

„Output“ statt „Input“?

Zu den Hoffnungen auf Bildungsstandards, Kerncurricula, Leistungstests¹

Die Ergebnisse der internationalen Leistungsstudien TIMSS und PISA haben die Öffentlichkeit alarmiert: Deutsche SchülerInnen und Schulen scheinen am Ende der Schulzeit nicht genug zu wissen und zu können, und zudem streuen die Leistungen breiter als in anderen Ländern. Vor allem SchülerInnen mit Migrationshintergrund und sozial benachteiligte Kinder schneiden deutlich schlechter ab. Schuld sei das Steuerungssystem, wird verkündet, und es werden auch gleich drei Rezepte angeboten, die Misere zu überwinden:

- *Kerncurricula*, die Lehrpläne entfrachten und auf wesentliche Inhalte konzentrieren, sollen eine gemeinsame Grundbildung für alle sichern.
- *Bildungsstandards*, die zentrale Kompetenzen bestimmen, sollen als verbindlicher Maßstab den Unterricht orientieren und seine Qualität fördern.
- Die Kontrolle der Ergebnisse durch regelmäßige flächendeckende *Vergleichstests* sollen ein Mindest- oder Regelniveau in grundlegenden Leistungen („literacy“) garantieren.

Bleibt die Frage, wie man zu solchen Vorgaben kommt – und welche Wirkungen sie tatsächlich haben werden.

„Kerncurricula“: Was ist Bildung ?

Wenn man Romane aus der Zeit um die Jahrhundertwende liest, findet man häufig das alte Bildungsideal oft bis zur Karikatur verbogen: formal gute Manieren, geflügelte Zitate im Munde, Klavier spielend, enzyklopädisch breit informiert, kurz: den Konventionen angepasst und zumindest verbal auf dem Stande geltenden Wissens und Könnens. Wer heutige Quizshows anschaut, merkt, dass sich der gesellschaftliche Fortschritt in Grenzen hält.

Die wohl gemeinte Aufgabe der Schule war aus dieser Sicht: Tradition des Bewährten, d. h. Sicherung der alten Werte und Vermittlung eines Kanons von Bildungsgütern. Dahinter steckt die Vorstellung einer *materialen* Bildung. Es ging um Inhalte, die von Generation zu Generation weitergegeben, von den Gebildeten den Ungebildeten vermittelt werden.

Die gesellschaftliche Instabilität, manche sprachen von einem rasanten „Fortschritt“, die Explosion des Wissens in der Neuzeit und eine veränderte Sicht des Menschen aus der Aufklärung, vor allem das ihm zugestandene Recht zur Entfaltung seiner persönlichen „Kräfte“, förderte schon im 18. Jahrhundert die Entwicklung der Idee einer *formalen* Bildung. Auf eine Kurzformel gebracht: Die Schule soll das Denken vermitteln, die Inhalte haben nur eine instrumentelle Funktion. Zum Teil waren die Inhalte dieselben, z. B. die lateinischen Klassiker, aber es waren nicht mehr die Aussagen der Texte, die Besitz des einzelnen werden sollten. Im

¹ Teile dieses Beitrags sind ausführlicher entfaltet in: Brügelmann, H.: Kerncurricula, Bildungsstandards und Leistungstests: Zur unvergänglichen Hoffnung auf die gute Schule „von oben“. In: Vierteljahrsschrift für wissenschaftliche Pädagogik (2004, i.D.).

Vordergrund stand nun – wie aus der Sicht einiger heute noch bei der Mathematik – die Entwicklung formaler Konzepte, z. B. grammatischer Strukturen, die über den konkreten Inhalt hinaus von Bedeutung sein sollten.

Durch Latein lernt man denken. Mit diesem Argument haben sich die Altphilologen seit der Anfechtung ihres Fachs im 19. Jahrhundert dagegen gewehrt, dass Latein aus dem Kanon des Gymnasiums verschwindet. In Deutschland erfolgreich, in den USA nicht. Dort wurde die kühne Behauptung schon in den 1920er Jahren auf den empirischen Prüfstand gestellt. Robert Thorndike verglich in zwei Studien die Leistungen von SchülerInnen mit und ohne Lateinunterricht und kam zu dem ernüchternden Ergebnis, dass sich keine Effekte auf die allgemeine Denkfähigkeit zeigten (vgl. die Zusammenfassung bei Strunz 2003, 562).

Damit erhielt die Theorie der formalen Bildung einen erheblichen Dämpfer. Das Gehirn ist kein Muskel, der durch Übung an beliebigen Gegenständen gekräftigt werden kann. Unser Wissen und Können ist bereichsspezifisch, d. h. in hohem Maße gebunden an konkrete Inhalte und bestimmte Verwendungskontexte. In Deutschland bestätigt eine neuere Studie die Ergebnisse aus den USA. Auch Haag/ Stern (2000) stellen fest, dass Lateiner weder in allgemeinen Problemlöseaufgaben noch in spezifische Denkleistungen besser abschneiden als Nichtlateiner – wenn man die Eingangsvoraussetzungen kontrolliert.

Der von den Lateinern, von MathematikerInnen und anderen erhoffte Transfer auf die Denkfähigkeit hat sich nicht nachweisen lassen. Trotzdem kann man am Konkreten Allgemeines lernen. Aber dieses Allgemeine muss thematisiert werden. Und das kann an demselben Gegenstand ganz Unterschiedliches sein. An dem Gedicht „Der Erlkönig“ lassen sich das Versmaß besprechen, die Komposition der Strophenfolge, die sprachlichen Bilder, die inhaltliche Thematik, der kulturelle Kontext. Je nach Blickwinkel können SchülerInnen etwas Anderes lernen. Darum ist die Wahl der Perspektive eine didaktisch zu begründende Entscheidung über die alternativen Kategorien, die an einem bestimmten Inhalt erschlossen werden können.

Noch schwieriger wird es, wenn man das Verhältnis von Gegenstand und Perspektive umkehrt. Es gilt nämlich auch: Dieselbe Kategorie lässt sich an verschiedenen Inhalten erarbeiten: ein bestimmtes Versmaß an unterschiedlichen Gedichten, Besonderheiten der Romantik an einem Gedicht, an einem Roman, an einem Musikstück.

Diese Einsicht stellt den Sinn der wieder aufgeflammten Kanon-Debatte in Frage: Warum müssen alle SchülerInnen dieselben „100 wichtigsten Bücher“ (DIE ZEIT) lesen? Eine solche Vorgabe unterstellt, dass zentrale Erfahrungen nur in der Auseinandersetzung mit *bestimmten* „Kulturgütern“ gewonnen werden können. Noch fraglicher ist die zweite Annahme, dass verschiedene Menschen in der Begegnung mit einem Inhalt auch tatsächlich dieselben Erfahrungen machen. Hat man die Vorstellung aufgegeben, dass Wahrnehmung ein passiver Vorgang ist, ist diese Annahme nicht mehr zu halten. Das Diktum von Marcel Proust, jeder Leser eines Buches sei einer Leser seiner selbst, bringt die Gegenposition auf den Punkt. Über einen Kanon ausgewählter Inhalte lässt sich nicht steuern, welche Vorstellungen in den Köpfen junger Menschen entstehen.

Wenn verschiedene Menschen „Dasselbe“ an ganz unterschiedlichen Inhalten und auch auf verschiedenen Wegen lernen können, macht ein Kerncurriculum keinen Sinn, in dem nicht nur ein Kanon von Inhalten vorgegeben wird, sondern auch noch die Abfolge, in der diese Inhalte zu bearbeiten sind.

Ist die Alternative ein *anything goes*? Ja und nein. Schule verfehlt ihre Aufgabe als Ort der Bekehrung oder der Belehrung. Sie hat aber eine zentrale Funktion als Forum der Begegnung. Nur so können Heranwachsende Distanz zur Zufälligkeit ihres familiären Milieus entwickeln. Bildung entsteht in der Auseinandersetzung des Einzelnen mit neuen Erfahrungen und fremden Sichtweisen.

Warum lässt man SchülerInnen nicht selbst wählen, welches Gedicht, welche Tiere, welches physikalische Problem sie interessieren? Im Austausch mit Anderen über die von *diesen* gewählten Gedichte, Tiere, physikalischen Probleme und in der wechselseitigen Diskussion der fremden Sichtweisen auf das *eigene* Gedicht, Tier, physikalische Problem könnten sich SchülerInnen allgemein bilden, ohne dass alle Schulen eines Bundeslands in derselben Woche auf dieselben Themen verpflichtet werden. Die Schweizer Didaktiker Peter Gallin und Urs Ruf (1990; 1995) haben diesen Gedanken auf die klare Formel gebracht: vom Singulären über das Divergierende zum Regulären.

Ein „Kerncurriculum... definiert einen Wissensbestand, der allen Kindern vermittelt werden muss“, formuliert dagegen Böttcher (2002, 26). Entfrachtung der Lehrpläne und Sicherung einer gemeinsamen Grundbildung sind die Hoffnungen, die mit der Forderung nach Kerncurricula verbunden werden. Das erste dieser beiden Ziele kann auch durch eine Individualisierung „von unten“ erreicht werden. Das zweite ist eine Illusion, wenn man darunter eine Gleichschaltung der Köpfe versteht. Sie ist pädagogisch nicht zu rechtfertigen und als Absicht auch den VertreterInnen des Kerncurriculum-Konzepts nicht zu unterstellen (vgl. neben Hirsch 1997 vor allem die Beiträge zu Böttcher/ Kalb 2002). Bildung als Erweiterung des eigenen Repertoires an Denk- und Handlungsmöglichkeiten muss an dem anknüpfen, was die einzelnen aus ihrer Lernbiografie mitbringen. Und das ist schon bei den Sechsjährigen erstaunlich unterschiedlich.

Aber kann man die Einheitlichkeit des Unterrichts nicht dadurch verbessern, dass man statt der Inhalte – oder neben ihnen auch – die *Lernergebnisse* vorschreibt?

„Leistungsstandards“:

Brauchen wir eine Output- statt einer Input-Kontrolle der Unterrichtsqualität?

Die Probleme sind offenkundig: Viele SchülerInnen verlassen die Schule, ohne zureichend für die Anforderungen der Berufswelt, des öffentlichen Lebens und des privaten Alltags gerüstet zu sein. Die Kluft zwischen leistungsstarken und leistungsschwachen SchülerInnen ist immens, und manche, die in ihrer Klasse zur Leistungsspitze gehört, würden in einer anderen Schule sitzen bleiben. Erst recht leiden SchülerInnen, die mit ihren Eltern in ein anderes Bundesland umziehen, unter der mangelnden Einheitlichkeit unseres Bildungswesens

Die *Input*-Steuerung des deutschen Bildungswesens habe versagt – so der seit PISA gebetsmühlenartig wiederholte Topos in der Reformdiskussion. Gemeint sind mit *Input*-

Steuerung sowohl Richtlinien und fachbezogene Lehrpläne als auch Qualifikationen von LehrerInnen, wie sie in Laufbahnprüfungen eingefordert werden. Zur *Input*-Steuerung zählt schließlich auch die Kontrolle des Unterrichts durch die Schulaufsicht, z. B. wenn es um die Bewerbung auf Funktionsstellen geht. Während in anderen Ländern die Fachleistungen der SchülerInnen, also der *Output*, entscheidendes Kriterium für die Leistungsfähigkeit des Systems seien, begnüge man sich in Deutschland damit, die Randbedingungen für Unterricht zu definieren – ohne zu wissen, ob z. B. Klassengröße, Studentafel bzw. bestimmte Methoden tatsächlich die erhofften Effekte hätten. So die Kritik. Dass die Leistungen den Erwartungen nicht entsprächen, zeigten die internationalen Ländervergleiche wie TIMSS und PISA.

„Wegen PISA“ werden immer neue Maßnahmen ergriffen. Zu ihnen zählt auch die Vereinbarung von Bildungsstandards. Redlicher sollte man nur von *Leistungsstandards* sprechen, denn die Vorgaben und die Tests zu ihrer Kontrolle beschränken sich auf wenige Leistungsbereiche in ausgewählten Fächern. Die Logik der ReformierInnen ist einfach: Andere Länder, die im internationalen Vergleich besser abgeschnitten haben, verfügen über landesweit verbindliche Leistungsstandards. Führen wir diese auch in Deutschland ein, sollten deutsche SchülerInnen beim nächsten PISA-Test genauso gut abschneiden wie Finnland, wie England und wie die USA.

So einfach, so falsch.

Denn anders als die 15-Jährigen haben die deutschen GrundschülerInnen in der internationalen Lesestudie IGLU überdurchschnittlich *gut* abgeschnitten. Und das, obwohl die deutschen Grundschulen – wie die Sekundarschulen – keine übergreifenden Bildungsstandards haben. Dasselbe Steuerungssystem also, aber ganz unterschiedliche Ergebnisse verschiedener Schulstufen innerhalb dieses Systems. Die fehlenden Standards können demnach nicht die Ursache sein, dass die 15-jährigen SchülerInnen so viel schlechtere Fachleistungen erbracht haben als ihre AltersgenossInnen in anderen Ländern.

Lassen sich Standards als Mindest- oder gar Regelanforderungen für alle vorschreiben?

Außerdem sollte man sich die angeblichen Vorbilder in anderen Ländern genauer anschauen. Denn Standards sind nicht gleich Standards. Schon der genauere Blick auf die deutschen Verhältnisse zeigt genügend Schwierigkeiten. Die Bundesländer definieren sie in den neuen KMK-Vorschlägen als Leistungs*niveaus*, die zu bestimmten Terminen zu erreichen sind: „Die Bildungsstandards legen fest, welche *Kompetenzen* die Kinder oder Jugendlichen bis zu einer bestimmten Jahrgangsstufe mindestens erworben haben sollen“ (Klieme u. a. 2003, 4). Auf den ersten Blick macht das Sinn: Alle SchülerInnen sollen, unabhängig von ihrer familiären Herkunft und von den besonderen Bedingungen ihrer Schule, dieselben Fähigkeiten und Kenntnisse erwerben. Aber ist das realistisch? Formuliert als mittlere Anforderungen, verurteilen solche „Regelstandards“ viele SchülerInnen von Vorneherein zum Scheitern.

Unsere Lesestudie LUST in NRW zeigt, dass die schwachen 10% eines Jahrgangs den starken 10% von Klasse 2 bis Klasse 4 um drei bis vier Schuljahre hinterherhinken (Brügelmann 2003). Die IGLU-Daten machen deutlich, dass die Streuung in anderen Ländern noch wesentlich höher ist (Bos u. a. 2003). Studien zum Schulanfang belegen, dass diese Bandbreite bereits in den

unterschiedlichen Vorerfahrungen der Kinder begründet ist (Brügelmann 1983, 203). Wie will man angesichts dieser Heterogenität sinnvolle Schwellenwerte für Ende 2. oder 4. Klasse festlegen, ohne die einen heillos zu unterfordern und die anderen von vornherein ins Abseits zu stellen?

Auch wenn man die Qualität von Unterricht fair bewerten will, müsste man wegen der unterschiedlichen Bedingungen in verschiedenen Klassen den Lernfortschritt der SchülerInnen von ihren jeweiligen Voraussetzungen her erfassen. Auf den Schulbezirk bezogene sozio-ökonomische Indikatoren sind zu grob. Man denke nur an die unterschiedliche Zusammensetzung von Parallelklassen in derselben Schule, deren eine die Kinder einer Plattenbausiedlung, die andere die Kinder aus dem Viertel mit Eigenheimen aufnimmt.

Referenzdaten zu Lernzuwachsen in verschiedenen Jahrgängen und Leistungsgruppen gibt es aber kaum. Hier besteht ein großer Forschungsbedarf, wenn man prozessbezogene Maßstäbe für den Lernerfolg gewinnen will. Erste Untersuchungen zeigen, dass die meisten Leistungsgruppen über die Schuljahre hinweg vergleichbare Fortschritte machen (sog. „Karawanen-Effekt“, vgl. Brügelmann/ Backhaus 2003). Die Leistungsunterschiede scheinen die individuelle Entwicklung also nicht zu behindern – es sei denn, man stellt Leistungshürden ohne Bezug auf das jeweilige Ausgangsniveau auf, wie die KMK-Standards sie vorsehen. Aber auch für Lernzuwächse gilt: Sind sie allein auf die Arbeit der Lehrperson zurückzuführen, hat sie wirklich in der Hand, ob ein Kind sich positiv entwickelt oder nicht?

Ein weiteres Problem kommt hinzu. Niemand kann wirklich begründen, auf welchem Niveau notwendige *Basisqualifikationen* zu definieren sind. Weder für den Erfolg in den höheren Schulstufen noch für die Bewältigung von Alltagsanforderungen gibt es empirisch belegte Schwellenwerte. In einer Pilotstudie haben wir unseren Test grundlegender Lesefähigkeiten in einer Handwerkskammer durchgeführt. Nach den dort erhobenen Leistungen könnten schon in der vierten Klasse fast 90% der GrundschülerInnen in Meister(!)kursen mithalten – selbst wenn man dort die schwächsten 5% außer Acht lässt. Kein Wunder, dass die Urteile von ExpertInnen, wie tragfähige Grundlagen für die Lesenwicklung am Ende der Grundschule oder der Pflichtschulzeit zu bestimmen sind, erheblich auseinander klaffen.

Vergleichende „Leistungstests“: Wer trägt die Folgen der Nichterfüllung?

Vor allem aber: Was soll geschehen, wenn die Mindestanforderungen nicht erfüllt werden? Die KMK-Gutachtergruppe um Eckart Klieme hat zu Recht darauf hingewiesen, dass Leistungsstandards ein bildungspolitisches Instrument sind, um die Leistungsfähigkeit des *Bildungssystems insgesamt* zu beobachten. Es wäre kontraproduktiv, ihre Nichterfüllung mit Sanktionen für Schulen, LehrerInnen und SchülerInnen zu verknüpfen. Ein Blick in die USA zeigt, dass die Schulleistungen in Bundesstaaten mit hohen Sanktionen für einzelne Institutionen und/ oder Personen schlechter sind als in Bundesstaaten mit niedrigen Sanktionen: Über dem nationalen Durchschnitt liegen 60-70% der Bundesstaaten mit niedrigen Sanktionen, aber nur 10-20% der Staaten mit hohen Sanktionen (Sacks 1999, 89-90). Und wo die Leistungen in standard-bezogenen Tests zunehmen – fallen sie mehrheitlich in unabhängigen Tests ab (Amrein/ Berliner 2002). Weil sich die Leistungen nicht so einfach steigern lassen, wie die BildungspolitikerInnen versprochen haben, werden in einzelnen US-

Bundestaaten bereits die Anforderungen gesenkt, um Sanktionen aus dem „No Child Left Behind“-Gesetz der Bundesregierung zu entgehen (.s. u. a. den Bericht in der New York Times vom 22.5.2003). In England beklagt selbst der Leiter der Evaluationsagentur OFSTED eine starke Einengung des Curriculums auf Mathematik und Sprache und als Folge eines weithin „flat and ordinary“ Lernangebots Niveausenkungen in Fächern wie Geschichte, Geographie, Kunst und Musik (Clare 2004).

Außerdem stellt sich die Frage, ob eine stärkere zentrale Kontrolle tatsächlich die entscheidenden Bedingungen vor Ort in den Griff bekommen kann. Nach der politischen „Wende“ 1989 ergab sich die seltene Gelegenheit, Schulleistungen von SchülerInnen aus zwei ganz verschiedenen Systemen zu vergleichen. In unsere Schreibvergleich Bundesrepublik-DDR haben wir dies für die Rechtschreibung im Diktat und in freien Texten getan (Brügelmann u. a. 1994). Dabei haben wir in der DDR in den Diktaten aus dem geübten Grundwortschatz in der Tat eine geringere Streuung der Fehlerquote festgestellt als in den gleichen Diktaten in der BRD. Aber in freien Texten war die Bandbreite der Rechtschreibleistungen in beiden Systemen etwa gleich groß – obwohl der Rechtschreibunterricht in der DDR bis ins Detail vorgeschrieben und kontrolliert wurde. Auch die internationale Lesestudie der IEA von 1991 stellte für die 9-Jährigen in der DDR eine vergleichbare Streuung fest wie für die gleichaltrige Vergleichsgruppe in der BRD (Lehmann u. a 1995, 21). So überrascht das Ergebnis einer Befragung in Baden-Württemberg nicht, in der mehr als zwei Drittel der LehrerInnen bekannten, ihre Arbeit werde sich durch die Einführung von Bildungsstandards „(eher) nicht“ ändern (Rauin 2003).

Schließlich dürfen negative Nebenwirkungen nicht übersehen werden, die sich aus der Bloßstellung derjenigen, die die gesetzten Ziele nicht erreichen, ergeben können (vgl. Leßmann 2003). In Neuseeland hat eine Studie ausdrücklich negative Effekte der Konkurrenzorientierung zwischen LehrerInnen auf die Qualität des Lernens in den Schulen festgestellt (Ladd/ Fiske 2003). Aus verschiedenen angelsächsischen Ländern wird berichtet, dass LehrerInnen versuchen, das Verfahren zu unterlaufen (Figlio/ Winnicki 2003) – bis hin zu expliziten Betrugsversuchen. In einer Befragung in den USA gaben 80% der LehrerInnen an, dass sie zunehmend mehr Zeit für getestete Fächer und weniger Zeit für nicht getestete Fächer aufwenden (Pedulla u.a 2003). Ein solches „teaching to the test“ verändert nicht nur die Inhalte des Curriculum, sondern auch den Stil des Lernens. Bildungsstandards, die ihren Namen verdienen, müssen deshalb mehr erfassen als nur Fachleistungen. Vor allem müssen sie auch Kriterien für die Qualität von Bildungsprozessen bestimmen, in denen die SchülerInnen sich tatsächlich aneignen können, was gefordert wird.

Bildungsansprüche der SchülerInnen – eine alternative Sicht

Der Grundschulverband hat in seinen „Bildungsansprüchen der Kinder“ deutlich gemacht, wie Unterricht auf zentrale Ziele ausgerichtet werden kann, ohne dass in diesen Entwicklungsdimensionen für alle Kinder gleiche Niveaus zu bestimmten Zeitpunkten formuliert werden, und dass sich die Qualität der angestrebten Lernerfahrungen nicht durch ihre Operationalisierung in Form von Testitems definieren lässt, sondern über eine Beschreibung der vorausgehenden Lernprozesse und –kontexte konkretisiert werden muss. Ein Beispiel zur Illustration(Grundschulverband 2003):

„ Lesen

- Die Kinder lesen Kinderbücher, die sie nach eigenem Interesse ausgewählt haben. Sie lesen selbstvergessen und genussvoll, auch ohne weitere Didaktisierung des Leseertrages.
- Sie finden in Texten gezielt Informationen.
- Sie denken über Texte nach, nehmen zu Gedanken, Handlungen, Personen Stellung; sie entwickeln innere Vorstellungsbilder.
- Sie nutzen verschiedene Medien, gewinnen daraus Anreize zum Schreiben, zum Lesen, zum Gestalten eigener Medienbeiträge.
- Sie praktizieren Methoden:
 - Sie äußern Vermutungen über den weiteren Fortgang,
 - sie finden durch überfliegendes Lesen eine gesuchte Textstelle,
 - sie belegen eigene Aussagen mit Textstellen.

[...]

Bezogen auf das Lesen unterscheiden sich die Kinder erheblich in ihrer Leseerfahrung, ihrer Lesekompetenz und ihren Lesevorlieben. Deshalb muss der Unterricht ein breites Leseangebot bereit halten, Lesezeiten zum individuellen Lesen einbeziehen und Verbindungen von schulischem zu außerschulischem Lesen herstellen. Am Ende der Klasse 4 ergibt sich eine Bandbreite in der Lesekompetenz der Kinder. Es sind Vielleser und Wenigleser in einer Klasse. Das ist zu akzeptieren. Aber alle Kinder verfügen über eine Verstehensfähigkeit, die sich allerdings beim Verstehen von unterschiedlich komplexen Texten bezogen auf deren Inhalt, Satzbau und Sprachmuster unterscheiden kann.

[...]

Lernbedingungen

Die Ziele sind nur erreichbar, wenn der Unterricht folgende Lernbedingungen schafft:

[...]

Schreib-Lese-Kultur: Der Unterricht vermittelt an die Kinder persönlich bedeutsame Begründungen zum Schreiben und zum Lesen, damit verbunden auch zum Rechtschreiben. Dazu ist die Entwicklung einer Schreib-Lese-Kultur in der Klasse mit vielfältigen Schreib- und Leseanregungen auf unterschiedlichem Niveau erforderlich. Für die Normorientierung sind Schreibgelegenheiten nötig, die Texte mit Normanspruch veröffentlichen, z.B. Klassenzeitung, Geschichtenbuch für andere Leser, Lesetipps für die Schulbücherei. Dabei gilt der Normanspruch erst nach dem Entwurf von Texten für deren Überarbeitung und Schlussfassung.

[...]

Lesen: Das Spektrum der Textsorten, der Textthemen und der Schwierigkeitsgrade ist breit gefasst, so dass die unterschiedlichen Vorlieben der Kinder (z.B. zu Texten mit Tieren, zu Science fiction, zu illustrierten Büchern) in der Schule durch ein breites Buch- und Textangebot bedient werden. Genießendes Lesen setzt entsprechende Buchangebote, Lesezeiten und Leseorte auch in der Schule voraus. Handelnde Methoden fördern die aktive Auseinandersetzung mit Texten, hierbei werden auch für die Kinder neuartige Texte und neue Textsorten einbezogen. Die Methoden des Textumgangs werden ebenso bei Nicht-Printmedien angewendet, z.B. bei Hörspielen, Filmen, Literatur-CD-ROMs.“

Es ist wichtig, die Entwicklung grundlegender Leistungen in zentralen Fächern im Blick zu behalten. Fatal kann sich aber auswirken, wenn diese Perspektive in der öffentlichen Aufmerksamkeit so dominiert und so reduziert wird, dass andere zentrale Aspekte von Schulqualität vernachlässigt werden. Genau dies aber ist zur Zeit der Fall.

Wie kann Rechenschaft sinnvoll eingefordert werden?

Aus *pädagogischer* Sicht erweist sich die Qualität einer Schule, eines Unterrichts also nicht nur an ihren „Produkten“, also an den erworbenen Fähigkeiten, Kenntnissen und Einstellungen. Sie wird wesentlich bestimmt durch die Qualität der Prozesse und der Rahmenbedingungen, unter denen SchülerInnen ihre Erfahrungen machen. Hartmut von Hentig hat diesen Anspruch mit dem Zitat von Paul Goodman, die Schule sei zu denken als „a place to grow up“. Das Wachstum einer Person wird herausgefordert und gestützt durch die Kultur, in der sie lebt. Die Schule als ein sozial und materiell gestalteter Raum hat Wirkungen, die sich nicht immer kurzfristig und nicht nur in fachlichen Leistungen erfassen lassen. In der Demokratie darf Schule nicht als bloße Produktionsstätte fachlicher Leistungen verstanden werden. Zudem muss verhindert werden, dass die Bildungsansprüche der Kinder und Jugendlichen durch die Qualifikationserwartungen der Gesellschaft verdrängt werden (Grundschulverband 2003).

Damit kann Rechenschaft nicht einseitig „von unten nach oben“ gefordert werden. Der Grundschulverband hat bereits 1999 für ein Evaluationssystem plädiert, das nach unterschiedlichen Ebenen und Funktionen differenziert wird.

Seit einigen Jahren stehen zentrale Leistungsmessungen im Vordergrund. Sie sind hilfreich für ein sog. „System Monitoring“. *Politik und Verwaltung* können auf diese Weise allgemeine Problemstellen ausmachen und die Wirksamkeit ihrer Maßnahmen überprüfen und allgemeine Problemstellen ausmachen. Für solche Daten reicht es, wenn sie (wie bei PISA und IGLU) an Stichproben und etwa alle fünf Jahre erhoben werden. Auch für einzelne Schulen und Lehrerinnen können die Befunde als Bezugswerte zur Interpretation eigener Daten hilfreich sein.

Die *Schulaufsicht* sollte durch eine Inspektion ergänzt (oder gar ersetzt?) werden, die den Unterrichtsbetrieb nicht rechtlich kontrolliert, sondern fachliche Rückmeldung gibt. Die Schule sollten den „Fremdblick“ durch eine interne Bestandsaufnahme vorbereiten: Was sind unsere Ziele, wo liegen unsere Stärken, welche Probleme haben wir? Und sie sollte die Berichte der externen BeobachterInnen öffentlich kommentieren: Welche Einschätzungen teilen wir, was wollen wir unternehmen, um unsere Arbeit weiter zu verbessern? Ohne Sanktionsbefugnisse kann die Inspektion nur durch ihre Kompetenz und durch den sozialen Druck, den die Berichterstattung bedeutet, wirken.

Als Rahmen für die Gespräche mit einzelnen Kolleginnen könnten die folgenden Fragen hilfreich sein:

- Was sind Ihre wichtigsten Ziele und Prinzipien?
Nachfrage: „Wie stehen Sie zu folgenden Vorgaben der Richtlinien/ Lehrpläne, zu folgenden Positionen der schulpädagogischen und (fach-)didaktischen Diskussion?“
- Wo steht ihre Lerngruppe, wo stehen einzelne Kinder in den zentralen Entwicklungsdimensionen?
Nachfrage: „Ist Ihnen aufgefallen, dass Marc...?“
- Wo sehen Sie die Stärken und die Schwächen Ihrer Arbeit, also Ihrer Versuche, die eigenen Ansprüche und die vorgegebenen Anforderungen umzusetzen?
Nachfrage: „Mir ist bei der Beobachtung Ihres Unterrichts aufgefallen, dass...“
- Welche Umstände erschweren es Ihnen, Ihre Ansprüche im Alltag umzusetzen?
Nachfrage: „Könnte es auch daran liegen, dass ...?“

- Was haben Sie sich als nächste Schritte zur Entwicklung Ihrer Arbeit vorgenommen?
Nachfrage: „Haben Sie auch an folgende Möglichkeiten gedacht:?“
- Welche Unterstützung/ welche Rahmenbedingungen wären nötig bzw. hilfreich?
Nachfrage: „Würde es Ihnen helfen, wenn...?“

Die Grundidee: Die Diskussion über die *Kriterien* für „guten Unterricht“ ist zu trennen von der Frage nach der Qualität seiner *Umsetzung*. Es macht wenig Sinn, den *tatsächlichen* Unterricht mit Ansprüchen zu bewerten, die die Bewerteten gar nicht teilen. Über diese *normativen* Fragen ist vorweg Verständigung zu erzielen.

Mit diesen Fragen ist zugleich schon angedeutet, welche Perspektiven für die Selbstreflexion im *Kollegium* wichtig sind. Aufgabe aller an der Schule Beteiligten ist, ein Schulprogramm mit eigenem Profil zu entwickeln und fortzuschreiben. Die Verständigung über gemeinsame Ziele und Maßnahmen muss dabei ergänzt werden durch eine kontinuierliche Bestandsaufnahme der eigenen Stärken und Schwächen.

Was für das Kollegium bzw. die Schulkonferenz gilt, ist ebenso bedeutsam auf der Ebene der einzelnen Klasse. Selbstevaluation von *Lehrerinnen* und *SchülerInnen* kann sogar als Schlüssel für die Entwicklung des Unterrichts betrachtet werden. Konkret stellen sich drei Aufgaben:

- Die Lehrperson sollte in Abständen ihren eigenen Unterricht und seine Wirkungen überprüfen.
- Sie sollte die Lernprozesse und Leistungen der SchülerInnen beobachten und im Blick auf eine individuelle Unterstützung auswerten.
- Jedes einzelne Kind sollte seine eigene Arbeit einschätzen und die Leistungen anderer kommentieren.

Diese Lernbeobachtung kann durch Materialien und Strukturen von außen gestützt werden. Eine Arbeitsgruppe des Grundschulverbands arbeitet zur Zeit an konkreten Hilfen, deren Intention ich mit einigen Stichworten zum Lernbereich „Sprache“ veranschaulichen möchte.

1 Lernstände feststellen und einordnen	2 Lernentwicklungen dokumentieren und bestätigen
3 Lerngespräche mit anderen führen	4 über eigene Lernwege nachdenken

(1) Lernstände feststellen und einordnen

Zwei Möglichkeiten bieten sich an. Über *punktueller* Erhebungen können gezielt Leistungen erhoben und die Ergebnisse in übergreifenden Erhebungen verortet werden. Auf diese Weise kann die Lehrperson den Stand ihrer Klasse vergleichend einschätzen und zugleich ihre

Maßstäbe für die Bewertung von Leistungen überprüfen. Konkrete Verfahren für das erste Schuljahr sind:

- Das „weiße Blatt“ am Schulanfang und in den ersten Schulmonaten („Was kannst du schon schreiben?“ (vgl. Dehn 1988) ;
- Benennen von Schrift auf Fotos aus der Umwelt („Weißt du, was da steht/ was das heißt?“) (vgl. Brügelmann/ Brinkmann 1998);
- Diktieren von Kunstwörtern („Marswörter“) oder den Kindern schriftlich noch nicht bekannten Wörtern alle drei Monate, um zu sehen, wie weit die Kinder lauttreu schreiben können, also das Prinzip der Buchstabenschrift schon erfasst haben (ebda.);
- Stolperwörter-Lesetest, um Lesetempo und -genauigkeit sowie das inhaltliche Verständnis von Sätzen zu erfassen (ab Ende 1. Klasse bis zum Ende der Grundschulzeit; vgl. Metze 2003);
- Hamburger Schreibprobe (HSP1 – HSP4/5), um die Entwicklung verschiedener Rechtschreibstrategien zu erfassen (vgl. May 2002).

Ergänzend können *begleitende* Beobachtungen und Arbeitsproben der Kinder ausgewertet werden (vgl. Brinkmann/ Brügelmann 1993):

- Kann ein Kind aus dem Sachbuch erzählen, in dem es gelesen hat?
- Wie weit sind seine freien Texte inhaltlich und sprachlich verständlich?
- Kann es selbstständig mit seiner Wörterkartei üben?

Wer die punktuellen Erhebungen selbst auswertet, differenziert dabei seinen Blick für die situativen Beobachtungen und deren Interpretation

(2) Lernentwicklungen dokumentieren und bestätigen

Fortschritte der Kinder können in verschiedenen Formen dokumentiert und gewürdigt werden:

- über einen LesePASS können die Titel der gelesenen Texte/ Bücher ausgewiesen werden;
- mit Hilfe eines Rechtschreibausweises lassen sich die selbst gewählten Termine erfolgreicher Proben festhalten (z. B. weniger als 3 Fehler in einem „Schleichdiktat“);
- in Lesetagebüchern können die persönlichen Bewertungen der individuellen Lektüre von einfachen Angaben (Autor/ Titel) und Einschätzungen (Smileys für die Bewertung unter vorgegebenen Kriterien wie „neues Wissen“, „Spannung“ u. ä.) bis hin zu kleinen Rezensionen gesammelt werden und die Lesebiografie der Kinder dokumentieren (vgl. Brinkmann 2004);
- in einem Rechtschreibforscher-Heft sammeln und ordnen die Kinder Beispielwörter zu Rechtschreibbesonderheiten, z. B. Wörter mit <ih> oder <qu> (vgl. Peschel/ Reinhardt 2001).

Während bei den punktuellen Erhebungen der Vergleich mit allgemeinen Anforderungen und verschiedenen Bezugsgruppen (Gleichaltrige, Mädchen und Jungen, Migrantenkinder) und quantitative Fortschritte im Vordergrund stehen, erlauben die zuletzt genannten Formen eine stärker individualisierte Würdigung der Schriftsprachentwicklung.

Ziel sollte schon hier sein, die Kinder in die Würdigung der Fortschritte und ins Nachdenken über Schwierigkeiten einzubeziehen. Wichtig ist, dass die Kinder sich selbst Ziele setzen und dass sie ihre Arbeit selbstständig planen, nachdem sie sich Klarheit verschafft haben, was sie schon gut können und wo sie noch Schwierigkeiten haben.

Diese Entwicklung kann zusätzlich über die beiden folgenden Unterrichtselemente gestützt werden.

(3) Lerngespräche mit anderen führen

Es sind verschiedene „Institutionen“ im Unterrichtsalltag denkbar, die den Kindern helfen, ihre eigene Arbeit kritisch-konstruktiv zu überprüfen und an den Arbeiten anderer ihre Maßstäbe zu schärfen:

- Schreibkonferenzen, in denen nach bestimmten Regeln Entwürfe vorgestellt, kommentiert und dann mit Hilfe Anderer überarbeitet werden (vgl. Spitta 1998);
- Rechtschreibgespräche in der Klasse, z. B. zum „harten Brocken des Tages“ (Erichson 2004), so dass die Kinder Schwierigkeiten, Hypothesen und Lösungsstrategien austauschen und damit voneinander lernen können;
- Metagespräche über Stärken und Schwächen von Präsentationen vor der Klasse, über Arbeitsergebnissen von Gruppen oder Einzelnen bis hin zu deren Bewertung durch das Plenum nach vereinbarten Kriterien.

(4) Über eigene Lernwege nachdenken

Schwache SchülerInnen haben oft Probleme, Aufgaben strategisch anzugehen und ihre Arbeit auch rein technisch zu organisieren. Zum anderen ist ihnen nur selten bewusst, welche Fortschritte sie schon gemacht haben, auch wenn sie im Vergleich zu ihren KlassenkameradInnen immer noch langsamer sind oder mehr Fehler machen.

Die Reflexion über die eigene Arbeit kann gefördert werden durch

- Portfolios, in denen die Kinder Produkte ihrer Arbeit (mit Datum!) sammeln, so dass Fortschritte im Rückblick unmittelbar sichtbar werden;
- Lerntagebücher, in dem mit Zeichnungen, Texten, Ausschnitten die Auseinandersetzung mit bestimmten Aufgaben prozessual dokumentiert wird (vgl. Gallin/ Ruf 1998; Bartnitzky 2003);
- Selbstzeugnisse, in denen die Kinder ihre Leistungen selbst einschätzen und so die Bewertung durch die Lehrperson ergänzen (Resch 2004, Röbe 2003).

Fazit

Die Beispiele machen deutlich, dass Rechenschaftspflichten sich nicht auf die Systemebene beschränken darf, dass andererseits die Instrumente der Systemevaluation nur begrenzt geeignet sind, um den Evaluationsanforderungen gerecht zu werden, die sich auf anderen Ebenen stellen. Wer die Lernkultur verändern will, muss Evaluation stärker auf die *Prozesse* des Unterrichts und der Lernarbeit fokussieren und er muss die Evaluationskompetenz *im* Unterricht stärken – bei LehrerInnen und SchülerInnen, statt sie auf ExpertInnen nach außen zu verlagern.

Literatur

- Amrein, A. L./ Berliner, D. C. (2002): High-stakes testing, uncertainty, and student learning. In: Education Policy Analysis Archives, Vol. 10, No. 18. [<http://epaa.asu.edu/epaa/v10n18/>].
- Bartnitzky, J. (2003): Einsatz eines Lerntagebuchs in der Grundschule zur Förderung der Lern- und

- Leistungsmotivation. Eine Interventionsstudie. Ms. einer Dissertation. Fakultät 14 der Universität: Dortmund.
- Böttcher, W./ Kalb, P. E. (Hrsg.) (2002): Kerncurriculum. Was Kinder in der Grundschule lernen sollen. Eine Streitschrift. Beltz: Weinheim/ Basel
- Bos, W., u. a. (Hrsg.) (2003): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Waxmann: Münster u. a.,
- Brinkmann, E. (2004): [Lesezeit ... Dok von entw](#)
- Brinkmann, E./ Brügelmann, H. (1993): Ideen-Kiste Schriftsprache 1 (mit didaktischer Einführung „Offenheit mit Sicherheit“). Verlag für pädagogische Medien: Hamburg.
- Brügelmann, H. (1983): Kinder auf dem Weg zur Schrift - eine Fibel für Lehrer und Laien. Ekkehard Faude: Konstanz (7. Auflage 2000).
- Brügelmann, H. (2003): Grundlegende Leseleistungen und der „Karawanen-Effekt“ in der Grundschule. Zentrale Befunde aus dem Projekt LUST an der Universität Siegen. In: Grundschulverband aktuell Nr. 84 (November 2003, 19-25). s. a. www.uni-siegen.de/~agprim/lust
- Brügelmann, H./ Backhaus A. (2003): Lesentwicklung von Grundschulkindern in verschiedenen Leistungsgruppen. Lese-Untersuchung mit dem Stolperwörter-Test (LUST-3). Antrag an die Deutsche Forschungsgemeinschaft. FB 2 der Universität: Siegen. s. a. www.uni-siegen.de/~agprim/lust (Zugriff am 7.2.2004).
- Brügelmann, H./ Brinkmann, E. (1998): Die Schrift erfinden – Beobachtungshilfen und methodische Ideen für einen offenen Anfangsunterricht im Lesen und Schreiben. Libelle: CH-Lengwil.
- Brügelmann, H., Lange, I., u. a. (1994): „Schreibvergleich Bundesrepublik-DDR“ 1990/91. In: Brügelmann/ Richter (1994, 129-134).
- Brügelmann, H./ Richter, S. (Hrsg.) (1994): Wie wir recht schreiben lernen. Zehn Jahre Kinder auf dem Weg zur Schrift. Libelle Verlag: CH-Lengwil (2. Aufl. 1996).
- Clare, J. (2004): Primary league tables „failing pupils“. In: News Telegraph v. 5.2.2004.
- Dehn, M. (1988): Zeit für die Schrift. Lesenlernen und Schreibenkönnen. Kamp: Bochum (4. Aufl. 1994)
- [Erichson 2004](#)
- Figlio, D./ Winicki, J. (2003): Food for thought: The effects of school accountability plans on school nutrition. [NBER Working Paper No. 9319](#).
- Gallin, P./ Ruf, U. (1995): Ich mache das so! Wie machst du es? Das machen wir ab. Lehrmittelverlag des Kantons: Zürich (CVK: Bielefeld).
- Gallin, P./ Ruf, U. (1998): Sprache und Mathematik in der Schule. Auf eigenen Wegen zur Fachkompetenz. Illustriert mit sechzehn Szenen aus der Biographie von Lernenden. Kallmeyer: Seelze (1. Aufl. Zürich 1990).
- Grundschulverband (2003): Bildungsansprüche von Grundschulkindern – Standards zeitgemäßer Grundschularbeit. Empfehlungen des Grundschulverbandes zur aktuellen Standard-Diskussion. Grundschulverband – Arbeitskreis Grundschule e.V. : Frankfurt.
- Haag, L./ Stern, E. (2000): Non scholae sed vitae discimus? Auf der Suche nach globalen und spezifischen Transfereffekten des Lateinunterrichts. In: Zeitschrift für pädagogische Psychologie, 14. Jg., 146-1577.
- Hirsch, E. D. (1997): Cultural literacy. What every American needs to know. Boston.
- Klieme, E., u. a. (2003): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Deutsches Institut für Internationale Pädagogische Forschung: Frankfurt.
- Ladd, H.F. & Fiske, E.B. (2003). Does competition improve teaching and learning? Evidence from New Zealand. Educational Evaluation and Policy Analysis, Vol. 25, No. 1, 97-112.
- Lehmann, R. H., u. a. (1995): Leseverständnis und Lesegewohnheiten deutscher Schülerinnen und Schüler. Beltz: Weinheim/ Basel.
- Leßmann, B. (2003): Einsichten in die US-amerikanische Schulwirklichkeit. Aussichten für PISA-Konsequenzen in Deutschland? Wegweisung oder Warnung?! In: Grundschulunterricht, H. 9/2003, 45-49.
- May, P. (2002): Hamburger Schreibprobe. Diagnose orthographischer Kompetenz. Verlag für Pädagogische Medien: Hamburg.
- Metze, W. (2003): Stolperwörter-Lesetest. Ergebnisse der Stichprobenerhebung (Stand 3.8.2003). Vervielf. Ms. Auch zugänglich über: www.lesetest1-4.de .
- Pedulla, J. J., et al. (2003): Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers. National Board on Educational Testing and Public Policy. Lynch School of Education: Boston College; s.a. www.bc.edu/research/nbetpp/reports.html [Zugriff 2.4.03]
- Peschel, F./ Reinhardt, A. (2001): Das Sprachforscher-Buch. Friedrich: Seelze-Velber.
- Rauin, U. (2003): Bildungsstandards in BW – Sachstand und Anforderungen. In: Die Grünen – Bündnis 90 (2003, 29-38). Die Grünen – Bündnis 90 (Hrsg.) (2003): Bessere Schulen durch Bildungsstandards? Dokumentation einer öffentlichen Anhörung im Landtag von Baden-Württemberg am 21.3.2003.
- Resch, U. (2004): Selbsteinschätzungsbögen. Vervielf. Ms. Grundschule Harmonie: Eitorf.
- Richter, S. (1992): Die Rechtschreibentwicklung im Anfangsunterricht und Möglichkeiten der Vorhersage ihrer Störungen. Phil. Diss. FB 12 der Universität Bremen. Verlag Dr. Kovac: Hamburg,
- Röbe, E. u. a. (2003): Kinderfragebogen „Wie ich lerne – was ich kann“ (2.-4.Schuljahr). Vervielf. Ms. Pädagogische Hochschule: Ludwigsburg.
- Sacks, P. (1999): Standardized minds. The high price of America's testing culture and what we can do to change it. Cambridge, MA: Perseus Publishing.
- Spitta, G. (Hrsg.) (1998): Freies Schreiben -- eigene Wege gehen. Libelle: CH-Lengwil.
- Strunz, F. (2003): Kein Ende um Latein. Der totesagte Park. In: Neue Sammlung, 43. Jg., H.4, 561-581.