

Jürgen Oelkers

### *Bildungsstandards und individuelle Förderung: Ein Widerspruch?\*)*

Das Thema meines Vortrages scheint mit einer Paradoxie verbunden zu sein: „Bildungsstandards“ gelten für *alle* Schülerinnen und Schüler gleich, „fördern“ aber kann man immer nur *einzelne*. Wer beides auf einander bezieht, scheint in die Sackgasse einer Paradoxie zu geraten. Die eine Seite schliesst aus, was die andere fordert. Die individuelle Förderung setzt Rücksicht auf das gegebene Potential der Kinder, die Unterschiede im Lerntempo und den je erreichten Lernstand voraus. Standards, so scheint es, egalisieren und geben eine Leistungsnorm vor, die gerade nicht individualisiert werden kann. Besteht daher zwischen beiden Seiten ein Widerspruch oder gar eine unauflösbare Paradoxie?

Über „Bildungsstandards“ wird in pädagogischen Kreisen und auch in der grösseren Öffentlichkeit seit einigen Jahren intensiv diskutiert, oft kontrovers und selten mit klaren Begriffen. Der Ausdruck *standards of education* stammt aus den Vereinigten Staaten und ist eine unglückliche oder missverständliche Wortwahl, weil jeder sofort an Industrienormen denkt und dann leicht der Verdacht geäussert werden kann, dass solchen Normen jetzt auch noch unsere Kinder unterworfen werden sollen. Das Wort der „McDonaldisierung“ der Bildung hat die Runde gemacht, ganz so trage sich die Bildungspolitik mit der finsternen Absicht herum, den Unterricht in normierte Häppchen zu zerlegen, die alle gleich schmecken sollen und auf den individuellen Geschmack nicht reagieren müssen.

Dieses Bild liegt nahe, weil der Ausdruck „Standards“ in aller Regel so gebraucht wird, als feste Norm für alle. Gleichwohl ist das Bild schief und nur für polemische Zwecke zu gebrauchen. Die Schulen werden nicht zu Fast-Food-Ketten, die Schülerinnen und Schüler erhalten keine Einheitsnahrung und die Lehrkräfte werden nicht zu Hamburgerbratern umgeschult. Das sage ich auch aus dem Grunde, weil diese Ketten bei den Schülern meistens sehr beliebt sind. Aber das erlaubt keine Analogie zum Unterricht, und das ist auch gut so. Bildung erwirbt man nicht unter Verzicht auf Anstrengungen, Bildung setzt persönliche Leistungen voraus und sie ist weder „fast“ noch „cool.“

Dies vorausgesetzt, gliedert sich mein Vortrag in drei Teile: Ich werde zunächst etwas zum Begriff „Bildungsstandards“ sagen, der vor allem zur Beschreibung von Leistungserwartungen verwendet wird und Stufen der Zielerreichung voraussetzt (1). In einem zweiten Schritt gehe ich auf das gesamtschweizerische Projekt HarmoS ein. Ein Teil dieses Projekts befasst sich mit Bildungsstandards und Leistungstests (2). Abschliessend komme ich die Frage des Förderns zurück und zeige, wie sowohl die Bildungsstandards als auch die Leistungstest die Praxis des Förderns bereichern und zum Teil auch auf eine neue Basis stellen können (3).

---

\*) Vortrag vor dem Verband der Schulpflege-Präsidentinnen und -Präsidenten Luzern (VSPL) am 22. Oktober 2007 in Emmenbrücke.

Meine These geht dahin, dass nur dann mit Aussicht auf Erfolg gefördert werden kann, wenn klar ist, *wohin* gefördert werden soll. Das „wohin“ bestimmen heute die Lehrkräfte je für sich, das gilt auch für den Umfang der Förderung, die Dauer und die Reichweite. Auch *welche* Schülerinnen und Schüler gefördert werden sollen, bestimmen allein die Lehrkräfte. Standards geben nun verbindlicher vor, wohin gefördert werden soll und was erreicht werden muss, damit bis zu einer bestimmten Linie niemand zurück bleibt. *No Child Left Behind* lautet ein Slogan der internationalen Bildungspolitik, dieser Slogan lässt sich nur mit Bildungsstandards umsetzen, was aber sind Bildungsstandards?

### 1. Zum Konzept der Bildungsstandards

Ein Kernstück von HarmoS sind Bildungsstandards. Bildungsstandards werden derzeit entwickelt, aber es gibt es noch kein Produkt, das man in Augenschein nehmen könnte. Genauer müsste ich sagen, es gibt noch keine Praxis, die unterscheidbar wäre. Das erklärt zum Teil die Diskussion, die sich nicht auf Erfahrungswerte beziehen kann und daher gelegentlich aufgeregt ist, weil die Befürchtungen die Runde machen. Die Lehrkräfte sehen die Notwendigkeit, Standards zu entwickeln, oft gar nicht ein, sie sehen ihre Freiheiten bedroht und reagieren so lange mit Abwehr oder mindestens mit starken Reserven, wie sie die Vorteile des neuen Systems noch nicht sehen. Diese Reaktion ist verständlich, denn wo genau liegt die Innovation und was genau wäre der erwartbare Gewinn für die Praxis?

In gewisser Hinsicht gibt es gar keine Innovation. Standards sind in bestimmter Hinsicht nichts Neues, legt man eine sehr allgemeinen Begriff zugrunde, dann *bestehen* Schulen geradezu aus Standards (Oelkers 2004).

- Die Notenskala ist ebenso ein „Standard“ wie der Gebrauch von Lehrmitteln oder die Lehrer:Schüler-Relation.
- Ohne das gemeinsame Instrument der Notenskala könnte es in einer Schule keine Leistungsbeurteilung geben,
- Lehrmittel standardisieren die Gehalte des Unterrichts
- und die Relation zwischen Lehrern und Schülern ist sogar ein statistischer Standard.
- Was ist dann also neu an den Bildungsstandards?

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, wie sich auch in der Schweiz im Vorfeld des HarmoS-Konkordats gezeigt hat. Die Karriere der „standards of education“ war zunächst mehr von einem bildungspolitischen Slogan getragen 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 Bildungszielen. 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 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 die Testaufgaben generiert und auf ihre Validität hin geprüft werden. Das geschieht zur Zeit in Konsortien, die aus Vertretern der verschiedenen Fachdidaktiken sowie Praktikerinnen und Praktikern zusammengesetzt sind. Die Resultate dieser aufwändigen Entwicklungsarbeit werden vermutlich erst Ende 2008 vorliegen. Ein zentrales Problem der Leistungstests ist, dass die Resultate die Lehrkräfte erreichen müssen, wenn sie verwendet werden sollen. Die Messung allein reicht nicht aus, vielmehr muss es eine wirksame Rückkoppelung geben. Generell wird das Feedback zu einem Schlüsselproblem der Systementwicklung.

Bildungsstandards gelten politisch als Teil der Qualitätssicherung, die in anderen Bereichen bereits weit fortgeschritten ist. Fast alle Schweizer Schulen 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 einer Schule betont zu sehen.

In dieser Hinsicht sind Standards gar nicht strittig. Die Schulen übernehmen einfach und passen auf ihre Verhältnisse hin an, was sie bei anderen Schulen sehen. Das ist der normale Vorgang der Schulentwicklung, der völlig unspektakulär daher kommt und auch nicht mit grossen Diskussionen verbunden ist. Davon zu unterscheiden ist die nationale Strategie der Qualitätssicherung. Diese Strategie betrifft erstmalig den „Bildungsraum“ Schweiz und nicht mehr nur die einzelnen Kantone. Das ist neu, merkwürdigerweise aber auch nicht strittig. Die grossen Diskussionen sind bislang ausgeblieben, was die Beobachter nicht irritiert hat. Die Rede ist vom Schweizer Projekt HarmoS, das mich in einem zweiten Schritt beschäftigen wird.

## 2. Das Schweizer Projekt HarmoS

Die *Interkantonale Vereinbarung über die Harmonisierung der obligatorischen Schule*, abgekürzt HarmoS (mit dem grossen „S“ am Ende), ist am 14. Juni 2007 in Bern unterzeichnet worden. Das wäre der Öffentlichkeit beinahe entgangen. Wer die Geschichte der Schweizerischen Volksschule kennt und die dabei nicht nur prägenden, sondern oft auch rivalisierenden Eigenheiten der Kantone vor Augen hat, ist vom dem Konkordat überrascht. „Harmonisierung“ klingt fast wie *Zentralisierung*, was sonst immer ein Grund ist, Abstimmungen zu verlieren. Ist die Schweiz plötzlich harmoniesüchtig geworden, und dies ausgerechnet im Bildungsbereich? Oder ist das Projekt missverstanden, wenn man es als grosse Reform versteht?

Das Projekt ist im Jahre 2003 lanciert worden, nicht ganz freiwillig, wie man weiss; die Vernehmlassung vor einem Jahr erbrachte mit Ausnahme der Gestaltung der Sekundarstufe I eine grosse bis überwältigende Zustimmung, nunmehr ist das Konkordat von der EDK verabschiedet, und zwar einstimmig. Die öffentliche Aufmerksamkeit hielt sich in Grenzen, was mit der Zustimmung zu tun haben kann, aber auch mit dem *courant normal* der eidgenössischen Bildungspolitik. Wenigstens ist erstaunlich, wie vergleichsweise gering die Resonanz auf das Konkordat war, wenn man vor Augen hat, wie stark der Eingriff in das bisherige System sein wird.

Das griechische Wort *harmonia* lässt sich mit „Zusammenfügung“ oder „Ebenmass“ übersetzen. Die Pythagoreer<sup>1</sup> verstanden bekanntlich die Harmonie der Welt als *schöne Proportion*, veranschaulicht in den Sphären von Mathematik und Musik. In der realen Welt geht es etwas profaner zu. HarmoS wird nicht zu einer schönen Proportion führen und auch kein Ebenmass herstellen, die Konkurrenz zwischen den Kantonen bleibt bestehen, heute sind Schulen Standortfaktoren, die Vorteile im Wettbewerb darstellen, wenn sie gut sind. Harmonie wird also nicht eintreten, wohl aber kommt es zu einer Strukturangleichung, die nach den bisherigen Massstäben der Schweizer Schulgeschichte sehr weit geht, so weit, dass manche Traditionalisten schon um die Identität ihres Kantons bangen. Aber das kann selbst HarmoS nicht bewirken.

Das neue Konkordat ersetzt das alte aus dem Jahre 1970. Um den Wandel der Bildungspolitik zu verstehen, lohnt ein Blick in das alte Konkordat, das seinerzeit vom Präsidenten der EDK, dem Zuger Erziehungsdirektor und späteren Bundesrat Hans Hürlimann, sowie dem Generalsekretär Emil Egger unterschrieben wurde. Das Konkordat<sup>2</sup> unterscheidet fein zwischen „Verpflichtungen“ und „Empfehlungen.“ Die Verpflichtungen beziehen sich auf die Angleichung der kantonalen Schulgesetze in vier Punkten: Festlegung des Schuleintrittsalter auf das vollendete 6. Altersjahr, Dauer der Schulpflicht, Dauer der ordentlichen Ausbildungszeit vom Schuleintritt bis zur Maturitätsprüfung sowie Beginn des Schuljahres zwischen Mitte August und Mitte Oktober.

Moritz Arnet (2000, S. 24ff.) hat beschrieben, welche Kämpfe diese bescheidenen Regelungen seinerzeit auslösten. Zankapfel war der Schuljahresbeginn. Wer erinnert sich noch daran, dass der Kanton Zürich dem Konkordat wohl beitrug, die Stimmbürgerschaft auch zustimmte, aber am 4. Juni 1972 dann einer Initiative der Gegner folgte, die den Frühjahrschulbeginn beibehalten wollten. Der Kanton Bern hatte seinen Beitritt zum Konkordat von der Abstimmung über den Schuljahresbeginn abhängig gemacht, die ebenfalls zugunsten der Gegner entschieden wurde. Zwei Plenarversammlungen der EDK im Juni und November 1972 dienten der Beurteilung dieser nunmehr ziemlich verfahrenen Lage.

Arnet beschreibt das so:

„Der Misserfolg bei der Schuljahresbeginn-Frage blockierte ... das Durchsetzungsvermögen vorab in Strukturfragen, versetzte die Kantone auf die Versagerbank und liess eine Bundeslösung befürchten. Es machte sich die Sorge breit, die koordinationswillige Westschweiz wende sich von der zerstrittenen Deutschschweiz ab. Die Erziehungsdirektoren machten sich dennoch Mut und beschworen den Schulterschluss“ (ebd., S. 25).

<sup>1</sup> Benannt nach der Schule von Pythagoras von Samos (zw. 580 u. 572 - zw. 500 und 490 v. Chr.).

<sup>2</sup> Grundlage des Konkordats waren Empfehlungen der EDK aus dem Jahre 1967.

Die Bildungsschweiz brach nicht auseinander, für die Schuljahrfrage wurden „pragmatische Übergangslösungen“ gefunden (ebd., S. 26) und eine Bundeslösung konnte weiterhin befürchtet werden. Das Konkordat von 1970 liess die Schulstruktur unangetastet und stellte auch keinen inhaltlichen Eingriff dar. Bei den Empfehlungen ging man weiter, aber das war kein Kunststück, weil die Verbindlichkeit offen blieb. Immerhin sollten die Konkordatskantone zuhänden aller Kantone Empfehlungen insbesondere für folgende Bereiche ausarbeiten:

- Rahmenlehrpläne,
- gemeinsame Lehrmittel,
- Sicherstellung des freien Übertritts zwischen gleichwertigen Schulen,
- Übertritt in die gegliederten Oberstufen,
- Anerkennung von Examensabschlüssen und Diplomen, die in gleichwertigen Ausbildungsgängen erworben wurden,
- einheitliche Bezeichnung der gleichen Schulstufen und gleichen Schultypen,
- gleichwertige Lehrerausbildung (Konkordat 1970)

Zur Ehrenrettung des alten Weges, möglichst *wenig* für verbindlich zu erklären, muss gesagt werden, dass die meisten dieser Punkte durch pragmatische Annäherung erreicht wurden. Die Lehrerinnen- und Lehrerbildung ist tertialisiert worden, die Lehrmittel wurden regional angenähert, heute kooperieren sogar die Lehrmittelverlage, Übertritte und Anerkennungen sind freizügig geregelt, die Bezeichnungen für Stufen und Typen sind weitgehend harmonisiert.

Der neue Staatsvertrag geht wesentlich weiter und greift in die Struktur der Volksschule ein. Deutlich heisst es in der Zweckbestimmung:

- Die Vereinbarungskantone harmonisieren die obligatorische Schule im Blick auf „die Ziele des Unterrichts und die Schulstrukturen.“
- Ausserdem verpflichten sich die Kantone, „die Qualität und Durchlässigkeit des Schulsystems durch gemeinsame Steuerungsinstrumente (zu) entwickeln und (zu) sichern“ (Interkantonale Vereinbarung Art. 1).

Beide Bestimmungen wären 1970 nicht nur unmöglich, sondern auch undenkbar gewesen. Gemeinsame Ziele, die irgendwie verbindlich gewesen wären, gab es nicht, Schulstrukturen festzulegen, war Sache allein der Kantone, und „Qualitätssicherung,“ heute in aller Munde, war ein unbekanntes Wort. Der erste amerikanische Aufsatz, der eine Systemsteuerung durch evidenzbasiertes Feedback vorschlug, erschien erst vier Jahre später (Levin 1974).

Hinter dem Projekt HarmoS steht eine weitgehende Umstellung der Bildungspolitik hin zu dem, was neudeutsch „Outputsteuerung“ genannt wird. Mit diesem ziemlich sperrigen Wort verbindet sich ein simples Modell, das den *Eingang* und den *Ausgang* eines Prozesses beschreibt. Das Bildungssystem lässt sich mit diesem Modell so abbilden:

- Bislang lag der Augenmerk der Politik auf dem *Input*, also dem, was in das System eingegeben wurde, nämlich Lehrpläne,

Lehrmittel, ausgebildete Lehrkräfte, Schülerinnen und Schüler, Ressourcen und gute Absichten.

- In Zukunft sollen die Resultate oder der *Output* eine zentrale Rolle spielen.
- Mit Resultaten sind vor allem die durch Unterricht hervorgebrachten *Kompetenzen* der Schülerinnen und Schüler gemeint.

Das Modell ist ebenso einfach wie erfolgreich. Eine deutsche Expertise spricht sogar von einem „Paradigmenwechsel“ (Klieme et. al. 2003) in der Bildungspolitik. Das ist, wenn man an Thomas Kuhn denkt, etwas sehr hoch gegriffen. Aber im internationalen Vergleich hat sich die Steuerung der Bildungssysteme tatsächlich seit etwa zwanzig Jahren sehr grundlegend verändert (Oelkers/Reusser 2007). Die EDK hat also keine Anomalie in die Welt gesetzt.

Es geht bei Harmos bekanntlich nicht nur um Bildungsstandards. Die Vereinbarung sieht die Einführung einer obligatorischen Basisstufe vor, mit der sich die Schulpflicht schweizweit um zwei Jahre nach unten verlängern wird. Sie soll in Zukunft mit dem vollendeten vierten Lebensjahr beginnen. Die Primarschule umschliesst fortan eine Vorschule oder Eingangsstufe. Der Kindergarten wird mit der Schule fusioniert, eigentlich ist das eine bildungspolitische Sensation, die wiederum von der Öffentlichkeit unaufgeregt zur Kenntnis genommen wurde. Vielleicht ist der Grund für diese Gelassenheit, dass Vorschulen eine europäische Norm darstellen, wengleich in aller Regel nicht als Obligatorium. Eine Anomalie ist die Eingangsstufe jedenfalls auch nicht.

De facto führt das zur Etablierung einer achtjährigen kantonalen Gesamtschule und zur Weiterführung der dreijährigen, gegliederten Sekundarstufe. Der organisatorischen Harmonisierung der kantonalen Systeme soll auch eine inhaltliche folgen. Die Grundbildung wird in Zukunft auf einheitlich fünf Lernbereiche oder Fachgruppen bezogen, es soll sprachregionale Lehrpläne geben, zudem Blockzeiten sowie Tagesstrukturen und eben landesweit verbindliche Standards. Der grösste Schritt ist die Basisstufe, wenn sie denn so einheitlich kommt, wie das vorgesehen ist. Was heute primär diskutiert wird, die Einführung von Bildungsstandards, ist Teil eines grösseren und sehr ambitionierten Projekts der Systementwicklung. Angesichts der kantonalen Bildungshoheit ist das tatsächlich ein grosser Schritt, der alles übertrifft, was in der Vergangenheit an Abstimmung zwischen den Kantonen erreicht wurde.

Im Bericht zur Vernehmlassung zu HarmoS, der zu Beginn des Jahres 2006 vorlag, wird gesagt, dass mit den Bildungsstandards *nicht* das Ziel verfolgt werde, auf der Basis einer nationalen Vorstellung die Bildungsprozess zu „vereinheitlichen“ (EDK 2006, S. 24). Vielmehr soll mit dem Instrument der Bildungsstandards versucht werden, die *erwarteten* Leistungsergebnisse so zu beschreiben, dass diese normative Beschreibung auf nationaler Ebene „verbindlichen Charakter“ annehmen kann (ebd.). Wie die Unterrichtsziele zu erreichen sind, wird nicht vorgegeben. Dasselbe gilt für die Aufteilung der Unterrichtszeit, den Einsatz der Ressourcen und alle weiteren Fragen der Schulorganisation. Hier bleibt die Zuständigkeit der Kantone, der Gemeinden und der einzelnen Schulen erhalten. Harmonisiert werden die *Leistungserwartungen*, nicht die Ausführungen. Insofern basieren manche Befürchtungen auf Unkenntnis.

Entwickelt werden *Mindeststandards*. Der Ausdruck beschreibt, was auf einer Skala von Kompetenzstufen in jedem Fach zu einem bestimmten Zeitpunkt von allen Schülerinnen

und Schülern erreicht werden muss. Wer einen solchen Standard nicht erreicht, soll oder muss gezielt gefördert werden. Die Kantone können die Mindeststandards höher oder tiefer ansetzen. Es handelt sich also nicht um eine Nivellierung nach unten. Die Standards werden in erster Linie einen Einfluss auf die Lehrplanarbeit und die Lehrmittel haben. Die im Projekt entwickelten Referenzrahmen erlauben eine bessere Strukturierung der Fächer, mit den skalierten Kompetenzniveaus wird es zudem möglich, eine Progression bei den fachbezogenen Lernzielen festzulegen. Die offiziellen Lehrmittel stützen sich in Zukunft auf Standard-kompatible Rahmenlehrpläne, so dass sich die Lehrpersonen unausweichlich auf Standards beziehen müssen.

Mit dem HarmoS-Projekt sind drei grössere bildungspolitische Festlegungen verbunden:

- Die Kantone entwickeln Unterstützungs- und Differenzierungsmaßnahmen für diejenigen Schülerinnen und Schüler, die die Basiskompetenzen *nicht* erreichen.
- Die Mindeststandards dienen *nicht* der schulischen Selektion. Durch standardisierte Tests im heutigen 2., 6. und 9. Schuljahr wird die Schülerbeurteilung verbessert.
- Die Wirksamkeit des schweizerischen Bildungssystems im Bereich der obligatorischen Schule wird am Erreichen der Standards gemessen, darauf bezieht sich das nationale Bildungsmonitoring (ebd., S. 25/26).

Das HarmoS-Konkordat ist verbunden mit einem Einstieg in die Testkultur, auch das muss deutlich gesagt werden. Regelmässige Leistungstests sind den Bildungssystemen im deutschen Sprachraum bislang fremd ist, aber sie werden in liberalen Systemen wie Schweden mit Erfolg praktiziert. Oft unterstellt die Diskussion Szenarien wie die, dass der ganze Unterricht nur noch auf den Test hin ausgerichtet wird und die individuelle Förderung dabei auf der Strecke bleibt. Aber bei nur drei Messzeitpunkten ist das eine stark übertriebene Befürchtung. Selbst wenn die Lehrkräfte ihre Schülerinnen und Schüler gezielt auf die Tests vorbereiten, sie werden überwiegend nicht das tun, was in der amerikanischen Literatur „teaching the test“ genannt wird.

Unter curricularen „Standards“ sind wie gesagt Aufgaben zu verstehen, die den Leistungsstand bestimmter Fächer in einem Kernbereich testen und die für diesen Zweck fortlaufend geeicht werden müssen. Es handelt sich um die Fachbereiche

- Erstsprache
- Fremdsprachen
- Mathematik
- Naturwissenschaften.

Es wird also zunächst *Standard-Fächer* und *Nicht-Standard-Fächer* geben, weil der Aufwand zu gross wäre, für alle Fächer gleichzeitig Bildungsstandards zu entwickeln. Darin liegt eine gewisse Gefahr, die bei der weiteren Entwicklung im Auge behalten werden muss. Aus der Tatsache, dass mit bestimmten Fächern begonnen wird, darf nicht geschlossen werden, dass es fortan eine Zwei-Klassen-Schulen geben wird, die aus „wichtigen“ Fächern besteht, nämlich solchen, die getestet werden und „unwichtigen“, wo das nicht der Fall ist. Das Konkordat lässt im Übrigen die Fächerfrage offen.

Standards sind, nochmals gesagt, keine neuen Erfindungen. Jede Schule verfügt über bestimmte Ressourcen und curriculare Standards sind einfach Lerninhalte, wie sie in Lehrmitteln und Lehrplänen vorhanden sind. Sie werden von den fachdidaktischen Konsortien zusammen mit den Lehrkräften so präzisiert, dass sie getestet werden können. Das Neue an der Entwicklung ist also die Testreihe bestimmter Aufgaben sowie die nationale Harmonisierung, wobei auch gesagt werden muss, dass viele curriculare Gehalte bereits heute zwischen den Kantonen angenähert sind. Anders hätte zum Beispiel der PISA-Test gar nicht konstruiert werden können. Zur Bestimmung des Leistungsstandes sieht die interkantonale Vereinbarung auch vor, dass die Schülerinnen und Schüler Portfolios führen, in und mit denen sie ihre Lernfortschritte dokumentieren. Sie werden also nicht nur getestet (Interkantonale Vereinbarung Art. 9).

Generell gesagt: Bildungsstandards sind keine bürokratischen Festlegungen, sondern stellen verbindliche Leistungserwartungen dar, die für den gesamten Bildungsraum Schweiz gelten sollen (HARMOS 2004). Ohne solche Erwartungen liesse sich von Outputsteuerung gar nicht reden. Sie verlangt externe Bezugsgrössen und gemeinsame Massstäbe, sonst könnte es die berühmte „Orientierung am Resultat“ nicht geben. Das Resultat muss mehr sein, als sich Noten und Zeugnisse beschreiben lässt, anders kann es keinen Vergleich geben. Damit sind Folgen für das schulische Ziel des Förderns verbunden, die mich abschliessend beschäftigen werden.

Der letzte Teil meines Vortrages setzt sich mit der Praxis des Testens auseinander und versucht eine Einschätzung, ob und wie sich regelmässige Tests auf das Fördern der Schülerinnen und Schüler auswirkt. Viel wissen wir darüber in der Schweiz nicht, aber erste und ganz ermutigende Befunde gibt es. Ich beginne mit einem Beispiel ausserhalb der Schweiz. Das Beispiel England zeigt, wie stark sich Schulen durch Bildungsstandards verändern können und dass für die Bestimmung der Grenzen die Bildungspolitik und letztlich der öffentliche Konsens verantwortlich sind.

### 3. Fördern und Testen

England ist das Land mit dem am weitesten entwickelten System von Bildungsstandards. Hier gibt es eine Diskussion, die darauf hinweist, dass inzwischen so etwas wie eine *Audit-Society* (Power 1997) entstanden sei, eine Gesellschaft, die auf allen Ebenen und ständig Rechenschaft (audit) abverlangt, ohne davon besonders viel Nutzen zu haben. „Audits work because organisations have literally been made auditable“ (ebd., S. 91). Rechenschaft verkommt zum Ritual, einfach weil die Umwelten der Organisationen darauf eingerichtet sind und formal Bilanzierungen abverlangen, die intern leer laufen (ebd.). Man macht Leitbilder, evaluiert sich selbst und lässt sich evaluieren, aber was sich damit ändert, lässt sich auch Mikrobereich kaum wahrnehmen. Was die Wirksamkeit prüfen soll, bleibt selbst ohne Effekt.

In England hat die staatliche Bildungspolitik seit dem Education Reform Act von 1988 die Praxis der Schulen mit vergleichsweise harten Gesetzen und engen Vorschriften zu reglementieren versucht. Zum ersten Male wurde ein nationales Curriculum eingeführt, die Inhalte aller Fächer auf allen Schulstufen werden mit Bildungsstandards und Kompetenzstufen beschrieben und an mindestens vier Stellen wird der Leistungsstand der Schülerinnen und Schüler getestet. Die Schulen werden regelmässig extern evaluiert, die Inspektorenberichte werden im Internet publiziert. Die Leistungen der Schulen werden mit

Hilfe von Rankings öffentlich gemacht, die einzelnen Schulen haben Kennziffern und Leistungsdaten, die Schulwahl ist frei gestellt, was sich so entwickelt hat, ist eine Art Marktssystem.

Die ökonomische Rhetorik löst den Verdacht aus, es handle sich dabei lediglich um „a messianic belief“ (Wolf 2000, S. 29), mit dem letztlich nur Unheil angerichtet werde. Diese Kritik wird häufig vertreten, einhergehend mit dem Verdacht, die Einführung von Bildungsstandards hätte ein verengtes Curriculum zur Folge, das von der Wirtschaft angetrieben sei (economy driven) und zur Uniformierung des Unterrichts zwingt. Dadurch sei die pädagogische Freiheit gefährdet und werde der Bürokratisierung Vorschub geleistet (Miller 1996 und diverse Andere). Aber das deckt sich nicht mit den Erfahrungen und stellt keine Erfassung der komplexen Wirklichkeit dar, die auch in England keineswegs nur negativ ist. Die Lehrgewerkschaften, etwa, sind nicht gegen Standards, sondern nur gegen Rankings, und zum alten System der „kindzentrierten“ Primarschule will niemand wirklich zurück.

Eine Kritik am Weissbuch der englischen Regierung von 1997, das eine erste positive Bilanz der Implementation von Bildungsstandards zieht,<sup>3</sup> geht denn auch eher dahin, die technischen Probleme des Testens ernst zu nehmen, besonders des Testens über längere Zeiträume, und dabei die Relativität der Aussagen vor Augen zu haben. „There is no absolute criterion which determine what pupils will be able to do“ (Plewis/Goldstein 1998, S. 19). Das Lernen reagiert nicht einfach auf gesetzte „benchmarks“ (ebd.), und es dürfte schwer sein, die Ziele ohne deren „shifting“ über die Zeit konstant zu halten (ebd.). Aber das sind keine grundsätzlichen Einwände, wie sie etwa in Deutschland oder Österreich regelmässig vorgebracht werden, auch weil überhaupt keine Praxis besteht.

Die Gefahr besteht darin, den Unterricht ganz oder weitgehend auf den Test einzustellen. Dieses *Teaching the Test* muss vermieden werden, 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 natürlich haben nationale Leistungstests auch Nachteile. Eine Richtung in der amerikanischen Kritik verweist auf die Gefahr, bei hohen Investitionen in die Beschreibung der *Ergebnisse* die Inhalte zu vernachlässigen und die Prozesse des Lernens einzig auf das Ergebnis zuzuschneiden. Beachtet werden dann nur noch Testresultate, einschliesslich der dazu nötigen Motivierungstechniken. Damit werden kurzfristige Erfolge gesucht und der langfristige Aufbau der Kompetenz in schulischen Lernbereichen vernachlässigt.

Je mehr man auf diese Weise testet, desto weniger stabil sind die Resultate. Die Kritiker argumentieren, dass das Lernen der Schülerinnen und Schüler auf die Inhalte verengt wird, die testrelevant sind, und dass die getesteten Inhalte nur einem kleinen Ausschnitt von dem darstellen, was die Lernenden wissen sollten. In Metanalysen zeigt sich allerdings, dass

---

<sup>3</sup> Daten zur Qualitätssteigerung finden sich auch in: Arbeitsgruppe Internationale Vergleichsstudie (2003), S. 167ff.

diese Befürchtungen übertrieben sind, oft mit schlechten Tests zu tun haben und gelegentlich auch nur Schutzbehauptungen darstellen. Mindestens übersehen sie die positiven Möglichkeiten der Überprüfung des Unterrichts mit Hilfe von Evaluationen und Rückmeldesysteme, die heute weltweit entwickelt werden (Evers/Walberg 2004).

In den Vereinigten Staaten und anderen angelsächsischen Ländern werden die Testergebnisse oft mit High-Stakes verknüpft. „High Stakes“ bezeichnen keine besondere Form oder Art des Testens, sondern die mit den Resultaten verbundenen Konsequenzen. Der amerikanische Ausdruck „stake“ lässt sich einerseits mit dem deutschen Wort „Grenzpfehl“ zusammenbringen, andererseits mit der Idee des Spiel- oder Wetteinsatzes. „High Stakes Testing“ hat ernsthafte Konsequenzen für die Person, die getestet wird, etwa dass es sich um einen finalen Test handelt, der nicht wiederholt werden kann und der über Promotionen oder Selektionen entscheidet.

Tests sind dann „high stakes“, wenn sie dazu verwendet werden, Urteile über Zugänge, Graduierungen oder Promotionen zu fundieren oder festzulegen. Sie dienen nicht primär der Förderung, sondern der Zu- und Abweisung innerhalb eines selektiven Systems. Das lässt sich zum Beispiel auch mit Lohnfolgen verbinden. Es gibt amerikanische Schuldistrikte, in denen die Lehrkräfte umso mehr verdienen, je besser ihre Schülerinnen und Schüler in den Tests abschneiden. Diese Art des *performance pay* - Bezahlung nach Leistung - ist naturgemäss sehr umstritten, wobei selbst die amerikanischen Lehrerverbände nicht unisono dagegen sind. Zumeist wird die Fairness des Verfahrens bezweifelt, nicht die Unterschiede in den Leistungen der Lehrkräfte.

Eine gibt noch eine weitere Folge, die oft nicht bedacht wird: Die Lehrkräfte ebenso wie die Schülerinnen und Schüler richten sich kognitiv und vor allem motivational darauf ein, *dass* und *wie* getestet wird (O’Neil/Abedi/Miyoshi/Mastergeorge 2004).<sup>4</sup> 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. Ohne klare Zweckbestimmung und ohne effektive Rückmeldesysteme besteht tatsächlich die Gefahr der Ritualisierung, die wenig Auswirkungen auf den Unterricht hat.

Amerikanische Erfahrungen zeigen ausserdem, dass rigide Testprogramme die soziale Ungleichheit verstärken und die ohnehin gegebene Benachteiligung bestimmter Gruppen erhöhen können. Gute Testprogramme müssen aufwendig entwickelt werden, die interne Kommunikation der Resultate ist ein andauerndes Problem (Orfield/Kornhaber 2001, S. 149f.) und der Aufwand macht nur Sinn, wenn er bei den politischen Entscheidungsprozessen auch genutzt wird. Datenbezogene Bildungspolitik ist aber auch in den Vereinigten Staaten immer noch die Ausnahme (ebd., S. 155), weil sich die Daten oft nicht mit den Interessen decken. Letztlich stellt also diese Art von Qualitätssicherung selbst ein Qualitätsproblem dar.

Eine weitere Variante der Kritik betrifft die politische Rhetorik hinter den Bildungsstandards (Baker 2004, S. 153ff.) Ein bekannter Slogan lautet: „No child left behind.“ Wenn zum Beispiel in der amerikanischen Administration behauptet wird, es sei möglich, die öffentliche Bildung so auszurichten und zu reformieren, dass „kein Kind zurückbleibt,“ so ist zunächst damit nicht mehr erfasst, als die Kommunikation eines politischen Slogans. In der Praxis erreichen nie alle Kinder gleiche Ziele, schon weil die

---

<sup>4</sup> Vermutet wird ein „senioritis-effect“, also eine Abnahme der Motivation am Ende der Schulzeit (O’Neil/Abedi/Miyoshi/Mastergeorge 2004, S. 199).

Schule soziale Nachteile nur begrenzt ausgleichen kann (Meier/Wood 2004). Slogans sind keine Ziele, aber sie werden vielfach genauso verstanden, als Eröffnung von Möglichkeiten, die oft praktisch ausgeschlossen sind. Andererseits ist es nur mit Hilfe von Standards möglich, die Linie zu bestimmen, die auch schwächere Schüler erreichen müssen. Förderprogramme setzen Standards und darauf bezogen Ressourcen voraus, wenn sie wirksam werden sollen.

Es gibt aber noch eine ganz Art von Kritik, die von den Befürwortern von Standards vorgebracht wird. Diese Kritik geht aus von den politischen Problemen der Implementation und beklagt, dass und wie die Standards unterlaufen werden. Basis dieser Kritik ist ein Vergleich der tatsächlichen Standardentwicklung in den amerikanischen Bundesstaaten (Finn Jr./Petrelli 2000). Die Probleme der Implementation werden so gefasst:

„It’s difficult to reach agreement on standards, hard to get the assessments properly aligned with those standards, and painful - mainly for political reasons - to impose meaningful consequences on students, teachers, and schools” (Finn Jr. 2002, S. 32).

Das Unterlaufen der Standards ist auf vielfältige und oft kaum sichtbare Weise möglich. Standards können stark oder schwach sein, leicht oder anforderungsreich, was oft aus den Formulierungen gar nicht hervorgeht. Die Qualität und Strenge der Tests stimmt oft nicht mit den publizierten Standards überein. Die Linie zum Bestehen der Tests kann unterschiedlich gezogen werden. Auch wenn der Test aus qualitativ guten Fragen besteht, können viele Schüler den Test bestehen, obwohl sie viele Fragen *nicht* beantwortet haben. Wenn die Bestehensrate ein Indikator für den Erfolg oder die Qualität einer Schule ist, wird die Rate steigen, sofern die Länder oder die Distrikte Einfluss nehmen können.

Grundsätzlich wird in den meisten Beiträgen nicht die Erweiterung der Perspektive hin zu den Resultaten von Schule und Unterricht in Frage gestellt, sondern eher die Rhetorik der zu grossen Erwartungen oder die technischen wie die politischen Probleme des Testens. Die Politik muss daher sehr genau bestimmen, was mit dem Bildungsstandards erreicht werden soll und wie sich die Erfahrungen verwenden lassen. Die Akzeptanz im Feld hängt von dieser Vorgabe ab. Je vager die Idee des Testens kommuniziert wird und je grösser der Interpretationsspielraum ist, desto mehr entstehen Unsicherheit und Abwehr. Dass die Zukunft längst begonnen hat, zeigt ein Blick in die Schweiz.

Tests als formative Rückmeldesysteme funktionieren inzwischen im Volksschulbereich recht gut, auch weil sie nicht als High-Stakes-Testing angelegt sind. Die beiden bekannten Beispiele sind *Klassenscockpit* und *Stellwerk*, die inzwischen viel genutzt werden. Die Tests sollen leistungsfördernd wirken, die Förderung des Leistungsverhaltens setzt voraus, dass der Leistungsstand der Klasse eingeschätzt werden kann, ohne dass die Bezugsnorm allein die Normalverteilung der jeweiligen Klasse wäre. Viele Lehrkräfte nutzen heute diese Möglichkeit und sie verschwenden keinen Gedanken mehr daran, dass sie sich früher gegen jeden Vergleich gewehrt haben.

Wenn der Vergleich nicht mit einem persönlichen Risiko behaftet ist, wird er akzeptiert, sofern sich der Nutzen zeigt. Die Schülerinnen und Schüler gehen damit inzwischen ganz selbstverständlich um. Das zeigt sich auch daran, dass sie zwischen den normalen Tests und dem Klassenscockpit unterscheiden. Wenn man sie fragt, wie der Tag war, sagen sie bei Gelegenheit, heute war „Klassenscockpit.“ Die Lehrkräfte stellen die Noten der einzelnen Schüler oft so dar, dass ihre eigene Bewertung neben dem Ergebnis von

Klassenscockpit erscheint, um mit einem Vergleich die Transparenz zu erhöhen, nicht zuletzt im Blick auf die Elterngespräche.

Die Schülerinnen und Schüler werden dann mit einem objektiven Resultat konfrontiert, das sie weder schön reden noch wegverhandeln können. Studien zeigen, dass sie das oft mehr beeindruckt als die Ermahnungen der Lehrkräfte. Aber auch deren Seite profitiert. In einer gerade vorgelegten Zürcher Evaluation von *Stellwerk* zeigt sich, dass neben dem Test selbst vor allem die Standortgespräche bei den Lehrkräften auf Zustimmung stossen (Sigrist/Kammermann 2006). Sie nehmen hier Einfluss, auch weil die Eltern sich dem Resultat der Tests nicht entziehen können. Allein die Tatsache, *dass* getestet wird, hat in der Praxis einen hohen Symbolwert, den die Lehrkräfte dann auch einzusetzen wissen. Sie können damit offenbar gezielter und nachhaltiger fördern als bisher.

Eine zentrale und scheinbar besonders heikle Frage, die mit der Einführung von Leistungstests verbunden ist, bezieht sich auf die Verwendung der Daten. Wie verhindert man die Schubladisierung? Die Steuerung über das Resultat wäre keine, wenn die Ergebnisse gar nie im Klassenzimmer ankommen und produktiv genutzt werden. Dieser Frage geht ein Projekt nach, das „Check Five“ heisst (Tresch 2007). So wird ein vergleichender Leistungstest im Kanton Aargau genannt, der im fünften Schuljahr der Primarschule durchgeführt wurde. Das Projekt ist vom Grossen Rat in Auftrag gegeben worden und hat den Zweck, ein Instrument zu entwickeln, mit dem sich die Ergebnisse von Evaluationsstudien an die Lehrkräfte rückvermitteln lassen, so dass auch tatsächlich Fördermassnahmen ergreifen.

Wenn derartig aufwändige Studien wie die PISA-Tests auch einen praktischen und nicht bloss einen politischen oder wissenschaftlichen Wert haben sollen, dann ist die Frage unausweichlich, was die Lehrkräfte mit Daten aus Leistungstests anfangen können. „Check Five“ ist darauf eine mögliche Antwort. Der Test wurde betraf vier Dimensionen, nämlich

- Mathematik,
- Deutsch,
- kooperatives Problemlösen
- und selbst reguliertes Lernen.

Weil hinter dem Test eine politische Absicht steht, rechnete das Projektteam mit erheblichen Widerständen in der Lehrerschaft und wäre zufrieden gewesen, wenn sich 40 Klassen gefunden hätten. Von den 370 Lehrpersonen, die im Kanton Aargau in der fünften Klasse unterrichten, meldeten sich 262, der Test wurde schliesslich mit 140 Klassen und 2.531 Schülern durchgeführt. Die Akzeptanz solcher Verfahren ist grundsätzlich offenbar vorhanden. Die Lehrkräfte wurden vor Durchführung des Tests befragt, mit welchen Einstellungen und Erwartungen sie an das Projekt herangehen, für das sie sich freiwillig gemeldet haben.

Jede einzelne Lehrkraft erhielt zwei Monate nach dem Test eine Rückmeldung, die zeigt, wie ihre Klasse im Vergleich zum Gesamtergebnis abgeschnitten hat. Im Anschluss an die Rückmeldung der Ergebnisse erfolgten noch zwei weitere Befragungen, die die Umsetzung erhoben. Sie hatten vier hauptsächliche Ziele:

- Wie gehen die Lehrkräfte mit den Daten ihrer Klasse um, zumal dann, wenn sie nicht gut sind?
- Wie analysieren sie die Ursachen?

- Wie reflektieren sie die Qualität ihres Unterrichts im Lichte der Daten?
- Und welche Massnahmen ergreifen sie zur Verbesserung der Qualität?

Das Projekt ist inzwischen abgeschlossen und die Ergebnisse liegen vor. Zentral sind auch in dieser Befragung die Leistungsunterschiede zwischen den Schülerinnen und Schülern, aber auch zwischen den einzelnen Klassen. Ein Teil der Schüler konnte nahezu alle Mathematikaufgaben richtig lösen, ein anderer Teil löste nur zwischen 10 und 20 Prozent der Aufgaben richtig. Zwischen den Klassen waren die Unterschiede am grössten in den Bereichen Grammatik und Arithmetik, also klar strukturierten Lernfeldern, die offenbar höchst verschiedenen unterrichtet werden. Weitaus geringer waren die Unterschiede im Textverständnis, auch weil das nicht nur im Deutschunterricht abverlangt wird. Die Unterschiede lassen sich nicht mit der sozialen Herkunft erklären,<sup>5</sup> sondern sind hausgemacht.

Die Lehrkräfte waren umso mehr aufgefordert zu reagieren und sich zu überlegen, mit welchen Massnahmen sie die Qualität ihres Unterrichts verbessern wollen. So verwendet, haben Leistungstests Potenzial für die Praxis des Förderns.

- Die Studie zeigt, dass die Testergebnisse mit den Eltern sowie den Schülerinnen und Schülern zum Teil aufwändig kommuniziert wurden.
- Die Lehrkräfte geben an, mit den Ergebnissen die fachlichen Leistungen ihrer Schüler besser einschätzen zu können.
- Sie ergänzen damit das klasseninterne Bezugssystem und kommen zu objektiveren Bewertungen.
- Hunderte von Massnahmen zur Verbesserung des Unterrichts wurden vorgelegt, von denen mehr als die Hälfte im Berichtszeitraum auch umgesetzt wurden.
- 131 der 140 Lehrpersonen würden ein zweites Mal an einem solchen Test teilnehmen, was man durchaus als Erfolgsindikator betrachten kann.

Zusammengefasst gesagt: Die Akzeptanz von Leistungstests steigt mit der positiven Erfahrung der Verwendbarkeit der Resultate. Akademische Vorbehalte teilen die Lehrkräfte nicht (Tresch/Moser 2005). Sie nutzen die Daten, wenn man ihnen eine Handreichung zum Umgang mit den Ergebnissen von Leistungstests anbietet (Departement Bildung, Kultur und Sport 2005). Das Verfahren stösst auf hohe Zustimmung auch bei den Schülerinnen und Schülern sowie bei den Eltern, was auch durch andere Studien bestätigt wird. Daraus lässt sich folgender Schluss ziehen: Objektivierete, transparente Verfahren der Leistungsmessung sind keine Horrorerfahrungen, sondern wirksame Instrumente, sofern sich die Daten ins Feld rückübersetzen lassen und die Lehrkräfte erreichen. Die Förderung der einzelnen Schüler kann so ein neues Niveau erreichen.

### *Literatur*

Arbeitsgruppe Internationale Vergleichsstudie (2003). Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten. Berlin: BMBF.  
 Arnet, M.: Das Schulkonkordat vom 29. Oktober 1970. Entstehung, Geschichte, Kommentar. Bern: EDK 2000.

---

<sup>5</sup> Die Segregation der Schülerpopulation nach bildungsrelevanten Merkmalen im Kanton Aarau ist weit geringer als etwa im Kanton Zürich (Tresch/Moser 2005, S. 68).

- Baker, E.L. : Reforming Education Through Policy Metaphors. In: Th. Fitzner (Hrsg.), Bildungsstandards. Internationale Erfahrungen - Schulentwicklung - Bildungsreform. Bad Boll: Evangelische Akademie 2004, S. 150-163.
- Department Bildung, Kultur und Sport des Kantons Aargau (BKS) (Hrsg.): Leistungen messen und beurteilen. Handreichungen zum Umgang mit Ergebnissen von Leistungstests. Aarau: BKS 2005.
- EDK: Interkantonale Vereinbarung über die Harmonisierung der obligatorischen Schule. HarmoS-Konkordat. Bericht zur Vernehmlassung. Bern: EDK 2006.
- Evers, W./Walberg, H. (Eds.): Testing Student Learning, Evaluation Teaching Effectiveness. Stanford Ca: Hoover Press 2004.
- Finn, Jr., Ch.E.: Real Accountability in K-12 Education. The Marriage of Ted and Alice. In: In: W. M. Evers/H.J. Walberg (Eds.): School Accountability. Stanford Ca.: Hoover Press 2002, S. 23-46.
- Finn Jr., Ch.E./M.J. Petrilli.: The State of the State Standards 2000. Washington DC: Thomas B. Fordham Foundation 2000.
- HARMOS: Zielsetzungen und Konzeption. Juni 2004. Bern: EDK 2004.
- Interkantonale Vereinbarung über die Harmonisierung der obligatorischen Schule vom 14. Juni 2007.
- Klieme, E. et al.: Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Herausgegeben vom Bundesministerium für Bildung und Forschung. Bonn: BMBF 2003.
- Konkordat über die Schulkoordination vom 29. Oktober 1970.
- Levin, H. M.: A Conceptual Framework for Accountability in Education. In: School Review Vol. 82, No. 3 (May 1974), S. 363-391.
- 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.
- Miller, R. (Ed.): Educational Freedom for a Democratic Society. A Critique of National Educational Goals, Standards, and Curriculum. Brandon, VT: Psychology Press/Holistic Education Press 1995.
- 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.
- Oelkers, J.: Zum Problem von Standards aus historischer Sicht. In: Neue Sammlung 44, 2004, S. 179-200.
- Oelkers, J./Reusser, K.: Expertise „Qualität entwickeln - Standards sichern - mit Differenz umgehen.“ Berlin: Bundesministerium für Bildung und Wissenschaft 2007.
- Orfield, G./Kornhaber, M.L. (Eds.): Raising Standards or Raising Barriers? Inequality and High-Stakes Testing in Public Education. New York: The Century Foundation Press 2001.
- Plewis, I./Goldstein, H.: The 1997 Education White Paper - A Failure of Standards. In: British Journal of Curriculum and Assessment Vol. 8, 1998, S. 17-20.
- Power, M.: The Audit Society: Rituals of Verification. Oxford: Oxford University Press 1997.
- Siegrist, M./Kammermann, M.: Begleitende und abschliessende Auswertung der Erfahrungen mit dem neu gestalteten Schuljahr an der Sekundarschule des Kantons Zürich. Zwischenbericht. Vervielf. Ms. Zürich: Interkantonale Hochschule für Heilpädagogik 2006.
- Tresch, S.: Potenzial Leistungstest. Wie Lehrerinnen und Lehrer Ergebnismeldungen zur Sicherung und Steigerung ihrer Unterrichtsqualität nutzen. Bern: h.e.p. Verlag 2007.
- Tresch, S./Moser U.: Check 5. Schlussbericht 2005 zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau. Vervielf. Ms. Zürich: KBL 2005.

Wolf, A.: A Comparative Perspective on Educational Standards. In: H. Goldstein/A. Heath (Eds.): Educational Standards. (Proceedings of the British Academy Vol. 102) Oxford: Oxford University Press 2000, S. 9-37.