, Vol. 10, No. 18. [<http://epaa.asu.edu/epaa/v10n18/>].

Ascher, C. (1996): Performance contracting: A forgotten experiment in school privatization. In: Phi Delta Kappan, Vol. 77, No. 9, 615-621.

Becker, D. E. (2001): Ein Kerncurriculum für die amerikanische Schule. Theoretische Grundlagen, Praxis und Kritik des Reformansatzes von E. D. Hirsch Jr. Vervielf. Ms.

Becker, D. H. (2002): Empirische Befunde zum Kerncurriculum: Implementierungsbedingungen und Effekte. In: Böttcher/ Kalb (2002, 84-122).

Böttcher, W. (2003): Bildung, Standards, Kerncurricula. Ein Versuch, einige Missverständnisse auszuräumen. In: Die Deutsche Schule, 95. Jg., H. 2 (im Druck).

Böttcher, W./ Kalb, P. E. (Hrsg.) (2002): Kerncurriculum – Was Kinder in der Grundschule lernen sollen. Beltz Pädagogik: Weinheim/ Basel.

Brügelmann, H. (1980): Experimental decision-making and responsive accountability. Expert report for "Basic Education Policies Project". OECD/ CERl: Paris.

Brügelmann, H., u. a. (1999): Was leisten unsere Schulen? Qualität und Evaluation von Unterricht in der Diskussion. Kallmeyersche Verlagsbuchhandlung: Seelze.

Carlson, D. (1979): Statewide assessment in California. In: Studies in Educational Evaluation, Vol. 5, No. 1, 55-75.

CERI (1978): National statement – United Kingdom. Vervielf. Ms. (SME/ET/78.80). OECD: Paris

Deutscher Bildungsrat (1974): Zur Förderung praxisnaher Curriculumentwicklung. Empfehlungen der Bildungskommission. Bundesdruckerei: Bonn.

Flitner, A./ Lenzen, D. (Hrsg.) (1977): Abitur-Normen gefährden die Schule. Piper: München.

House, E. R. (1975): Accountability in the U.S.A. In: Cambridge Journal of Education, Vol. 5, No. 2, 71-78.

Jacobsen, S. (1989): Merit pay incentives in teaching. In: Weis et al. (1989, 111-128).

Kay, B. W. (1979): Processes of accountability in education -- England and Wales. Working Paper for O.E.C.D. Vervielf. Ms. (SME/ET/79.41). OECD: Paris.

Murnane, R./ Cohen, D. (1986): Merit pay and the evaluation problem: Why most merit pay plans fail and a few survive. In: Harvard Educational Review, Vol. 56, No. 1, 1-17.

NAEP (2001): The Nation's report card. Fourth grade reading highlights 2000. National Center for Education Statistics. U. S. Department of Education: Washington.

O'Reilly, B. (2002): Why Edison doesn't work. In: Fortune, December 4, 2002 (www.fortune.com/fortune/articles/0,15114,395208,00.html).

Schwarz, B./ Prange, K. (Hrsg.) (1997): Schlechte Lehrer/innen. Zu einem vernachlässigten Aspekt des Lehrberufs. Beltz: Weinheim/ Basel.

Shepard, L. (1979): Purposes of Assessment. In: Studies in Educational Evaluation, Vol. 5, No. 1, 13-16.

Simon, J. (1979): What and who is APU? In: Forum, Vol. 22, No. 1, 7-11.

Stenhouse, L. (1975): An introduction to curriculum research and development. Heinemann Educational Books: London et al. (dt. Zusammenfassung in: Zeitschrift für Pädagogik, 19. Jg., H. 3, 447-452).

Terhart, E. (1997): Gute Lehrer -- schlechte Lehrer. In: Schwarz / Prange (1997, 34-85).

Weis, L. et al. (eds.) (1989) : Crisis in teaching. Perspectives on current reforms. New York.