

Jürgen Oelkers

*HarmoS wirkt?!
Pädagogische Konsequenzen auf der Sekundarstufe II^{*)}*

Der Titel meines Vortrag ist sicher übertrieben, denn wie kann etwas wirken, was noch gar nicht existiert oder, freundlicher gesagt, was mehr oder weniger energisch erst auf den Weg gebracht wird? Wer, wie ich, im Thurgau wohnt, weiss zudem, dass dieser Weg durchaus steinig sein kann. Allerdings konnte sich der Thurgau den neuerlichen Kampf gegen den alten Berner Bildungsvogt leisten, weil praktisch alles, was HarmoS an Strukturreformen vorsieht, bereits vorhanden ist, interessanterweise getragen von einer Partei, die einen abgewählten Bundesrat in ihren Reihen weiss und ein neues Politikfeld entdeckte. Aber der Pulverdampf scheint verrauchte zu sein, Indianer und Kavallerie kämpfen an anderen Orten und der legendäre Bildungsvogt scheint offenbar nur noch für die Weltwoche schlagzeilenträchtig zu sein, mit welchem Effekt für die Auflage, ist nicht bekannt.

Gleichwohl, „HarmoS wirkt,“ mindestens in dem Sinne, dass Tagungen wie diese stattfinden, die zur Vorbereitung auf das Kommende dienen, nicht auf das kommende Unheil, aber auch nicht auf das Heil, denn Heilserwartungen sind unangebracht. Das sage ich, weil ich aus einer Disziplin komme, die ständig Heilserwartungen produziert. Aber HarmoS ist keine pädagogische Utopie, sondern ein gross angelegter bildungspolitischer Versuch, den man durchaus als gewagt bezeichnen kann. Kritiker bemängeln schon die Sprache dieses Versuches und malen die Gefahr an die Wand, dass mit den Bildungsstandards eine Art Industrienorm eingeführt wird. Und in manchen Gymnasialkollegien macht das Wort der „McDonaldisierung“ des Unterrichts die Runde. Such in diesem Sinne „wirkt“ HarmoS - als Befürchtung.

Andererseits betrifft HarmoS zunächst ja gar nicht die Sekundarstufe II, sondern nur die obligatorische Schulzeit, einschliesslich der neuen Basisstufe. Wenn von einer politischen Wirkungsabsicht die Rede ist, dann bezieht sie sich auf die Volksschule, wo es bislang weder Mindeststandards noch verbindliche Leistungstests noch entwickelte Förderprogramme gibt. Lehrpläne und Lehrmittel basieren nicht auf Kompetenzstufen, unterstellt wird nach wie vor, dass die Ziele des Unterrichts von allen Schülerinnen und Schülern *gleich* erreicht werden können, und Qualitätssicherung muss von den angeblich so autonomen Schulen mühsam gelernt werden. Das gilt natürlich nur für den Kanton Zürich und die dortige Umsetzung des Volksschulgesetzes.

Die Frage meines Vortrages klingt so gesehen reichlich hypothetisch. Was sollen die Wirkungen von HarmoS in einem Bereich sein, für den das Projekt gar nicht vorgesehen ist? Man könnte etwas lahm von der Verbesserung der Kompetenzentwicklung in Kernfächern der Volksschule sprechen, von der irgendwie auch die Gymnasien und die Berufsschulen profitieren werden. Aber bis man das weiss, ist meine Pensionsgrenze längst überschritten, denn dann müsste eine Kohorte das neue System durchlaufen haben und dabei auch gut

^{*)} Vortrag auf dem „SchiLf-Netz-Tag 2009“ am 27. März 2009 im GiB Thun.

beobachtet worden sein. Wenn also „HarmoS wirkt“, dann mit vielen Unsicherheiten, weil grosse Zeiträume in Anspruch genommen werden und unterwegs viel passieren kann.

Nach diesen Relativierungen könnte ich eigentlich gehen, denn was will ich jetzt noch zum Thema sagen? Andererseits bedeutet HarmoS auch für die Sekundarstufe II eine Zäsur in der Schulkultur, für die Gymnasien vermutlich mehr als für die Berufsschulen, die ja seit längeren unter bildungsvogtlicher Aufsicht stehen und die Effizienzanforderungen der Betriebe vor sich haben. Mit Zäsur meine ich die Ergebnisorientierung oder die Einführung von Regeln und Verfahren zur „Outputsteuerung“. Ich gebe zu, dass dieses Wortungetüm ästhetische Abwehr gegenüber HarmoS weckt, andererseits bezeichnet es die Sache ziemlich genau, auch wenn es wohl nie in der Sprache der Deutschlehrer heimisch werden wird.

Im Folgenden werde ich zunächst fragen, was an der Zäsur wirklich neu ist und so auf die Sekundarstufe II erst zukommt und nicht bereits realisiert ist. „Standards“ an sich sind weder neu noch stellen sie eine Zumutung dar (1). In einem zweiten Schritt werde ich etwas über die Schwierigkeiten des Unterrichts *vor* der Sekundarstufe II sagen und mich dabei auf Leistungsdaten aus dem Kanton Zürich beziehen, die natürlich mit dem üblichen Vorbehalt zu verstehen sind. Die anderen Kantone müssen sie nicht beachten (2). Abschliessend gehe ich auf die neuen Instrumente ein und frage, wie Leistungstest sinnvoll einzusetzen sind und ob Rankings zu den pädagogischen Konsequenzen auf der Sekundarstufe II zählen. Ich gebe zu, auch das ist wieder eine Zürcher Besonderheit (3).

1. Standards auf der Sekundarstufe II

Der Verlauf der politischen Diskussion über HarmoS hat den Eindruck erweckt, dass „Bildungsstandards“ etwas grundlegend Neues seien und gerade Gymnasien aufgefordert seien, sie zu entwickeln. Aber Appetenzverhalten ist nicht angebracht. Für die Gymnasien sind Standards keine neue Erfahrung oder dies nur in einer bestimmten Hinsicht. Die Lehrkräfte hören den Ausdruck „Bildungsstandards“ nicht gerne, aber wogegen sie sich wenden, ist eigentlich nur eine übertriebene Sicht der psychometrischen Steuerung des Unterrichts, also ständige Tests, die regierungsamtlich genutzt werden. Standards selbst dagegen sind Alltag in den Schulen.

- 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 in gewisser Hinsicht die Inhalte des Unterrichts
- und die Relation zwischen Lehrern und Schülern ist sogar ein statistischer Standard.

Gymnasien und Berufsschulen unterscheiden sich von anderen Schulen nicht nur durch das Fächerangebot, sondern auch durch eine differenzierte Niveauforderung. Lehrkräfte des Gymnasiums benutzen den Ausdruck „gymnasiale Standards“ wie selbstverständlich, wenn sie die Besonderheit ihres Unterrichts aufzeigen wollen. Es macht für sie geradezu die Identität der Gymnasien aus, dass sie spezielle Anforderungen stellen können. Wenn der Ausdruck „Bildungsstandards“ gerade bei Gymnasiallehrkräften oft negativ besetzt ist, dann hängt mit Befürchtungen zusammen, die eine bildungspolitische Diskussion ausgelöst hat.

Der Tatbestand von gymnasialen Standards ist davon unberührt; die Frage ist, wie sie formuliert, festgelegt und überprüft werden.

Angesichts der gelegentlich etwas irreführenden Diskussion halte ich eingangs fest:

1. Jedes Schule auf der Sekundarstufe II hat und vertritt *Standards*, versteht man darunter die Qualitätsanforderungen des Fachunterrichts.
2. Der Erfolg, also *schulische 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.

Soweit gibt es nichts Neues. Am Ende eines jeden Lehrgangs, den Gymnasien und Berufsschulen anbieten, stehen Kompetenzen, allerdings nie bei allen Schülern gleich. Das zu sagen, ist eigentlich trivial und erhält nur deswegen eine herausgehobene Bedeutung, weil die heutige Diskussion über Bildungsstandards neue Begriffe - manche sagen auch einen neuen Jargon - benutzt. Statt wie früher von „Fähigkeiten und Fertigkeiten“ spricht man heute von fachlichen und überfachlichen „Kompetenzen“, statt Leistungsanforderungen ist von „Standards“ die Rede, und statt „Niveauunterschied“ wird der Ausdruck „Kompetenzstufen“ benutzt. Die neue Sprache allein verweist aber nicht schon auf etwas, das in der Praxis neu wäre (Oelkers 2004).

Die Standardisierung der Schule zeigt sich nicht zuletzt in den Verfahren der Leitungsbeurteilung und so der Notengebung. Historisch wurde als Notensystem fast immer eine Skala von fünf oder sechs Stufen verwendet. Dieser Befund ist sehr stabil. Der „Gothaer Schulmethodus“ von 1642¹ sah zum ersten Male im deutschen Sprachraum ein einheitliches Schema für die gestufte Bewertung der Leistungen im Blick auf *Unterrichtsgegenstände* vor. Gleichzeitig sollten *Ingenium* und *Mores*, also die Geistesgaben und das sittliche Verhalten, der Schüler beurteilt werden. Für das *Ingenium* genühten vier Stufen, die beiden anderen Bereiche sollten mit dem Fünfersystem bewertet werden.

Ingenium	Unterrichtsgegenstände	Mores
Sehr fein	fein	fromm
Gut	fertig	fleissig
Ziemlich	ziemlich	still

¹ Der heutige Schema der Schulnoten geht wesentlich auf den von Andres Reyher (1601-1673) verfassten Gothaer *Schulmethodus*. Reyher war von 1639 an Rektor des Gymnasiums in Gotha. Er studierte in Leipzig, machte dort seinen Magister und lehrte an der Philosophischen Fakultät, bevor er eine Karriere als Gymnasialrektor machte. Der Herzog von Gotha, Ernst der Fromme (1601-1675), berief Reyher von Lüneburg nach Gotha, um die Reform der Volksschule voran zu bringen. „Schulmethodus“ wurde der „special- und sonderbahre Bericht“ zur neuen Volksschulordnung genannt, der 1642 veröffentlicht wurde. Diese Ordnung führte in Gotha die Schulpflicht vom fünften bis zum zwölften Lebensjahr ein, regelte die Rangfolge der Elementarfächer sowie die Einteilung der Schüler nach Klassen, schrieb die Lehrbücher vor, verfügte Regeln der Unterrichtsmethode und der Schuldisziplin, sah Realien vor und fasste schliesslich das Schema der Notengebung. In diesem Sinne handelte es sich um die erste moderne Standardisierung des Volksschulunterrichts im deutschen Sprachraum.

Schlecht	etwas/wenig	unfleissig
	schlecht	ungehorsam

Offenbar ist dieses Schema - zeitgemäss verändert - geeignet, Leistungs- und Verhaltensdifferenzen innerhalb einer bestimmten, grösseren Lerngruppe zu beschreiben. Die vergleichsweise intensive Forschung hat bislang keine Alternativen geliefert, die bei gleicher Ökonomie praktikabler wären. In diesem Sinne handelt es sich um einen Standard, den die Schule selbst hervorgebracht hat und der auch in Zukunft Verwendung finden wird.

Das lässt sich verallgemeinern. Es gibt in jeder Schule eine Beurteilungskultur, die verschieden geartet ist, jedoch in aller Regel von den Beteiligten als bewährt angesehen und somit als Standard akzeptiert wird. Das gilt angesichts klarer Normen und Verfahren besonders für die Sekundarstufe II, also für Gymnasien und Berufsschulen. Von „Beurteilungskultur“ kann auch in einem weitergehenden Sinne gesprochen. Jeder Schüler und jede Schülerin beurteilt alle seine oder ihre Lehrkräfte, wenngleich weitgehend informell und oft ohne Nutzen für den Unterricht. Studien zu den Erwartungen der Schülerschaft, darunter Absolventenbefragungen aus Zürich, zeigen, dass die Jugendlichen es inzwischen für eine professionelle Selbstverständlichkeit halten, dass *sie* mit einem formalisierten Verfahren ihre Lehrkräfte beurteilen müssten.

Wo die Schulen dieser Forderung nachkommen, geschieht das zu ihrem Vorteil. Sie erzielen einen Transparenzgewinn, wenn ich mich so ausdrücken darf. Die Ergebnisse sind beachtlich und weiterführend, ohne die zuvor vorhandenen Befürchtungen zu bestätigen. In vielen Schulen geht die Bereitschaft soweit noch nicht, aber auch hier ist Beurteilung nicht mehr allein die Sache der Lehrkräfte. Das Stichwort „Evaluation“ ist inzwischen auch im Bereich der Gymnasien kein Fremdwort mehr. Etwas unfein wird gesagt, dass dies die Folge eines Generationenwechsels sei,² aber der Tatbestand selbst ist unstrittig, Qualitätssicherung hat nach den Berufsschulen auch die Gymnasien erreicht (Oelkers 2008).

Die Lehrkräfte müssen sich auf eine neue Situation ihres Berufsfeldes einstellen, die sich im Alltag zeigen wird. Die Schulen selbst sind Teil und Objekt der Beurteilungskultur. Auch die Schweizer Gymnasien und Berufsschulen werden inzwischen extern evaluiert werden und können die Daten zur Schulentwicklung nutzen.³ Strittig ist das im Blick auf das Verhältnis von Aufwand und Ertrag sowie auf die Darstellungsform. Wenn Vergleiche zwischen Schulen in Rankings abgebildet werden, entsteht Aufregung, weil die Ausgangsbedingungen verschieden sind. Ich komme darauf zurück. Festgehalten sei, dass Evaluationen nur dann sinnvoll sind, wenn Standards gegeben sind und auch beachtet werden.

Zunächst stellt sich die Frage, wie sich das Thema der Standards über das hinaus, was die Schulen ohnehin tun, fassen lässt. Die politische Diskussion der „Bildungsstandards“ geht auf einen amerikanischen Krisenreport zurück, der 1983 erschien und den bezeichnenden Titel trug *A Nation at Risk*. Man kann das Bildungssystem der Vereinigten Staaten nur sehr begrenzt mit dem der Schweiz vergleichen. Das lässt sich an zwei kardinalen Unterschieden zeigen:

- In der Schweiz gibt es keine High School und in den Vereinigten Staaten kein duales System in der Berufsbildung.

² NZZ Nr. 67 v. 21/22. März 2009, S. 47.

³ Ein Angebot ist die *Interkantonale Fachstelle für externe Schulevaluation* (IFES) an der Universität Zürich.

- Andererseits sind die bildungspolitischen Entwicklungen in den Vereinigten Staaten immer stark beachtet worden.
- Die Theorie der Bildungsstandards basiert im Wesentlichen auf amerikanischen Diskussionen, was auch für den Zusammenhang mit Tests und Evaluationen gilt.

Die 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 und er ist auch in der Schweiz im Vorfeld des Harnos-Konkordats beachtet worden. Die Karriere des Konzepts der „Bildungsstandards“ war zunächst getragen mehr von einem bildungspolitischen Slogan als von einem genau definierten Begriff. Slogans stehen für pädagogische Bewegungen und werden zugespitzt politisiert, ohne die Assoziationen beschränken zu müssen. Das würde die Besetzung des Ausdrucks „Standards“ oder „Standarisieren“ mit der Industrienorm erklären. Aber kein Unterricht gleicht dem anderen, insofern hat sich ein unglücklicher Begriff eingebürgert.

Diane Ravitch versuchte zu explizieren, was den Kern des Konzepts ausmacht. 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
- Diese Bedingung wird so gefasst: „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*.

Damit wird die amerikanische Theoriediskussion 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. Das elaborierteste System in Europa ist seit 1988 in England aufgebaut worden, und

selbst in den liberalen Gesamtschulsystemen Skandinaviens werden inzwischen Standards und Tests verwendet.

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 bisherigen Verstehenshorizontes.

Schulen können also nicht „aus dem Nichts“ Wissen aufbauen, sondern nur die gegebenen Erfahrungen und das vorhandene Wissen erweitern (Ravitch 1995, S. 103). Allerdings ist die „Konstruktion“ des Wissens nicht so zu verstehen, dass es sich gleichsam um Erfindungen der Lernenden handelt. Das Fach oder neuerdings in den empirischen Forschungen die „Lerndomäne“ bestimmt nach wie vor die Anforderungen des Unterrichts, also die Aufgaben und Leistungen. Im HarmoS-Projekt sind Bildungsstandards Leistungserwartungen und in diesem Sinne normative Grössen, die sich im Erreichungsgrad unterscheiden (HARMOS 2004, EDK 2006, Interkantonale Vereinbarung 2007).

- 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.

Gegenüber den bisherigen Lehrplänen ist neu, dass dafür Stufen der Erreichbarkeit zur Verfügung stehen. Stoffverteilungen oder die Bestimmung der Wissensanforderungen reichen nicht mehr aus. Das ist insofern sinnvoll, als Ziele nie gleich erreicht werden, was im Lehrplan alter Art immer vorausgesetzt wurde. Nunmehr werden verschiedene Stufen unterschieden, die auch für die Notengebung relevant werden. Zu diesem Zweck müssen Aufgaben generiert und auf ihre Validität hin geprüft werden. Erst dann sind sie testrelevant. Für sich genommen reichen Tests aber nicht aus, ihre Ergebnisse müssen die Schule und den Unterricht erreichen, wenn wirklich eine Steuerung durch Ergebnisse angestrebt werden soll. Das *Feedback* wird daher zu einem Schlüsselproblem der Systementwicklung (Maradan/Mangold 2005).

Der Vorrang *fachlicher* Ansprüche in den vorliegenden Konzepten der Bildungsstandards bestätigt in mehrfacher Hinsicht das Bildungsprofil der Sekundarstufe II. Es ist verblüffend, wie stark sowohl bei Ravitch als auch in dem 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 es nicht längst zu tun hätten. Für die Primarschulen sind Bildungsstandards etwas grundlegend Neues, nicht jedoch für die Gymnasien und Berufsschulen, die sich, die sich also nicht durch bildungspolitische Suggestionen aufregen lassen sollten.

Auf der anderen Seite sind die Gymnasien auch in der Schweiz Objekte der Kritik, in der Kernelemente des Gymnasiums auf den Prüfstand gestellt werden. Die allgemeine Hochschulreife wird ebenso in Zweifel gezogen wie die Ausbildung der Lehrkräfte, die Budgets werden angegriffen, vom Sparen sind eher die weniger teuren Schulen ausgenommen und mit Hinweis auf die Quote der Studienabbrecher wird auch die Qualität der gymnasialen Bildung als fraglich erachtet. Die inzwischen gescheiterte Idee, im Aargau ein Elitegymnasium zu gründen, gehörte ebenfalls in diesen Zusammenhang. Und schliesslich ist mit den Daten beider PISA-Studien die Systemfrage neu gestellt worden: Wozu brauchen wir Gymnasien, wenn integrierte Gesamtschulen im Ausland erfolgreicher sind?

Auf diese Frage werde ich heute nicht eingehen. Sie wird bekanntlich heftig in den Kantonen nördlich des Rheins diskutiert, wo neue linke Mehrheiten ein altes linkes Projekt neu zum Leben erweckt haben. Die Gymnasien in der Schweiz sind in einer anderen Situation, so dass die neuerliche *querelle allemande* über das Bildungswesen vermutlich keine Auswirkungen haben wird. Ausserdem zeigt die deutsche Erfahrung, dass man ständig die Systemfrage stellen kann, ohne sie durch Wiederholung einer Antwort näher zu bringen (Oelkers 2006). Es ist auch seltsam, dass die finnische Gesamtschule als Musterbeispiel erhalten muss, ohne die stark selektiven Gymnasien in Finnland und die hohe Jugendarbeitslosigkeit in Rechnung zu stellen.

Aber das ist nicht mein Thema. In meinem zweiten Teil geht es um die Entwicklung des Lernstandes, der über die Jahrgänge hinweg aufgebaut wird und so die gesamte Schulzeit betrifft. Gymnasien und Berufsschulen kommen erst am Ende, dann, wenn es für grundlegende Korrekturen zu spät ist. Insofern müssen die Schulen auf der Sekundarstufe II ein Interesse daran haben, was vorher geschieht. Harmos, könnte ich auch sagen zwingt zu einer Gesamtsicht, weil der gestufte Aufbau von Kompetenzen angestrebt wird. Doch wie sieht es damit im Schulalltag aus? Diese Frage wird mich in einem zweiten Schritt beschäftigen.

2. *Unterricht und Lernstand*

Die Bildungsdirektion des Kantons Zürich hat im Jahre 2004 ein Forschungsprojekt in Auftrag gegeben, das eine Kohorte von anfänglich rund 2.000 Schülerinnen und Schülern während ihrer gesamten Schulzeit beobachtet. Die Schüler werden bei Schuleintritt getestet und dann nachfolgend in der dritten, sechsten und neunten Klasse erneut, um so die Entwicklung des Lernstandes erfassen zu können. Dieser Längsschnitt ist einmalig in der Schweiz, auch im deutschsprachigen Ausland ist ein solches Forschungsvorhaben bislang nicht unternommen worden. Vorliegende Studien erfassen Momente, aber nicht Prozesse, also nicht Veränderungen an verschiedenen Zeitpunkten. Die Ergebnisse der ersten beiden Studien liegen inzwischen vor (Moser/Stamm/Hollenweger 2005; Moser/Keller/Tresch 2008).

- Bei Schuleintritt hat etwa vier Fünftel der Kinder den Lernstoff der ersten Klasse in Mathematik teilweise bewältigt und ein knappes Fünftel befindet sich leistungsmässig bereits in der zweiten Klasse.
- Die Lesekompetenzen liegen weiter auseinander. 5 Prozent der Kinder kennt keine Buchstaben, etwa 35 Prozent erfassen die Verbindung zwischen Buchstaben und Lauten nur im Ansatz, 15 Prozent lesen bereits schwierige Wörter und längere Sätze.

- Noch disparater ist der Wortschatz, also die Fähigkeit, Bilder mit eigenen Wörtern richtig zu bezeichnen.

Nur knapp ein Drittel der Kinder beherrscht die richtigen Bezeichnungen, ein weiteres Drittel überbrückt geschickt die Lücken, das letzte Drittel beherrscht nur ganz elementare Bezeichnungen, also kann sagen, dass ein Hammer auf einem Bild ein „Hammer“ ist, mehr jedoch nicht. Korrekte Bezeichnungen für andere Bilder können nicht abgerufen werden.

Nicht alle Kinder kommen überhaupt in die Regelschule. Im Jahre 2005 sind 8.5 Prozent der Kinder in „Kleinklassen“ eingewiesen worden, weil sie aus Sicht des Kindergartens nicht als schulreif galten. Kleinklassen werden als Förderklassen geführt, stellen faktisch aber eine soziale Selektion dar. Die Studie zeigt, dass der Leistungsstand in den Kleinklassen tatsächlich deutlich niedriger ist als der in den Regelklassen. Rechnet man die Kleinklassenschüler dazu, dann ist die Lücke zwischen den Leistungsstarken und den Leistungsschwachen in allen Bereichen noch grösser als die bisher genannten Zahlen dies anzeigen. Ob in Kleinklassen die Förderabsicht tatsächlich Erfolg hat, ist nicht untersucht worden. Die Schüler gewinnen dadurch nichts, weil sie nach einem Jahr Kleinklasse mit dem ersten Schuljahr beginnen, bzw. nach zwei Jahren Kleinklasse in die zweite und nicht in die dritte Klasse versetzt werden.

- Die zweite Studie nach drei Jahren Schulzeit zeigt, dass Unterricht starke, aber auch ungleiche Effekte hat.
- Die Schülerinnen und Schüler erreichen die erste Klasse der Volksschule mit grossen Unterschieden im Lernstand, Unterschiede können ausgeglichen werden, sofern qualitativ guter Unterricht stattfindet.
- Er macht den Unterschied, was mit der Kompetenz der Lehrkräfte, der Zusammensetzung der Klasse und den je erreichten Lernfortschritten zu tun hat.

Die Testleistungen bei Beginn der Schule erlaubten nur begrenzt Vorhersagen für die Leistungen, die drei Jahre später erzielt wurden. Die zweite Studie zeigt, dass viele Kinder schlechte Chancen beim Start ausgleichen konnten und so vom Unterricht profitiert haben. In diesem Sinne ist Schule kein Schicksal.

Allerdings gibt es deutlich auch Risikogruppen, solche Kinder nämlich, die die Lernziele nicht erreicht haben und mit hoher Wahrscheinlichkeit im weiteren Verlauf ihrer Schulkarriere weiter zurückbleiben. Der oft benutzte Slogan „No Child Left Behind“ erweist sich so als Zweckoptimismus. Die Zürcher Studie wirft die Frage auf, wie mit der Risikogruppe umgegangen werden soll, und das verweist auf die Notwendigkeit einer konsequenten Leistungsentwicklung, für die ein anderer Slogan zur Verfügung steht, nämlich „Fördern und Fordern.“ Der Slogan selbst hat keinen Inhalt und besagt eigentlich nur, dass irgendwie hälftig verfahren werden soll. Die wirklichen Probleme sind deswegen schwer zu bearbeiten, weil sie in den Klassen entstehen und trotz erheblichem Einsatz der Lehrkräfte auftreten.

Der Abstand zwischen den leistungsstarken und den leistungsschwachen Schülern muss im Verlauf der Schulzeit also keinesfalls geringer werden, eher ist anzunehmen, dass er mit den Anforderungen grösser wird.

- Bestimmte Gruppen von Schülerinnen und Schülern verbessern ihren Lernstand nicht, sondern bleiben auf dem einmal erreichten Niveau stehen.

- Das gilt besonders für die letzten Schuljahre, und dann, wenn lohnende Aufgaben und Lernanlässe fehlen.
- Dieses Problem der negativen Differenzierung lässt sich nicht einfach durch punktuelle Fördermassnahmen bearbeiten, sondern stellt sich im Blick auf den Verlauf der gesamten Schulzeit.

Die Studien verweisen auch darauf, dass die Entwicklung des Lernstandes zwischen verschiedenen Klassen auch ein- und derselben Schule höchst unterschiedlich sein kann. Das gilt ebenso für die Effekte der Förderung. In den einen Klassen gelingt es, die Leistungsschere zu verringern, in anderen jedoch nicht, was auch damit zu tun, dass sehr verschieden unterrichtet wird. Verbunden damit ist allerdings auch ein strukturelles Problem, das deutlich benannt werden muss.

- Fortlaufender Unterricht, der in einem bestimmten Zeitraum und ohne zusätzliche Ressourcen Ziele erreichen soll, die für alle gleich gelten, kann nur sehr begrenzt Nachteile einzelner Schüler ausgleichen.
- Kinder bleiben auch deswegen zurück, weil erreichbare Ziele für sie gar nicht bestehen.
- Sie werden dann als „leistungsschwach“ bezeichnet und oft aus dem Regelunterricht herausgenommen.

In der Schweiz ist die Zahl derjenigen Schülerinnen und Schüler, die ausserhalb der Regelklassen unterrichtet werden, in den letzten zwanzig Jahren ständig angewachsen (Kronig 2007). Eine Zuweisung in Kleinklassen oder anderen Formen der Auslagerung mindert die Chancen in erheblicher Weise und kann doch als „Fördermassnahme“ bezeichnet werden kann. Die Zuweisung erfolgt in bester Absicht, nämlich zum Vorteil der Schülerinnen und Schüler, ähnlich wie das „Sitzenbleiben“ als eine zweite Chance gewertet wird, was oft gar nicht der Fall ist. Die Schüler werden einfach nur ein Jahr länger in der Schule gehalten. Wenn sie ihre Leistungen verbessern sollen, dann sicher nicht so, dass sie das gleiche Programm zum zweiten Mal absolvieren müssen.

Die Schulstruktur ist auf dieses Problem der Zieldifferenzierung bislang kaum eingestellt. Im Kanton Zürich gibt es auf der Stufe der Sekundarschule fünf verschiedene Möglichkeiten, die Schülerinnen und Schüler zuzuweisen. Die Oberstufe der Volksschule kennt drei verschiedene Niveaus, in die die Schüler bei Beginn der siebten Klasse eingeteilt werden. Diese Niveaus sind entstanden aus den früheren Typen der Sekundar-, Real- und Oberschule. Die Gemeinden können zwischen diesem Modell und dem der gegliederten Sekundarstufe wählen.⁴ Hinzu kommen noch Kleinklassen und das Untergymnasium. Wenn man die Sonderschulen dazu rechnet, sind es sogar sechs Varianten (Bericht und Antrag 2005), Das ist einmalig in der Schweiz, wie ich in Thun wohl sagen darf.

In heutigen Schulevaluationen wird deutlich, dass viele Lehrkräfte „Fördern“ mit „Individualisieren“ gleichsetzen und damit oft nicht mehr meinen als die dosierte Zuteilung von Aufgaben. Leistungsstarke Schüler werden dann einfach dadurch „gefördert“, dass sie in

⁴ In der gegliederten Sekundarschule werden die Schülerinnen und Schüler auf zwei Stammklassen verteilt, deren Anspruchsniveau unterschiedlich ist. Unterschieden werden Stammklassen mit grundlegendem Niveau (G) und Stammklassen mit erweiterten Anforderungen (E). In zwei Fächern wird der Unterricht in drei Leistungsniveaus erteilt, wobei zwischen Mathematik, Deutsch, Französisch und Englisch gewählt werden kann. Hier wird unterschieden zwischen einem Niveau mit grundlegenden (g), mit mittleren (m) und mit erweiterten (e) Anforderungen.

der gleichen Zeit mehr Aufgaben lösen dürfen als die Schwächeren. „Fördern“ ist dann leicht einmal von „bestrafen“ kaum noch zu unterscheiden. Dahinter verbirgt sich ein systematisches Problem. „Fördern und fordern“ wird letztlich immer nur als Appell an die Lehrkräfte verstanden, die institutionellen Bedingungen und Ressourcen werden vernachlässigt, obwohl nicht zuletzt sie zur Erklärung von Erfolgen oder Misserfolgen beitragen können.

Mit Blick auf die Praxis gesagt: Die heutigen Klassengrößen, das Stundendeputat oder der Betreuungsaufwand sind harte Fakten, die jede weitergehende Form von individueller Förderung zunächst einmal begrenzen. Die gegebenen Umstände erlauben nicht mehr, und wenn die institutionellen Rahmenbedingungen gleich bleiben, dann ist auch kaum ein anderer Schluss möglich.

- Das Fördern findet seine Grenze schnell an den täglichen Belastungen, wenn nicht mehr geschieht, als an die Lehrkräfte zu appellieren.
- An sich lohnende Vorhaben, die zusätzliche Anstrengungen erfordern, werden dann nicht realisiert, die Arbeitszeit ist nicht beliebig steigerbar,
- und wenn es für den Aufbau einer integrativen Förderkultur keine Ressourcen, Orte und Verfahren gibt, dann erscheint sie leicht als nebensächlich oder schlimmer - als „nicht machbar.“

Begrifflich sollte man sich von einer Vorstellung trennen, die unter „Fördern“ einen Prozess versteht, der am Ende das Erreichen gleicher Ziele für alle gewährleistet. In keiner Lerngruppe erreichen alle Schüler ein- und denselben Leistungsstand. Mit dieser Idee wird eigentlich nur eine Lehrplanillusion kommuniziert. Ähnlich schwierig ist der Ausdruck „Fordern.“ Man kann darunter sehr allgemein die Ziele des Lehrplans verstehen, aber die bestimmen bekanntlich nicht wirklich den Unterricht. Hier jedoch, im Klassenzimmer, konkretisiert sich jede Leistungsforderung. Man kann auch nicht zur einen Hälfte „fördern“ und zur anderen „fordern.“ Die Formel ist ein politischer Kompromiss und soll nur anzeigen, dass beide Standpunkte, der „konservative“ des Forderns und der der „fortschrittliche“ des Förderns verträglich sind.

Bildungsstandards sind im Blick auf die Verbesserung der Förderpraxis in den Schulen ein grosser Hoffnungsträger, weil damit erstmalig für die Volksschule konkrete Zielgrösse und verbindliche Leistungserwartungen festgelegt und überprüft werden. „Standardisierung“ darf allerdings tatsächlich nicht im Sinne einer Industrienorm verstanden werden, insofern trifft der Vorbehalt der Lehrkräfte zu.

- Unterricht ist keine Fließbandarbeit und der Zeittakt der Schule ist keine Gewähr für eine effiziente Nutzung der Ressourcen.
- Am Ende der Schule steht auch kein „Produkt,“ schon gar nicht ein irgendwie gleiches,
- sondern Schülerinnen und Schüler, die im Erfolgsfalle gelernt haben, anspruchsvoll und selbständig zu arbeiten.

Die heutige betriebswirtschaftliche Sprache der Bildungsreform darf also nicht dazu führen, Machbarkeitsillusion zu entwickeln; umso wichtiger ist dann die Frage, was die Bildungsstandards und Tests zum Problem der schulischen Integration beitragen können.

Ein weiteres ungelöstes Problem erwächst aus der Tatsache, dass Bildungsstandards für die Regelschule oder für den Normalbetrieb entwickelt wurden. Was sie für die

Sonderbeschulung und deren Förderbedarf bedeuten, ist nicht abzusehen, und zwar unabhängig davon, ob die Sonderbeschulung integrativ erfolgt oder nicht und wer genau dazu gehört. Wie man ADS-Kinder oder aggressive Jugendliche auf Bildungsstandards und Leistungstests einstellen soll, ohne das Problem Ritalin zu überlassen und den Unterricht auf Teaching-the-Test umzustellen, wäre eine gute Frage an die HarmoS-Macher.

HarmoS geht aus von einem Bildungsraum Schweiz (Criblez 2008). Im Bericht zur Vernehmlassung zu HarmoS, der zu Beginn des Jahres 2006 vorlag, wird allerdings deutlich 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 auch 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, mangels Erfahrung unterstellt die Diskussion dann oft

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 Aufgaben zu verstehen, die den Leistungsstand bestimmter Fächer in einem Kernbereich testen und die für diesen Zweck fortlaufend geeicht werden müssen. 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..

Das alles berührt die Gymnasien und die Berufsschulen nicht direkt, dennoch hat HarmoS auch für die Sekundarstufe II Folgen, die sich mit einem Idealbild so skizzieren lassen:

- Der Aufbau der Kompetenzen der Schülerinnen und Schüler lässt sich genauer beschreiben als bisher.
- Die Schnittstellen und Anschlüsse sind besser geregelt.
- Die Übergänge in die aufnehmenden Schulen lässt sich objektiver gestalten, weil mehr und bessere Daten vorliegen.
- Der Fachunterricht der Sekundarstufe II kann auf einem Deutschschweizer Lehrplan aufbauen, so dass der Nacherwerb von Kompetenzen reduziert werden kann.

Das sind allerdings nur Hypothesen, die tatsächlichen Effekte von HarmoS sind derzeit natürlich noch nicht absehbar. Und es ist nicht gesagt, dass die Effekte nur positiv ausfallen werden.

Abwartend kann und sollte man den Prozess aber nicht begleiten. Die Schulen der Sekundarstufe II sind aufgefordert zu handeln, damit sie für die absehbare Zukunft gut aufgestellt sind. Das Europäische Sprachenportfolio zeigt, dass die Entwicklung von Standards bereits weit vorangeschritten ist und in Zukunft für alle Fächer und auf allen Stufen genauere Beschreibungen der Leistungserwartungen anstehen. Die Schulabschlüsse und damit der Verwendungsnutzen werden sich darauf beziehen und wenn das so ist, können die Schweizer Gymnasien und Berufsschulen nicht abseits stehen. Die Frage nach den konkreten Folgen, soweit sie absehbar sind, wird mich abschliessend beschäftigen.

3. Konkrete Schritte

Meine generelle Einschätzung sieht so aus: An den fachlichen Standards des Unterrichts auf der Sekundarstufe II wird sich kaum etwas ändern. Das Profil des Fachunterrichts bleibt grundsätzlich erhalten, und so auch die Besonderheit der Gymnasien und Berufsschulen.

- Neu sind Instrumente der Qualitätssicherung wie Tests oder Evaluationen,
- neu sind explizitere Kriterien der Leistungsbeurteilung
- sowie neue Formen der curricularen Beschreibung
- 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. Die Gymnasien entwickeln sich in diesem Sinn selbst, was natürlich auch längst der Fall ist.

Das zeigt eine kurze Übersicht: 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 2008). Zu einem Standard geworden ist auch die Homepage jeder einzelnen Schule, die sozusagen den Markenkern darstellt und vielfach genutzt wird. „Schulentwicklung“ ist also keine blosser Forderung mehr sondern Praxis.

Qualitätssicherung im Gymnasium hat zu tun mit fünf einzelnen Handlungsbereichen, die so fassen lassen:

- Evaluation der Schulqualität
- Transparenz der Leistungsanforderungen
- Entwicklung der Lerneinstellungen
- Analyse und ggf. Verbesserung der Prüfungs- und Notenpraxis
- Standards und Tests.

Wenn Qualitätssicherung mehr sein soll als ein unverbindliches Schlagwort und auch mehr als ein Webauftritt, dann müssen diese Bereiche bearbeitet werden. Dabei spielen neue Verfahren und Instrumente eine Rolle. Es ist sicher kein Zufall, dass in der Schulentwicklung zunehmend elektronische Medien verwendet werden. Auch Leistungstests gewinnen an Bedeutung. Tests sind nicht alles, sie haben auch deutliche Grenzen, aber sie werden für die Beurteilung der Leistungen immer wichtiger und sie werden ihren Platz nicht zuletzt in der Kommunikation mit Schülern und Eltern finden.

Eine zentrale Frage für die innere Entwicklung der Schulen 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.

Der Akteur für diese Entwicklung sind die einzelnen Fachgruppen in den Gymnasien und Berufsschulen, gestützt durch die kantonalen und interkantonalen Organisationen der Fachdidaktik. In den Fachgruppen werden die Inhalte des Unterrichts präzisiert und die Kriterien der Beurteilung abgestimmt. Orientierungsarbeiten sind dabei ebenso sinnvoll wie eine präzise Fassung der Notenstufen. Wichtig ist die Abstimmung quer zu den Jahrgangsstufen sowie kantonsweite oder gar überkantonale Referenzgrößen, die allerdings

bislang nicht vorliegen. Das kann sich ändern und sollte dann auch für die überfachlichen Kompetenzen, insbesondere für die Lernstrategien, in Angriff genommen werden.

Die von den Schülerinnen und Schülern eingesetzten Strategien des Lernens und Arbeitens sind keine didaktische Nebensächlichkeit.

- Von der Beherrschung anspruchsvoller Lernstrategien wird die Leistung wesentlich mitbestimmt.
- Die Erwartungen an das Fach ist die eine Seite, die tatsächlich eingesetzten Strategien des Lernens die andere.
- Sie müssen *im Unterricht* verwendet werden, aber das setzt voraus, anspruchsvolle Strategien werden als eine Leistungsanforderung verstanden.

Dafür müssten die Lehrkräfte eine Repertoireanpassung vornehmen, also von den Schülern auch andere als konventionelle Lernstrategien abverlangen. Das geht nur, wenn der tägliche Unterricht darauf eingestellt ist. Psychologische Kurse, die manche Schweizer Gymnasien zu Beginn des Schuljahres durchführen, leisten dies nicht. Wer zum Beispiel gelernt hat, was „mind-mapping“ als Verfahren ist, kann damit nicht schon im Unterricht umgehen. Die Form des Unterrichtens muss sich verändern, und die Lehrkräfte müssen ihr Repertoire auf die Lernstrategien der Schüler einstellen. Dass sich das Wohlbefinden verbessert hat, reicht nicht aus, um auch den Ertrag des Lernens zu verbessern.

Das *Know How* des Lernens ist wenig entwickelt. Die Schüler verwenden ihre Arbeitsformen wenig koordiniert und mehr oder weniger situativ. Häufig gebraucht werden einfache Strategien wie das Hervorheben wichtiger Aussagen in Texten oder Analogieschlüsse zu früheren Aufgaben gleicher oder ähnlicher Art. Wenn der Fortgang der Arbeit von den Schülern kontrolliert wird, dann nur, um keine Fehler zu machen. Metakognitive Strategien wie Planung und Evaluation der eigenen Leistung sind dagegen eher selten. Anspruchsvolle Lernstrategien haben aber einen Einfluss auf die Leistung, wobei sich viele Lehrkräfte über diesen Zusammenhang nicht im Klaren sind. Sie beurteilen, *was* und nicht *wie* gelernt wurde; wenn die Unterrichtsqualität verbessert werden soll, dann muss überlegt werden, wie die Lernstrategien in die Beurteilung mit einfließen.

Bildungsstandards im Sinne von Ravitch und der deutschen Expertise sind gebunden an externe Leistungstests. Auch für die Gymnasien bedeutet das den 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 „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)

Ob es eigene Bildungsstandards für die Sekundarstufe II 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 etwa 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. Im Kanton Zürich gibt es noch keine Orientierungsarbeiten wie in Baselland, aber dieser Weg ist sicher richtig und führt zu einer sinnvollen neuen Prüfungsform, die Vergleiche zulässt. Einheitliche Aufnahmeprüfungen wie in Zürich sind nicht ausreichend, es muss auch vor der Maturität eine vergleichende Leistungskontrolle geben, wie dies in der Volksschule mit *Stellwerk* der Fall ist. Anders als dort sind die Gymnasien wesentlich freier, die Inhalte selbst zu bestimmen, und es spricht auch nichts dagegen, eine gemeinsame Korrekturarbeit festzulegen.

Die Nachfrage nach Testsystemen im Volksschulbereich, neben „Klassencockpit“ und „Stellwerk“ aus St. Gallen auch das aargauische „Check Five“, zeigt, dass diese Chance von den Lehrkräften genutzt wird und die Instrumente sowie die Ergebnisse im Unterricht Verwendung finden (Department 2005; Tresch 2007). Die Gefahr besteht darin, den Unterricht ganz oder weitgehend auf Test einzustellen. 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, ob die Tests, wenn sie schon da sind, zur schulischen Selektion genutzt werden oder wirklich nur der Förderung dienen.
- Beachtet werden sollte auch, 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).⁵
- 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 2008). 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.

⁵ 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.

Vorteile von testbasierten Leistungsmessungen wären etwa:

- Objektivierung der Notengebung
- 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 überprüfbaren Resultaten führen. Das im Kern ist unstrittig, anders würden die Gymnasien und die Berufsschulen 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 „Bildung“ keinen Keil treiben und nicht von einer Unverträglichkeitserklärung ausgehen. Vielmehr gilt es, die Chancen der Entwicklung zu nutzen und die Risiken zu minimieren. Wenn Sie diesen Prozess heute einleiten, sind Sie auf dem richtigen Weg.

Ein konkretes Beispiel ist das Zürcher Projekt HSGYM: Hinter diesem etwas kryptischen Titel - er heisst einfach „Hochschule und Gymnasium - verbirgt sich der Versuch einer inhaltlichen Abstimmung. Fachgruppen, die aus den Vertretern aller Gymnasialfächer und der dazu gehörigen Universitätsfächer zusammengesetzt waren, haben fachliche Standards erarbeitet, die am Ende der Gymnasialzeit und so bei Eintritt in das Studium erreicht sein sollen. So ist für eine curriculare Festlegung gesorgt, hinter der beide Seiten stehen und die die Schule implementieren können. Damit wird ausgeschlossen, dass einfach nur Vorwürfe die Runde machen. Das Projekt ist gerade abgeschlossen worden und gibt einen ausgearbeiteten Rahmen vor, der auf Dialog beruht und keinen staatlichen Lehrplan darstellt. Möglich war das übrigens, weil der Ausdruck „Bildungsstandards“ vermieden wurde.

Und Rankings? Die Aufregung um die ETH-Studie, die ein Ranking war und keines sein sollte, entstand nur deswegen, weil ein gezielter Tabubruch vermutet wurde und keine unglücklichen Präsentation von an sich sehr vorzeigbaren Ergebnissen. Die Studie nämlich bescheinigt den Gymnasien eine sehr hohe Qualität mit vergleichsweise geringen Unterschieden zwischen den einzelnen Schulen. Die Rangierung in Form einer numerischen Abstufung dargestellt in Plätzen wie in einer Fussballigatabelle wird dem nicht gerecht. Das gleiche gilt für die vor einer Woche vorgestellte fünfte Ehemaligenbefragung im Kanton Zürich. Auch ist zu beachten, dass diese Befragungen keine Leistungsdaten enthalten und schon von daher nur sehr begrenzte Aussagen zulassen.

Gehören nun aber Rankings zu den pädagogischen Konsequenzen für die Sekundarstufe II, zu denen ich ja etwas sagen sollte? Rankings machen nur dann Sinn, wenn sie die Leistungsfähigkeit der Schulen wirklich erfassen und abbilden können. In der Schweiz ist das wegen der grossen Unterschiede zwischen den Schulen allein im Blick auf die Ressourcen besonders schwierig. Modelle, die wohl vergleichen, aber auch die Unterschiede richtig erfassen und so fair sein können, wären erst zu entwickeln. Wo es Schul-Rankings gibt, etwa in England und Wales, werden Gesamtschulen verglichen und doch entsteht jährlich Streit über die Fairness der Beurteilung. Die Idee wird sich an diesen Erfahrungen messen

lassen und auch an der Frage, ob wirklich Rankings die Schulwahl beeinflussen oder bei geringen Qualitätsunterschieden und gleicher Berechtigung nicht vielmehr doch die S-Bahn.

Erlauben sie mir abschliessend noch eine Bemerkung zum Unterricht: Wenn es um die Entwicklung der Unterrichtskultur der Gymnasien geht, dann müssen die Ziele der Schulen und die verschiedenen Lernformen in ein vernünftiges Verhältnis gesetzt werden. Was immer etwas abfällig „Frontalunterricht“ genannt wird, ist eine sehr ökonomische Form der Zielerreichung, die ergänzt und auch als Form weiterentwickelt werden kann, ohne dass es klug wäre, damit einen unüberbrückbaren Gegensatz zur Projektmethode oder zum „Selbstlernsemester“ zu verbinden. Der Einsatz der Lernformen erfolgt pragmatisch, und er muss zielverträglich sein.

In den Lehrgängen der Gymnasien und Berufsschulen sind nur begrenzt Phasen des Ausprobierens möglich. Neue elektronische Formate des Unterrichts wird es in Zukunft geben. Für sie gilt dasselbe wie für die erweiterte Lehr- und Lernformen des Gymnasiums. Sie müssen sich im Blick auf Gütekriterien als sinnvoll herausstellen und sind kein Selbstzweck. Mathematik ist über Jahrhunderte erfolgreich mit den Medien Tafel, Kreise, Schwamm unterrichtet worden. Wenn es heute bessere Möglichkeiten gibt, wird die Schule sie nutzen. Aber einfach nur die Schülerinnen und Schüler mit einem Laptop auszurüsten, ist noch keine Verbesserung der Unterrichtskultur.

Literatur

Bericht und Antrag des Regierungsrates des Kantons Zürich an den Kantonsrat zum Postulat KR-Nr. 54/2003 betreffend Bericht zur Situationsanalyse der Sekundarstufe I vom 20. Juni 2005.

Criblez, L. (Hrsg.): Bildungsraum Schweiz. Historische Entwicklung und aktuelle Herausforderung, Bern/Stuttgart: Haupt-Verlag 2008.

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.

Glaser, R.: The Future of Testing. A Research Agenda for Cognitive Psychology and Psychometrics. In: American Psychologist Vol. 36 (September 1981), S. 923-936.

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.

Kronig, W.: Die systematische Zufälligkeit des Bildungserfolgs. Theoretische Erklärungen und empirische Untersuchungen zu Lernentwicklung und zur Leistungsbewertung in unterschiedlichen Schulklassen., Bern/Stuttgart/Wien: Haupt Verlag 2007.

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.

- Moser, U./ Stamm, M./Hollenweger, J.: Für die Schule bereit? Lesen, Wortschatz, und soziale Kompetenzen bei Schuleintritt. Oberentfelden: Sauerländer Verlage AG 2005.
- Moser, U./Keller, F./ Tresch, S.: Schullaufbahn und Leistung. Bildungserfolg und Lernverlauf von Zürcher Schülerinnen und Schülern am Ende der dritten Volksschulklasse. Bern: h.e.p. Verlag 2008.
- Oelkers, J.: Zum Problem von Standards aus historischer Sicht. In: Neue Sammlung 44 (2004), S. 179-200.
- Oelkers, J.: Gesamtschule in Deutschland. Eine historische Analyse und ein Ausweg aus dem Dilemma. Weinheim/Basel: Beltz Verlag 2006.
- Oelkers, J.: Die Qualität der Gymnasien. Eine Expertise für den Kanton Zürich. Bern: h.e.p. Verlag 2008.
- Oelkers, J./Reusser, K.: Expertise „Qualität entwickeln - Standards sichern - mit Differenz umgehen.“ Berlin: Bundesministerium für Bildung und Wissenschaft 2008.
- 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.
- Volksschulgesetz des Kantons Zürich vom 7. Februar 2005.