

Jürgen Oelkers

Bildungsstandards im Gymnasium: Wovon reden wir?)*

Die amerikanische Historikerin Diane Ravitch hat vor mehr als zehn Jahren einen Vorschlag gemacht, wie sich der diffuse Ausdruck „Bildungsstandards“ begrifflich präziser fassen und differenzieren lässt. Der Vorschlag hat sich als nützlich erwiesen, was sich auch in der Schweiz im Vorfeld des Harnos-Konkordats gezeigt hat. Die bildungspolitische Karriere des Konzepts der Bildungsstandards war zunächst getragen mehr von einem bildungspolitischen Slogan als von einem genau definierten Begriff. Ravitch fasst zusammen und expliziert, was sich im bildungspolitischen Diskurs der Vereinigten Staaten seit den siebziger Jahren entwickelt hat.

Ihre dreiteilige Theorie der Standards ist zur Problembestimmung hilfreich und sorgt für die notwendigen Differenzierungen.

- Allgemein wird der Ausdruck „Standard“ mit einer doppelten Bedeutung gebraucht, als Ziel und als Mass der Zielerreichung.
- „A standard is both a *goal* (what should be done) and a *measure* of progress toward that goal (how well it was to be done)” (Ravitch 1995, S. 7).
- Wenn der Ausdruck Sinn machen soll, muss er eine bestimmte Minimalbedingung erfüllen: „A Standard is not useful or meaningful unless there is some way to measure whether it is reached“ (ebd., S. 11).

Das unterscheidet Standards von den bisherigen Zielen. Für sie war wichtig, wie sie zustande kamen und auf welchen Ebenen sie formuliert wurden. Ob sie erfüllt oder verfehlt wurden, war vielleicht eine Frage, aber es gab nur sehr begrenzt Verfahren, diese Frage auch zu beantworten. Vor allem konnte nicht gesagt werden, ob über die einzelne Klasse und die jeweiligen Noten hinaus Ziele der allgemeinen Bildung in der Breite überhaupt angestrebt und dann auch erreicht wurden.

Diane Ravitch unterscheidet drei Typen von Standards:

- Festlegung der Gehalte: *content standards*;
- Festlegung der Leistungsniveaus: *performance standards*;
- Festlegung der Ressourcen, von denen die Möglichkeiten des Lernen bestimmt werden: *opportunity-to-learn standards*.

Diese drei Typen spielen zusammen und sind untereinander verknüpft, können also nicht einzeln betrachtet werden:

„Content standards without performance standards are meaningless. Content standards define what is to be taught und learned; performance standards

*) Statement auf dem HerbstForum 2007 der Kantonsschule Zürcher Oberland am 18. September 2007 in Wetzikon.

describe how well it has been learned. Similarly, opportunity-to-learn standards cannot stand on their own, without content and performance standards” (ebd., S. 13).

Damit wird die amerikanische Diskussion bis Mitte der neunziger Jahre aus einer bestimmten Sicht zusammengefasst, wobei schon zu diesem Zeitpunkt die Frage des Testens, also die empirische Überprüfung der Standard-Erreichung, in den Mittelpunkt rückte. Seitdem ist in den Vereinigten Staaten intensiv versucht worden, für alle Unterrichtsfächer auf den verschiedenen Stufen des Schulsystems Standards zu entwickeln und die Ergebnisse zu testen.

Auch Ravitch schlägt den Gebrauch von Tests vor, die sich aber nicht einfach auf Schulwissen beziehen sollen, das wie im konventionellen Unterricht abgefragt wird, sondern mit denen die am Ende eines Lehrgangs erreichten Kompetenzen zur Problemlösung gemessen werden. Das geht auf einen Vorschlag des amerikanischen Psychologen Robert Glaser (1981) zurück und setzt ein aktives, kognitives Konzept von Lernen voraus. Schülerinnen und Schüler „empfangen“ im Unterricht nicht einfach Wissen, sondern *konstruieren* es, auf der Basis ihrer eigenen Erfahrungen, ihres Vorwissens und ihres Verstehenshorizontes. Schulen können also nicht „aus dem Nichts“ Wissen aufbauen, sondern nur die gegebenen Erfahrungen und das vorhandene Wissen erweitern (Ravitch 1995, S. 103).

Bildungsstandards sind Leistungserwartungen, also normative Grössen. Grundlegend für deren Entwicklung sind Kompetenzmodelle, die fachbezogen verstanden werden. Bildungsstandards werden mit Hilfe von Kompetenzmodellen als Leistungsdispositionen für einzelne *Fächer* oder inhaltliche *Lernbereiche* entwickelt. Die Kompetenzen beziehen sich auf die Fähigkeit zu Problemlösungen in Fächern, nicht auf isolierte psychische Dispositionen. Die Standards konkretisieren sich in Aufgabenstellungen, deren Bearbeitung anzeigt, welche Kompetenzstufe erreicht wurde. Das Ergebnis wird mit Hilfe von Leistungstests erfasst. Testaufgaben prüfen, ob und wie die Lernenden die Leistungserwartungen erfüllt haben oder nicht. Neu ist, dass dafür Stufen der Erreichbarkeit zur Verfügung stehen. Zu diesem Zweck müssen Aufgaben generiert und auf ihre Validität hin geprüft werden. Die Überprüfung allein reicht aber nicht aus. Die Ergebnisse müssen Schule und Unterricht auch erreichen, das *Feedback* wird daher zu einem Schlüsselproblem der Systementwicklung.

Der Vorrang *fachlicher* Ansprüche in den vorliegenden Konzepten der Bildungsstandards bestätigt in mehrfacher Hinsicht das gymnasiale Bildungsprofil. Es ist verblüffend, wie stark sowohl bei Ravitch als auch in dem inzwischen viel zitierten deutschen Gutachten (Klieme et. al. 2003) die *Fachlichkeit* des Lernens im Vordergrund steht. In dieser Hinsicht stellen „Standards“ für die Gymnasien in den meisten Hinsichten kein Problem dar, mit dem sie nicht längst zu tun hätten. Die heutige Diskussion verwirrt diesen Tatbestand und suggeriert, dass Bildungsstandards etwas grundlegend Neues seien. Das ist nicht der Fall (Oelkers 2004), wenigstens nicht für die Gymnasien, die sich also nicht durch bildungspolitische Suggestionen aufregen lassen sollten.

Gymnasien unterscheiden sich von anderen Schulen nicht nur durch das Fächerangebot, sondern zugleich durch die Niveauforderung. Lehrkräfte des Gymnasiums benutzen den Ausdruck „gymnasiale Standards“ wie selbstverständlich; wenn sie die Besonderheit ihres Unterrichts aufzeigen wollen. Es macht geradezu die Identität der Gymnasien aus, dass sie spezielle Anforderungen stellen können. Wenn der Ausdruck „Bildungsstandards“ gerade bei Gymnasiallehrkräften oft negativ besetzt ist und der

Ausdruck der „McDonaldisierung“ der Bildung im Raum steht, dann hängt das mit Befürchtungen zusammen, die eine bildungspolitische Diskussion und eine ökonomische Rhetorik ausgelöst haben. Der Tatbestand von Standards an sich ist davon nicht berührt.

Angesichts der Diskussion halte ich fest:

1. Jedes Gymnasium hat und vertritt *Standards*, versteht man darunter die Qualitätsanforderungen des Fachunterrichts.
2. *Gymnasiale Bildung* zeigt sich als fachliche und daneben auch überfachliche Kompetenz, die im Unterricht aufgebaut wird.
3. Die Schülerinnen und Schüler beherrschen fachliche Anforderungen unterschiedlich, die Unterschiede zeigen sich in *Leistungsstufen*.
4. *Tests* und *Noten* halten auch bisher schon Fortschritte, Stagnation und Rückschritte im Leistungsverhalten fest.

Bildungsstandards gelten politisch als Teil der Qualitätssicherung. Auch das ist nicht besonders aufregend. Alle Schweizer Gymnasien verwenden heute in der einen oder anderen Form Strategien der Qualitätssicherung. Die Strategien reichen von der Entwicklung und Kommunikation eigener Leitbilder (Mandel 2006) bis hin zur Beauftragung externer Evaluationen. Auch interne Evaluationen sind zu einem bestimmten Standard geworden ebenso die Kommunikation transparenter Leistungsanforderungen, etwa im Blick auf die Maturaarbeit (Oelkers 2007). Zu einem Standard geworden ist auch die Homepage jeder einzelnen Schule, die sozusagen den Markenkern darstellt und vielfach genutzt wird. Untersucht man die Websides, dann kann man sicher sein, die hohen Qualitätsansprüche oder eben die Standards betont zu sehen.

Die Frage ist, ob die am 14. Juni 2007, also vor wenigen Monaten, von der EDK einstimmig beschlossene *Interkantonale Vereinbarung über die Harmonisierung der obligatorischen Schule* daran etwas ändert. Zunächst bleibt festhalten: Die Gymnasien sind vom HarmoS-Konkordat ausgenommen. Das Konkordat regelt lediglich die in Zukunft elf Jahre des Schulobligatoriums, betroffen wären damit also nur die noch bestehenden Untergymnasien, nicht die Maturitätslehrgänge. Aber das macht die Frage nach den Standards der Gymnasien nicht überflüssig. Die Frage kommt auch aus einer anderen Richtung auf die Gymnasien zu, nicht nur - indirekt - von der EDK, sondern auch von den Universitäten, also den Abnehmern der gymnasialen Bildung.

Meine generelle Einschätzung sieht so aus: An den fachlichen Standards des Gymnasialunterrichts wird sich kaum etwas ändern. Das Profil des Fachunterrichts bleibt grundsätzlich erhalten, und so auch die Besonderheit des Gymnasiums.

- Neu sind Instrumente der Qualitätssicherung wie Tests oder Evaluationen,
- neu sind explizitere Kriterien der Leistungsbeurteilung
- und ebenfalls neu sind höhere Grade an Verbindlichkeit und Transparenz.

Noten erhalten einen Referenzrahmen, aber werden nicht etwa überflüssig. Ein nationales Curriculum für Gymnasien wird es jedoch ebenso wenig geben wie Methodenvorschriften oder externe Prüfungsbüros.

Eine zentrale Frage für die innere Entwicklung der Gymnasien sind die Prüfungspraxis und speziell die Notengebung. Sie bestimmt nachhaltig das Leistungsverhalten der Schülerinnen und Schüler, die den Kern der Schulqualität ausmachen. Die Schulen sollten die

Leistungen mit hoch transparenten Verfahren beurteilen und ihre internen wie externen Evaluationen darauf einstellen. Die Ergebnisse sollten verstanden werden als Anreize für die weitere Entwicklung. Wenn inhaltliche Standards als Rahmen vorgegeben sind, können Schulen sehr individuelle Wege gehen, sie zu erreichen. Was angenähert einheitlich sein muss, sind die Kriterien für die Leistungsbewertung. Mindestens in den einzelnen Fachgruppen sollten die Kriterien abgestimmt sein und offen kommuniziert werden.

Erst hier kommen Bildungsstandards im Sinne von Ravitch und der deutschen Expertise ins Spiel. Sie sind gebunden an externe Leistungstests (Maradan/Mangold 2005). Wenn die Gymnasien an HarmoS beteiligt werden, dann bedeutet das auch für sie der Einstieg in die Testkultur. Ich verwende diesen Ausdruck mit bedacht, denn das „Testen“ im Sinne von Aufgaben, die den Leistungsstand überprüfen, ist selbstverständlicher Bestandteil der gymnasialen Schulkultur. Man stelle sich einen Maturitätslehrgang ohne regelmässige Tests vor - das Gymnasium wäre ziemlich schnell am Ende.

Vermieden werden muss ein durchgehendes *Teaching the Test*, während die Vorbereitung auf Tests zum normalen Schulalltag gehört und nicht dadurch anders wird, dass Testpsychologen und Fachdidaktiker die Aufgaben konstruiert haben. Jede Klassenarbeit oder, wie man in der Schweiz bekanntlich und sehr feinsinnig sagt, jede „Probe“ ist ein Test, jede Benotung setzt eine Serie von Tests voraus und kein Urteil der Lehrkräfte kommt ohne Ergebnisse von Tests zustande. Die Ergänzung durch externe Tests kann man auch als einen Qualitätsgewinn erwarten. Amerikanische Autoren sagen das so:

„Teaching to the test and teaching the test are two different things. If a teacher is teaching to the test and doing good teaching that enhances learning, what’s wrong with that?“ (Marzano/Mayeski/Dean 2000).

Aber es gibt natürlich einen Unterschied. Mit den Bildungsstandards kommen die Aufgaben von Aussen und die Lehrkräfte sind nicht mehr allein dafür zuständig. Das gilt auch für die Auswertung. Beides, Aufgabenstellung und Auswertung, werden fachdidaktisch konstruiert und einem psychometrischen Verfahren unterworfen, zudem werden die Leistungen über den Klassendurchschnitt hinaus verglichen. Allerdings würde sich der Test nur auf einen bestimmten Zeitpunkt beziehen und hätte ansonsten keine Auswirkungen auf die Lehrfreiheit. Testsysteme dieser Art sind den Bildungssystemen im deutschen Sprachraum bislang fremd, aber werden zum Beispiel in Schweden mit einigem Erfolg praktiziert.

Ob es eigene Bildungsstandards für die Gymnasien geben wird, hängt einerseits von ihrem Nutzen ab. Sie müssen die vorhandene Qualität deutlich verbessern, wenn ihre Entwicklung Sinn machen soll. Andererseits ist diese Entwicklung teuer und es fragt sich, ob sich der Aufwand lohnt. Testfähige Aufgaben für einzelne Fächer herzustellen, und sei es nur für die PISA-Fächer, ist nicht leicht, stellt eine wiederkehrende Notwendigkeit dar, weil sich die Aufgaben mit jedem durchgeführten Test verbrauchen, und wäre als blosser Leistungsinformation vermutlich wenig wirksam. Die Gymnasien haben eine ausgesuchte Schülerschaft, die nicht mit Hilfe von Mindeststandards gezielt gefördert werden muss, damit wie im obligatorischen Bereich allgemeine Bildungsziele erreicht werden.

Das schliesst aber nicht aus, dass Tests eine grössere Rolle als bisher spielen können. Für die Gymnasien ist das allerdings noch kaum angedacht, geschweige denn auch nur in einem Vorstadium der Entwicklung. Das zeigt ein abschliessender Blick auf HarmoS. Im Projekt der Bildungsstandards sind Tests bislang vorgesehen am Ende der zweiten, der sechsten und der neunten Klasse. Das Ziel ist, den Leistungsstand aller Schüler an drei

Zeitpunkten zu messen und an zwei Zeitpunkten Leistungsrückstände korrigieren zu können. Der Zeitpunkt Ende der neunten Klasse ist gedacht als Abschlusstest, wie es heute mit PISA gemacht wird. Soll auch hier noch Förderung greifen, müsste entweder auf den Abschlusstest verzichtet und auf Projekte wie „Stellwerk“, also einen kantonalen Test Mitte der achten Klasse, der zur Lernstandserhebung dient, zurückgegriffen werden oder aber der Abschlusstest erhalte eine Funktion für den Schulübergang, was bislang aber nicht vorgesehen ist.

Gelegentlich ist der politischen Diskussion gefordert worden, diese Art Qualitätssicherung auch auf die Sekundarstufe II auszudehnen und entsprechende Tests in der zwölften Klasse durchzuführen. Damit sind Chancen verbunden, aber auch die Risiken sollten deutlich benannt werden. Der zentrale Vorteil ist die Entwicklung einer Förderkultur, die datenbasiert ist. Ein zentrales Ergebnis beider PISA-Studien geht dahin, dass Länder besser dastehen, die ihre Schülerinnen und Schüler gezielt fördern können. „Gezielt“ meint, die Förderung kann sich auf mehr beziehen als die Einschätzung der je beteiligten Lehrkräfte. Wenn die Testergebnisse vorliegen, wird sichtbar, wo genau der Förderungsbedarf liegt. Die Nachfrage nach Testsystemen im Volksschulbereich, etwa der „Klassenscockpit“ aus St. Gallen oder das aargauische „Check Five“, zeigt, dass diese Chance von den Lehrkräften genutzt wird und die Instrumente im Unterricht Verwendung finden (Tresch 2007).

Die Gefahr besteht darin, den Unterricht ganz oder weitgehend auf Test einzustellen. Ein „Teaching-the-Test“ wie in amerikanischen Schulen muss vermieden werden. Ein Problem ist auch, dass die Tests technisch gut sein müssen, also nicht einfach auf den Idealschüler zugeschnitten sind und die Effekte des Unterrichts von herkunftsbedingten Einflüssen unterscheiden können. Eine zentrale Frage ist schliesslich, ob die Tests, wenn sie schon da sind, zur schulischen Selektion genutzt werden oder wirklich nur der Förderung dienen. Schliesslich sollte noch beachtet werden, dass sich die Lehrkräfte und die Schüler kognitiv und vor allem motivational darauf ein, dass und wie getestet wird (O’Neil/Abedi/Miyoshi/Mastergeorge 2004).¹ Tests sind nicht einfach nur Leistungsmessungen, sondern selbst Lernobjekt, das von den Schülern mehr oder weniger ernst genommen wird. Sie lernen, sich regelmässig auf Tests einzustellen, so wie sie gelernt haben, regelmässig Klassenarbeiten zu schreiben und am Ende eine Durchschnittsnote entgegen zu nehmen.

Ein weiteres Problem ist, wie Rückmeldesysteme organisiert werden sollen. Die Resultate von Evaluationen bleiben unbeachtet, wenn sie nicht die Ebene des Unterrichts erreichen (Oelkers/Reusser 2007). Zudem sind die Tests mit einem hohen Aufwand verbunden, und es wäre Verschwendung, wenn mit den Daten nicht mehr angefangen wird als dass sie die Grundlage abgeben für nationale Bildungsberichte, die politisch relevant sind und aber nicht gelesen werden. Grössere Versuche im Volksschulbereich, wie das erwähnte „Check Five“, zeigen, dass die Lehrkräfte die Daten nutzen, wenn sie Vorteile davon haben. Die entsprechenden Strategien müssen aufgebaut und entwickelt werden, blosser Appelle sind wie Bildungsberichte wirkungslos.

Vorteile von testbasierten Leistungsmessungen wären etwa:

- Objektivierung der Notengebung

¹ Vermutet wird ein „senioritis-effect“, also eine Abnahme der Motivation am Ende der Schulzeit (O’Neil/Abedi/Miyoshi/Mastergeorge 2004, S. 199), was zum Beispiel für die PISA-Studie, die Leistungen am Ende der Schulzeit testet, erhebliche Konsequenzen hätte.

- Bessere Absicherung der Kommunikation der Noten gegenüber Eltern und Schülern
- Vergleich mit anderen Klassen zur Ermittlung des eigenen Standorts
- Bestimmung des Leistungsstandes zur gezielten Förderung einzelner Schülerinnen und Schüler
- Absicherung der erreichten oder zu erreichenden Unterrichtsqualität.

„Leistungsmessung“ klingt technisch, also fern dem gymnasialen Bildungsideal. Aber gemessen wird nur, was der Unterricht aufbaut, und die Tests erfassen das, was an Bildung hervorgebracht wird. Man kann nur dann von einer Leistungsschule sprechen, wenn die Leistungen zu überprüfbar Resultaten führen. Das im Kern ist unstrittig, anders würden die Gymnasien nicht so grossen Wert auf ihr Prüfungswesen legen. Die Tests sind einfach nur ein Mittel zur Objektivierung der Leistungsbeurteilung. Man sollte also zwischen „Test“ und „gymnasialer Bildung“ keinen Keil treiben und nicht von einer Unverträglichkeitserklärung ausgehen. Vielmehr gilt es, die Chancen der Entwicklung zu nutzen und die Risiken zu minimieren.

Literatur

- EDK: Interkantonale Vereinbarung über die Harmonisierung der obligatorischen Schule. HarmoS-Konkordat. Bericht zur Vernehmlassung. Bern: EDK 2006.
- Glaser, R.: The Future of Testing. A Research Agenda for Cognitive Psychology and Psychometrics. In: American Psychologist, Vol. 36 (September 1981), S. 923-936.
- Klieme, E. et al.: Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Herausgegeben vom Bundesministerium für Bildung und Forschung. Bonn: BMBF 2003.
- Mandel, S.: Wirksamkeit der Leitbilder von Schweizer Gymnasien. Das Leitbild als Mittel zur Qualitätssicherung, Qualitätsentwicklung und Qualitätsevaluation. Bern et. al.: Peter Lang Verlag 2006. (= Explorationen. Studien zur Erziehungswissenschaft, hrsg. v. J. Oelkers, Bd. 52)
- Maradan, O./Mangold, M.: Bildungsstandards in der Schweiz. Das Projekt HarmoS. In: ph/akzenteHeft 2 (2005), S. 3-7.
- Marzano, R.J./Mayeski, F./Dean, C.B.: Research into Practice Series: Implementing Standards in the Classroom. Aurora, CO: Mid-continent Research for Education and Learning (McREL) 2000.
- Oelkers, J.: Zum Problem von Standards aus historischer Sicht. In: Neue Sammlung 44 (2004), S. 179-200.
- Oelkers, J.: Expertise Mittelschulen. Zuhanden der Bildungsdirektion des Kantons Zürich. Bern: h.e.p. Verlag 2007.
- Oelkers, J./Reusser, K.: Expertise „Qualität entwickeln - Standards sichern - mit Differenz umgehen.“ Berlin: BMBF 2007.
- O’Neil, H.F./Abedi, J./Miyoshi, J./Mastergeorge, A.: Monetary Incentives for Low-Stakes Tests. In: Th. Fitzner Hrsg.): Bildungsstandards. Internationale Erfahrungen - Schulentwicklung - Bildungsreform. Bad Boll: Evangelische Akademie 2004, S. 164-201.
- Ravitch, D.: National Standards in American Education: A Citizen’s Guide. Washington, DC: Brookings Institution 1995.
- Tresch, S.: Potenzial Leistungstest. Wie Lehrerinnen und Lehrer Ergebnismeldungen zur Sicherung und Steigerung ihrer Unterrichtsqualität nutzen. Bern: h.e.p. Verlag 2007.

