

Modernes Bildungsmonitoring

Durch stärkere Output-Kontrolle der Schulen zu höheren Kompetenzniveaus der Schüler?

Die Ausgangssituation

Schon bald nach den deprimierenden Ergebnissen bei den internationalen Vergleichsstudien (TIMSS, Pisa) glaubte man in Deutschland die richtige Antwort gefunden zu haben. In vielen Ländern hatte man das Curriculum durch Einführung von Bildungsstandards reformiert und Input-Kontrolle durch Output-Kontrolle ersetzt. Es war also an der Zeit, sich über die Grenzen der Bundesländer hinweg auf Bildungsstandards zu einigen und ihre Realisierung durch zentrale Abschlussprüfungen, Vergleichsarbeiten und Schulinspektionen zu kontrollieren. Kurz: Man ersetzte die Input-Orientierung durch die Output-Orientierung. Schließlich hatte in der Vergangenheit eine Erhöhung des Inputs wenig gebracht: Eine Verminderung der Klassengröße hatte viel gekostet, ohne dass das Kompetenzniveau der Schüler merklich anstieg. Was lag also näher, als die externen Kontrollen zu verschärfen, die Schwachstellen in den Schulen zu identifizieren und die Schulen bei unbefriedigenden Leistungen zur Rechenschaft zu ziehen. Dazu konnte man moderne Testverfahren einsetzen, führte zusätzlich bei etwa der Hälfte der Lehrer im Rahmen der Schulinspektionen Unterrichtshospitationen durch und konnte auf diese Weise Schwachstellen der Schulen identifizieren. Im Rahmen einer vergrößerten Autonomie waren dann die Schulen selbst für eine Beseitigung der gefundenen Schwachstellen verantwortlich.

Viel sprach für eine solche Bildungsreform: Die Bildungsstandards konnten für mehr Leistungsgerechtigkeit sorgen, vor allem, wenn diese zusätzlich durch Vergleichsarbeiten erläutert wurden. Lehrer hatten nun konkrete handfeste Informationen darüber, welches Leistungsniveau ihre Klasse hatte. Außerdem hatten viele Länder, die bei TIMSS und Pisa viel besser abgeschnitten hatten, schon vor langer Zeit ihr Bildungssystem auf diese Output-Orientierung umgestellt. Waren nicht die guten Leistungen der skandinavischen Staaten auf dieses moderne „Bildungsmonitoring“ zurückzuführen? Und auch von den Niederlanden konnte man lernen, die ja auch gute Ergebnisse bei den internationalen Vergleichsstudien erzielten und die ein neues Modell der Schulinspektion entwickelt hatten, das man weitgehend auf Deutschland übertragen konnte.

Diese Argumentation klingt so überzeugend, dass sie anscheinend keiner Prüfung bedarf. Doch scheint eine Prüfung dringend erforderlich, weil diese „Modernisierung“ des Bildungswesens nicht nur mehr „Leistungsgerechtigkeit“ im Sinne gleiche Leistung - gleiche Note verspricht. Sie soll auch zu einer Erhöhung des Kompetenzniveaus der Schüler führen. Wenn aber dieses weit reichende Versprechen höherer Kompetenzen der Schüler auf tönernen Füßen steht, dann hat dies auch gravierende volkswirtschaftliche Konsequenzen.

Führen moderne Bildungsmonitoringsysteme zu höheren Kompetenzniveaus?

Im Klieme-Gutachten (2003, S. 13) steht: *„Die internationalen Ergebnisse der PISA-Studie, insbesondere die sehr guten Ergebnisse der skandinavischen Staaten und einiger anglo-amerikanischer Staaten, lassen vermuten, dass Länder, die systematische Qualitätssicherung betreiben – sei es durch regelmäßige Schulleistungsstudien oder durch ein dichtes Netz von Schulevaluationen – insgesamt höhere Leistungen erreichen.“*

Es wird hier nicht in Abrede gestellt, dass die Einführung von Bildungsstandards und die Durchführung didaktisch erläuteter Vergleichsarbeiten wichtig sind, um Lehrern eine Grundlage für ihre Zensurpraxis zu geben. Dies kann zu mehr Leistungsgerechtigkeit beitragen. In diesem Artikel geht es um das zweite Versprechen: dass die Einführung eines modernen Systems des Bildungsmonitorings quasi automatisch zu „insgesamt höheren Leistungen“ der Schüler führt.

Dem widerspricht, dass Länder wie die USA und Deutschland, die einseitig, d. h. weitgehend ohne zusätzliche Unterstützungsstrukturen, ein Kontrollsystem etabliert haben, höchstens durchschnittliche Leistungen in internationalen Vergleichsstudien erreichen. Lehrer lernen schnell, wie man ein solches Testkontrollsystem umgehen kann, um drohende Sanktionen zu vermeiden: Man schiebt z. B. schwächere Schüler auf Sonderschulen ab oder lässt sie das Schuljahr wiederholen oder macht ein verschärftes „Teaching to the test“, um auf diese Weise in einem engeren Bereich eine Leistungssteigerung zu erreichen. Mit echter Kompetenzsteigerung hat dies natürlich wenig zu tun.

Kann man aber nicht die skandinavischen Länder als Beleg für die Wirksamkeit eines modernen Systems der Output-Kontrolle anführen? Die skandinavischen Staaten erzielen deutlich höhere Kompetenzniveaus als Deutschland und die USA. Dennoch eignen sich die skandinavischen Länder kaum für einen Beleg der Wirkung eines modernen Monitoringsystems. In diesen Staaten wurden in vorbildlicher Weise Unterstützungsstrukturen aufgebaut, die bei der deutschen und der US-Variante fehlen. Unterstützungsstrukturen bestehen aus vielen einzelnen Bausteinen, z. B. a) die Einstellung von

Förderlehrern mit spezieller Zusatzausbildung, b) die Festlegung von Mindeststandards anstelle von Regelstandards, c) eine andere Zensierpraxis mit der Möglichkeit, Tests mehrfach zu schreiben, bis ein Mindestniveau erreicht wurde, und d) die Verfügbarkeit von Unterrichtsmaterialien, die bezüglich ihrer didaktisch-methodischen Angemessenheit empirisch validiert wurden (vgl. Wellenreuther 2010), und e) eine andere Ausrichtung der Lehrerbildung zusammen mit einer anderen Stellung der Schulleiter. In Schweden erhalten die Schulleiter eine dreijährige Zusatzausbildung erhalten; sie sind von allen Unterrichtspflichten befreit und können sich somit ihren Managementaufgaben widmen. Diese Schulleiter haben z. B. die Aufgabe, sich den Unterricht der Lehrer ihrer Schule anzusehen und haben bei der Einstellung und Entlassung von Lehrern ein Mitspracherecht. Sie haben durch die Unterrichtshospitationen Informationen, um die die Ergebnisse der Vergleichsarbeiten auf diesem Hintergrund zu bewerten. Es spricht viel dafür, dass das hohe Kompetenzniveau, das in skandinavischen Staaten erreicht wird, vor allem auf diese Unterstützungsstrukturen zurück zu führen ist. Man kann diese Argumentation auch in tabellarischer Weise verdeutlichen:

Modernes Bildungsmonitoring, Unterstützungsstrukturen und in der Schule erreichtes Kompetenzniveau

...Unterstützungs-strukturen (z. B. Wiederholbarkeit von Tests, Mindeststandards, Speziallehrer, empirisch validierte Unterrichtsmaterialien, mehr Autonomie der Schulen)	Modernes Bildungsmonitoringsystem: Outputkontrolle durch Vergleichsarbeiten, externe Tests, Schulinspektionen	
	NEIN	JA
nicht voll ausgebaut	NIEDRIG, insbesondere große Gruppe mit sehr niedrigen Leistungen Deutschland früher (allerdings gab es schon früher Schulinspektionen; Besuche durch den Schulrat)	NIEDRIG, insbesondere große Gruppe mit sehr niedrigen Leistungen Deutschland heute USA
voll ausgebaut	HOCH Skandinavische Staaten (hier werden zunehmend Methoden der externen Kontrolle durch Maßnahmen der internen Kontrolle bzw. Evaluation ersetzt werden)	HOCH Skandinavische Staaten

Herrmann (2004) hat schon kurz nach Erscheinen des Bildungsgutachtens zur Einführung nationaler Bildungsstandards gewarnt:

„In der Art und Weise, wie das Projekt „Bildungsstandards“ heute vom BMBF, von der KMK und einzelnen Länderministerien präsentiert wird, handelt es sich also nicht um eine ursachenorientierte PISA-Befund-Bearbeitungsstrategie, sondern ganz im Gegenteil um ein Ablenkungsmanöver. Dies geschieht in völliger Verkennung der Tatsache, dass nicht die vergleichende Erhebung von Schülerleistungen Verbesserungen bewirkt – „Das Schwein wird nicht vom Wiegen, sondern vom Füttern fett“ – sondern eine andere Lern- und Arbeitskultur...“ (S. 9 f.)

In den skandinavischen Staaten ist eine Tendenz erkennbar, das System externer Kontrollen von zentralen Instanzen auf dezentrale Instanzen zu verlagern sowie den relativen Anteil der Selbstevaluation im Vergleich zur Fremdevaluation zu erhöhen. Dennoch sinkt die Leistungsfähigkeit des Bildungssystems nicht, wie Vertreter eines einseitigen Bildungsmonitoringsystems erwarten müssten.

Wer diese Argumentation noch stärker untermauern will, kann direkte empirische Belege suchen, ob eine verstärkte Output-Kontrolle zu höheren Kompetenzniveaus führt. Für solche empirischen Belege gibt es zwei Quellen: a) Forschungen über den Zusammenhang von erhöhter externer Kontrolle und Leistung und b) Empirische Untersuchungen in den USA über die Wirkung eines „high stakes“ Testprogramms.

Erhöhte externe Kontrolle – höhere Leistung?

Hohe Anforderungen können durchaus zu höheren Leistungen beitragen, wenn im Unterricht gute Lernvoraussetzungen geschaffen werden. Bei schwächeren Schülern führen hohe Leistungsanforderungen bei summativen Tests jedoch eher zu niedrigen Leistungen, geringer Motivation und zu Prüfungsangst. Man kann dem entgegen wirken, indem man verstärkt kleine formative Tests schreibt, die in Parallelform auch mehrfach geschrieben werden können, bis der Schüler eine definierte Mindestleistung erbracht hat (vgl. Wellenreuther 2009 b). Auch explizites erklärendes Feedback bei einer möglichst zeitnahen Rückgabe der Tests bzw. bei einem sofortigen Selbstüberprüfen der eigenen Lösungen anhand ausführlicher

Lösungsbögen sowie ein fehlertolerantes Klima führen zu erhöhten Testleistungen (vgl. Harlen 2005; Wellenreuther 2009 a). Diese Art der Nutzung von Tests entspricht allerdings nicht der Praxis, die durch ein modernes System der Output-Orientierung nahe gelegt wird. Rick Stiggins kommentiert die Erwartungen an ein „High-stakes Test-Programm“:

„Seit langer Zeit, beginnend in den 1960er Jahren, ...haben wir geglaubt, dass der Weg zur Schulreform durch Verwendung von mehr und besseren standardisierten Tests geebnet wird. Der Fehler, den wir auf allen Stufen gemacht haben, ist zu glauben, dass einmal im Jahr durchgeführte Tests allein hinreichende Informationen und Motivationen liefern können, um das schulische Lernen zu verbessern.“

Tatsächlich hat dieser Glaube in die Macht standardisierter Tests Bildungspolitikern und Schulführern daran gehindert, sich mit einer vollkommen anderen Art und Weise des Testens auseinanderzusetzen – und zwar mit den von Tag zu Tag durchgeführten Tests – von denen gezeigt werden kann, dass mit ihrer Hilfe erstaunliche Gewinne in der schulischen Leistung erzielt werden können..“ (Stiggins 2004, S. 22 f., Übersetzung M. W.)

Empirische Untersuchungen zur Wirkung eines „high stakes“ Testprogramms in den USA

In diesen Untersuchungen werden entweder Bundesstaaten mit solchen Testprogrammen mit denen ohne solche Testprogramme verglichen oder die Entwicklung von Staaten mit und ohne solche „high-stakes“ Testprogramm über einen längeren Zeitraum verfolgt. Die dazu durchgeführten Studien zeigen entweder sehr geringe Leistungsverbesserungen aufgrund solcher Testprogramme (vgl. Raymond & Hanushek 2003; Rosenshine 2003), wobei die positiven Effekte zum größten Teil durch eine Fokussierung auf die Testinhalte erklärt werden können. Die neueste Untersuchung (vgl. Nichols, Glass & Berliner 2005) kommt hingegen zu negativen Auswirkungen eines solchen Testprogramms. Die Autoren fassen ihre Ergebnisse in folgender Weise zusammen (Übersetzung M. W.):

„Die durchgeführte Studie findet hingegen, dass Druck durch solche Tests fast keine Wirkung auf schulische Leistung ausübt...“

...Diese Analysen zeigten, dass:

- *Bundesstaaten mit einem hohen Anteil von Schülern aus Minoritätsgruppen einen höheren Druck ausüben. Dies lässt vermuten, dass irgendwelche Probleme, die durch diese folgenschweren Tests verursacht werden, disproportional Schüler aus Amerika's Minoritäten betreffen.*
- *Je stärker der durch die Tests ausgeübte Druck ist, umso kleiner ist der Anteil, der von der 8. oder von der 10. Klassenstufe zur 12. Klassenstufe gelangt. Die Studienergebnisse deuten darauf hin, dass mit einer Erhöhung des Drucks durch „high stakes tests“ der Anteil der Schüler, die sitzenbleiben oder die frühzeitig die Schule verlassen, ansteigt.*
- *Ansteigender Testdruck produzierte keinen Lernzuwachs ...*

Die Autoren folgern, dass es keine überzeugenden Daten gibt, nach denen der durch „high-stakes-Tests“ produzierte Druck irgendwelche Vorteile bezüglich der Schülerleistung hat. Sie plädieren für ein Moratorium der Politik, das öffentliche Bildungssystem zu zwingen, sich auf high-stakes Tests zu verlassen.“

Die Förderung leistungsschwacher Schüler in Deutschland

Das zentrale Problem des deutschen Bildungswesens besteht in der unzureichenden Förderung der leistungsschwächeren Schüler. Sowohl bei TIMSS wie bei Pisa wurde festgestellt, dass in Deutschland der Anteil der Schüler, die völlig unzureichend entwickelte Kompetenzen haben, besonders groß ist. Um dem entgegen zu wirken, wurde im Klieme-Gutachten die Einführung von verbindlichen *Mindeststandards* vorgeschlagen. Gleichwohl konnte sich die Bund-Länderkommission nur auf die Einführung von *Regelstandards* einigen. Damit wurde der Status Quo zementiert. Schulen müssen sich nicht darum kümmern, dass ihre Schüler bestimmte Mindestziele erreichen, Lehrern wird dadurch eine Verantwortlichkeit für solche Mindeststandards nicht zugemutet. Das passt zur deutschen Bildungsphilosophie, nach der Schüler für die Leistungserbringung selbst voll verantwortlich sind.

Doch nicht nur Mindeststandards, auch eine längere gemeinsame Beschulung, der Einsatz von Speziallehrern für die Förderung sowie ein anderes System des Testens scheint für die Förderung lernschwacher Schüler wichtig zu sein. Solche Maßnahmen werden in den skandinavischen Staaten umgesetzt und führen dort zu deutlich höheren Kompetenzniveaus der schwächeren Schüler.

Die Illusion der Selbstverbesserung

Modernes Bildungsmonitoring, wie es in Deutschland praktiziert wird, unterstellt eine Selbstheilungsthese: Wenn durch externe Tests oder durch Schulinspektionen möglichst zweifelsfrei festgestellt wurde, wo die Schwachstellen einer Schule sind, dann kann die Schule darauf adaptiv selbstheilend reagieren. So schreibt Maag Merki (2010, S. 153)

„Bezug nehmend auf Modelle der Kybernetik (Wiener 1948), wird davon ausgegangen, dass über die Rückmeldung der Ergebnisse aus den standardisierten Lernstandserhebungen oder den zentralen Abschlussprüfungen sowie einem Vergleich von Ist- und Soll-Ergebnissen bestimmte Handlungsschritte im jeweiligen System realisiert werden, die kurz- oder langfristig zu einer besseren Anpassung der schulischen Prozesse und der Ist-Ergebnisse an den erwarteten Output führen. Zudem wird angenommen, dass Schulen zu einem spezifischen Teil verantwortlich sind für die Lernergebnisse der Schüler/innen, diese somit als Indikator für die Effektivität schulischer Prozesse herangezogen werden können (Hamilton, Stecher & Klein 2002). . .

Die Schulen bzw. Lehrpersonen und Schulleitungen werden dabei als aktive Akteure verstanden, die in der Lage sind, über rezeptive und reflexive Prozesse der Auseinandersetzung mit einer möglichen Ist-Soll-Differenz geeignete Schulentwicklungsmaßnahmen abzuleiten und diese in funktionaler Weise in der schulischen Praxis umzusetzen.“

Hier wird ohne nähere Begründung vorausgesetzt, dass Schulen, Lehrpersonen und Schulleitungen sich selbst am eigenen Schopf aus dem Sumpf ziehen können. Dies ist nichts weiter als frommes Wunschdenken! Vielleicht wird dies schon an folgendem Beispiel klar:

Fallbericht

Es wird durch eine Vergleichsarbeit festgestellt, dass in der vierten Klasse eines sehr engagierten Lehrers, der diese Klasse vier Jahre lang in Deutsch unterrichtete, die Heterogenität der Leseleistungen sehr hoch ist: Es gibt einige ganz ausgezeichnete Leser, aber es gibt auch eine vergleichsweise große Gruppe von Schülern, die noch nicht angemessen sinnverstehend lesen können. Der Schulleiter spricht mit dem Lehrer über dieses Ergebnis. Dabei stellt sich heraus, dass der Lehrer ein entschiedener Vertreter der Methode „Lesen durch Schreiben“ nach Reichen ist. Angesprochen auf die vielen Schüler mit unzureichenden Lesefähigkeiten wendet der Lehrer ein, man könne diesen schwächeren Schülern dieses Wissen nicht überstülpen, wenn sie noch gar nicht in der entsprechenden Entwicklungsphase wären. Außerdem hätte er diese Methode auf der Universität gelernt und hätte bisher damit immer gute Ergebnisse erzielt. Der Schulleiter verweist auf neuere Forschungen zum Lesenlernen, durch welche gezeigt wird, dass die Reichen-Methode gerade für leistungsschwächere Schüler problematisch ist und die belegen, dass frühes systematisches Üben der phonologischen Bewusstheit bei diesen Schülern zu deutlich besseren Ergebnissen führt. Daraufhin wird der Lehrer sehr ungehalten und verbittet sich jegliche Vorschläge, die seine Methodenfreiheit einschränken würden.

Entscheidend: Die Unterrichtskompetenz von Lehrern

In der ersten Phase der Schulforschung wurde festgestellt, dass zwischen Schulen beträchtliche Qualitätsunterschiede bestehen. Der Kompetenzgewinn differierte beträchtlich zwischen den Schulen. Bei genauerer Analyse zeigte sich in der zweiten Phase der Schulforschung, dass diese Differenzen verschwinden, wenn man die soziale Zusammensetzung der Schüler kontrolliert. Für eine Qualitätsverbesserung von Schulen waren diese Ergebnisse wenig ergiebig. In einer dritten Phase der Schulforschung sah man sich die Unterschiede innerhalb der Schulen an, die im Kompetenzzuwachs zwischen vergleichbaren Klassen bestehen. Diese Analysen zeigten, wie entscheidend die Lehrkompetenz ist. Nach Analysen in den USA benötigen durchschnittliche Lehrer für einen Kompetenzzuwachs, der von den besten Lehrern in einem halben Jahr erreicht wird, ein Jahr, die schwächsten Lehrer brauchen dafür sogar zwei Jahre, also viermal so viel Zeit wie die besten Lehrer (Hanushek 2004; Williams 2007). Dies heißt: Es ist illusorisch, von einer Gleichheit der Bildungschancen auszugehen, wenn die Lerngelegenheiten von Lehrer zu Lehrer so unterschiedlich sind. Man muss deshalb an der Erweiterung des Kompetenzniveaus der Lehrer ansetzen, statt die Verantwortung für Minderleistungen dem „schlechten“ Schülermaterial zuzuschreiben.

Doch wie kann die Lehrkompetenz von Lehrern verbessert werden? Ohne gezielte Unterstützungen wird es nicht gehen. Zum Verhältnis von solchen Unterstützungsleistungen in Relation zur Schulinspektion in den Niederlanden schreibt van Bruggen (2006, S. 121):

„Niederländische Schulen haben ein Recht darauf, sich Unterstützungen einzukaufen. Dafür steht im Eintopfbudget Geld zur Verfügung, teilweise ist auch ein „Schulentwicklungsbudget“ vorgeschrieben. Bei einem Netzwerk von ca. 50 Schulbegleitungsdiensten mit insgesamt ca. 2000 Experten kann jede Schule Hilfe, Schulbegleitung, Fortbildung und Coaching, etc. kaufen. . .

Zum Vergleich: Das Inspektorat [das die Schulinspektionen durchführt.] zählt ca. 200 Inspektoren (plus ca. 230 Mitarbeiter); die Unterstützungseinrichtungen haben ca. 2500 Mitarbeiter. Das ist ein absolut passendes Verhältnis.“

Offensichtlich hat man in den Niederlanden erkannt, dass nach der Identifikation von Schwachstellen an Schulen die eigentliche Arbeit erst beginnt. Hier stellt sich die Frage nach den Gesichtspunkten, denen eine effektive Weiterqualifizierung genügen sollte. Viele Hinweise sprechen dafür, dass die Vermittlung des theoretischen Wissens in Lehrerfortbildungen nicht ausreicht, um dieses im Unterricht anwenden zu können. Schon von der Fahrschule her wissen wir, dass das theoretische Wissen vergleichsweise leicht zu erwerben ist. Das Einüben neuer Unterrichtsroutinen in den verschiedenen Unterrichtssituationen benötigt deutlich mehr Zeit und ist auf die Unterstützung und Rückmeldung anderer Personen angewiesen. Wenn die Lehrkompetenz von Lehrern erhöht werden soll, sind deshalb zwei Komponenten zu berücksichtigen: (1) Die Vermittlung der theoretischen Kenntnisse, die eine hohe Unterrichtskompetenz beschreiben. Diese theoretische Ausbildung zur Grundlegung einer breiten Wissensbasis kann in Lehrerfortbildungen, Workshops, und Vorträgen gelegt werden, wobei auch Methoden der Verdeutlichung anhand von Videoanalysen des Unterrichts einbezogen werden können, und (2) die schrittweise Entwicklung der entsprechenden Unterrichtskompetenzen im Rahmen von Lerngemeinschaften an Schulen.

Auswahl von Schwerpunkten für Lehrerfortbildungen

Eine Qualitätssicherung auf der Schulebene kann sich ganz unterschiedliche Schwerpunkte setzen. Solche Schwerpunkte sollten nach der empirischen Evidenz bezüglich ihrer Lernwirksamkeit ausgewählt werden. Folgende Schwerpunkte erfüllen m. E. dieses Kriterium:

- Methoden der Vermeidung der Überlastung des Arbeitsgedächtnisses (z. B. Verwendung von Lösungsbeispielen, Vermeiden des Aufmerksamkeitsteilungseffekts, Nutzung visueller Verdeutlichungen; vgl. Wellenreuther 2009 a; Pashler et al. 2007)
- Methoden der Förderung in einem Bereich, z. B. in Deutsch und Mathematik, Einsatz von Screening-Tests zur Identifikation von Schülern, die zusätzliche Förderung benötigen, Einsatz empirisch validierter Förderprogramme, Verwendung direkter Instruktion sowie von Lösungsbeispielen und visuellen Veranschaulichungen (vgl. Gerster et al. 2009; Helmke 2009; Schipper 2005; Wellenreuther 2009 a, b).
- Adaptives Reagieren auf Lernergebnisse von Schülern: Hier handelt es sich vor allem um Methoden, Leistungsmessungen in den Dienst des Lernens stellen (vgl. Williams 2007; Black, Harrison, Lee, Marshall & Wiliam 2003).
- Methoden eines effektiven Klassenmanagements: Dazu zählen z.B. vorbereitende Aktivitäten der Vorbereitung des Klassenraums, die Vermittlung von Regeln und Routinen sowie die Vermittlung von Handlungsstrategien, die Schüler auf das Unterrichtsgeschehen orientieren und die verfügbare Lernzeit erhöhen (vgl. Kounin 1976. Evertson, Emmer, Clements & Worsham 1994)

Wenn Lehrern die entsprechenden Inhalte und Handlungsrountinen in nachhaltiger Weise vermittelt werden, können alle hier angeführten Schwerpunktprogramme zu einer Steigerung des Kompetenzniveaus der Schüler führen.

Ein Beispiel: Schwerpunktprogramm „adaptives Unterrichten“

Eine Verbesserung des adaptiven alltäglichen Unterrichtens führt zu einer deutlichen Leistungssteigerung der Schüler (vgl. Black, Harrison, Lee, Marshall & Wiliam 2003). Bei diesem Schwerpunkt geht es um Techniken, die meist im Zusammenhang mit Methoden der formativen Leistungsbewertung bzw. des „Assessment for Learning“ diskutiert werden und die zu einer stärkeren Berücksichtigung des tatsächlichen Verständnisses aller Schüler befähigen sollen. Diese Techniken beziehen sich auf den alltäglichen Unterricht, wie er Minute für Minute und Tag für Tag stattfindet (vgl. Leahy, Lyon, Thompson & Wiliam 2005).

Klassenarbeiten eignen sich kaum für ein solches adaptives Nachjustieren des Unterrichts. Dazu werden die Ergebnisse der Klassenarbeiten zu spät den Schülern mitgeteilt. Zudem werden Schülern, welche bestimmte Aufgaben unbefriedigend gelöst haben, nicht an Parallelaufgaben Möglichkeiten geboten, sich erneut zu prüfen. Meist werden die Ergebnisse nur kurz kommentiert. Man weiß jedoch, dass Schüler vor allem durch ausführliche Kommentierungen der richtigen Lösungen aus Fehlern lernen können. Da die Leistung bei Klassenarbeiten abschließend bewertet wird, wird den Schülern die Motivation genommen, sich weiter noch intensiv mit ihren Leistungen auseinander zu setzen; was vielleicht noch interessiert ist, wie gut der Nachbar bzw. Konkurrent abgeschnitten hat (vgl. Butler 1988). Auch wird bei Klassenarbeiten nicht die Leistung bei den einzelnen Aufgaben in Bezug auf die relevanten Bildungsstandards protokolliert. Deshalb können solche spezifischen Informationen auch nicht für eine präzise Protokollierung der Kenntnisse und Kompetenzen genutzt werden. Wer aber an adaptivem Unterricht interessiert ist, muss die spezifischen Leistungen protokollieren.

Adaptives Unterrichten ist auf Informationen über das Verständnis des gerade vermittelten Lehrinhalts angewiesen. Nur wenn Lehrer und Schüler Einblick in das erreichte Verständnis gewinnen, können nötige Anpassungen zwischen Lehrziel und anvisiertem Können erfolgen. In den Aufsätzen der Gruppe um Paul Black und Dylan Wiliam (vgl. Black, Harrison, Lee, Marshall & Wiliam 2003, Leahy et al. 2005 sowie Wiliam 2007) werden folgende Techniken vorgestellt:

(1) *Klärung und Verdeutlichung von Absichten und Bewertungskriterien:* Bei komplexen Aufgabenstellungen ist es sehr schwierig, genau zu beschreiben, welche Maßstäbe an eine gute Ausarbeitung angelegt werden. In solchen Fällen ist es sinnvoll, mehrere fertige Berichte unterschiedlicher Qualität zu verteilen und über deren Qualität mit den Schülern zu diskutieren. Zunächst können die Schüler versuchen, allein oder in Partnerarbeit die Ausarbeitungen nach ihrer Qualität in eine Rangfolge zu bringen und diese Rangfolge zu begründen. Schüler können durch solche Lösungsbeispiele am besten verstehen, was sie lernen sollen und auf welche Punkte sie besonders achten sollen. Man kommt anhand solcher konkreter Beispiele dann dazu, über Bewertungskriterien zu sprechen, an denen sich die Schüler dann im Unterricht orientieren können.

(2) *Führung eines effektiven fragend-entwickelnden Unterrichts:* Lehrer sollten beim Beobachten der Schüler sich weniger darauf konzentrieren, ob die Antwort der Schüler richtig ist oder wie „gut“ die Antwort ist; wichtiger im Lernprozess ist, dass der Lehrer sich bemüht, die Gedanken der Schüler zu verstehen, um darauf mit Lernimpulsen zu reagieren. Vorab sollte sich der Lehrer Fragen überlegen, mit denen leicht Fehlverständnisse aufgedeckt werden können. Nach Leahy et al. (2005, S. 21) kann ein Lehrer z. B. fragen, wie viel Wasser, das eine Pflanze durch die Wurzeln aufnimmt, durch Schwitzen verloren geht, um darauf dann klären zu können, dass der Transport von Wasser dazu dient, Mineralien und Nährstoffe in der Pflanze zu verteilen. Ähnlich produktive Fragen wäre in der Bruchrechnung die Frage, welcher Bruch zwischen $\frac{1}{6}$ und $\frac{1}{7}$ liegt.

Herausfordernde Fragen sind für das Lernen sehr wichtig, wie viele experimentelle Forschungen belegen (vgl. Pashler et al. 2007, S. 29 f.). Um schwierige Fragen stellen zu können, müssen sich Lehrer darüber klar werden, was die kniffligen Fragen einer Unterrichtseinheit sind. Häufig findet der fragend-entwickelnde Unterricht nur zwischen dem Lehrer und einem kleinen Häuflein von 3 – 6 Schülern statt. Dabei stellt der Lehrer eine Frage und wartet darauf wenige Sekunden, um dann einen Schüler aufzurufen, der sich gemeldet hat. Um diese Praxis zu ändern, kann der Lehrer für diese Phasen vereinbaren, dass sich Schüler zu den vom Lehrer gestellten Fragen nicht melden sollen; Schüler sollten nur dann sich melden, wenn sie eine Frage an jemanden richten wollen. Der Lehrer ruft dann jeden Schüler nach einem Zufallsprinzip auf, wobei er tunlichst vermeidet, einen Schüler für die gegebene Antwort bloß zu stellen. Bei komplexen Fragen kann er die Wait-Time Technik verwenden: Er bittet die Schüler, sich mit ihrem Tischnachbarn z. B. drei Minuten lang eine Antwort zu überlegen. Danach ruft der Lehrer dann irgendwelche Schüler auf.

Eine andere Technik besteht darin, Mehrfachwahlfragen zu stellen. Schüler können ihre Antworten durch verschiedene Antwortkarten (z. B. A, B, C, D) mitteilen. Hierdurch kann der Lehrer leicht feststellen, wie der Lernstand zu dem gegebenen Inhalt ist. Wenn er sieht, dass alle Antworten richtig sind, kann er mit dem Inhalt fortfahren; wenn viele falsche Antworten gegeben werden, kann er den jeweiligen Inhalt nochmals auf andere Weise erläutern. Eine andere Möglichkeit wäre, Schüler auf Tafeln ihre Antworten schreiben zu lassen. Die Schüler halten dann auf ein Zeichen des Lehrers die Tafel hoch, so dass sich der Lehrer schnell ein Bild über das Verständnis der Schüler machen kann.

(3) *Lernwirksames Feedback geben:* Wenn Schüler eine Klassenarbeit mit einer Benotung zurückbekommen, dann ist ein optimales Feedback für Schüler mit relativ schlechten Leistungen nicht mehr möglich. Für die Schüler ist die Information über die schlechte Leistung so stark „ich-bedrohend“, dass sie kaum in der Lage sind, noch konkrete Informationen aufzunehmen. Nach Forschungen von Butler (1988) werden Informationen zu den Aufgabenlösungen nur dann aufgenommen, wenn keine Bewertungen im Sinne von Noten gegeben werden.

Feedback sollte zum Denken anregen. Wenn ein Schüler 10 Aufgaben bearbeitet hat, und zwei davon falsch gelöst sind, dann kann man dem Schüler rückmelden, dass zwei der 10 Aufgaben falsch gelöst sind. Er kann dann selbst, u. U. mit seinem Partner, herausfinden, welche Aufgaben falsch gelöst sind und wo der Fehler liegt. Dies bringt sicherlich viel mehr als die Rückmeldung „gut gemacht“. Im Zweifel bringen inhaltliche Rückmeldungen, die Punkte benennen, was gut gelungen ist und was noch Schwierigkeiten bereitet, erheblich mehr als die Rückmeldung, welche Aufgaben richtig oder falsch sind.

Bei komplexen Aufgaben (Aufsatz, Schreiben eines Versuchsprotokolls) sollten die Rückmeldungen Anregungen für eine Überarbeitung der erbrachten Leistung geben. Eine problematische Praxis ist es, immer neue Aufgaben zu stellen und Ausarbeitungen dazu herstellen zu lassen, die alle noch reichlich unvollkommen sind. Wichtig wäre es, Zeit für mehrfache Überarbeitungen einzuplanen. Erst wenn der Schüler nach mehreren Überarbeitungen eine akzeptable Leistung erbracht hat, hat er das Gefühl, solche Aufgaben gut bewältigen zu können.

(4) *Aktivierung der Schüler:* Für einen Lehrer ist es schwierig, nach einer Phase des fragend-entwickelnden Unterrichts abzuschätzen, wie gut die wesentlichen Inhalte von den Schülern verstanden wurden. In der Forschung wird z.B. darauf hingewiesen, dass im Unterricht in der Regel immer noch zu wenig mit konkreten Lösungsbeispielen gearbeitet wird, wobei solche Lösungsbeispiele längere mündliche Erklärungen ersetzen könnten (vgl. Zhu & Simon 1987; Carrol 1994). Auf jeden Fall ist es wichtig, dass der Lehrer in Erfahrung bringt, wie die Schüler ihr Verständnis des Gegenstands selbst be-

urteilen. Dazu können Schüler Ampel-Karten hochhalten: Eine rote Karte bedeutet „Nichts bzw. wenig verstanden“, eine grüne Karte dagegen „habe alles bzw. das Meiste verstanden“. Natürlich könnte man auch drei Farben verwenden, z. B. neben rot und grün noch gelb. Gelb würde dann für „zur Hälfte verstanden / teilweise verstanden“ stehen. Der Lehrer kann dann z. B. nach der Einführung fragen, wie gut das Erklärte verstanden wurde. Wenn dann überwiegend rote Karten hochgehalten werden, wird der Lehrer das Wesentliche nochmals an weiteren Lösungsbeispielen erläutern (wobei man die Vollständigkeit der Lösung variieren sollte: Eines vollständig gelöst, eines halb gelöst, eines ohne Lösung). Wenn aber nur wenige Schüler die rote Karte nehmen, kann er diese Schüler an einem Tisch versammeln, und mit diesen Schülern gemeinsam die ersten Aufgaben eines Arbeitsblatts bearbeiten. Auf jeden Fall wird der Lehrer durch solche Karten darauf aufmerksam gemacht, wie das Verständnis des jeweiligen Lerngegenstandes ist, und kann dadurch Maßnahmen einleiten, die sicherstellen, dass alle Schüler über die Mindestkenntnisse, die für ein Voranschreiten im Lehrplan erforderlich sind, auch verfügen.

(5) *Schüler zum Lernhelfer machen*: Es ist für Schüler schwieriger, die eigenen Leistungen und Arbeiten kritisch zu bewerten als die Arbeit von Mitschülern. Dies kann sich der Lehrer zu Nutze machen, indem er die Schüler wechselseitig die Aufgabenbearbeitungen kontrollieren lässt. Sie sollen dann inhaltlich auf Punkte hinweisen, die schon gut oder weniger gut gelungen sind, um dadurch Hinweise zur Überarbeitung zu geben. Das ist der erste Schritt, um zu lernen, nach welchen Gesichtspunkten die Güte von Leistungen zu beurteilen ist. Das Bewerten bzw. Benoten der Leistung sollte aber dem Lehrer vorbehalten bleiben.

(6) *Informationen zum Lernen für ein adaptives Unterrichten verwenden*: Nach Verwendung der Ampelkarten kann der Lehrer auch Partnergruppen zusammenstellen, von denen der eine den Inhalt gut, der andere ihn noch nicht hinreichend verstanden hat. Wichtig erscheint, dass der Lehrer sicherstellt, dass alle Schüler die relevanten Mindestkompetenzen, auf denen weiteres Lernen aufbaut, erworben haben. Um dies abschließend zu kontrollieren, kann der Lehrer am Ende einen kleinen Test durchführen, nach dem dann entschieden werden kann, ob die Klassenarbeit geschrieben werden kann oder dies noch nicht sinnvoll ist.

Vom theoretischen Wissen zur Weiterentwicklung der Unterrichtskompetenz

Lehrerfortbildungen werden meistens in Form von einwöchigen Kursen oder Workshops angeboten. Dabei wird im Grunde vorausgesetzt, dass Lehrer im Wesentlichen nur über das theoretische Wissen informiert werden müssen, das an vielen Beispielen aus der Unterrichtspraxis zu erläutern ist. Doch reicht ein solches Wissen, auch wenn es an Videobeispielen erläutert wird, bei weitem nicht aus, die Unterrichtskompetenzen der Lehrer in dem gewünschten Sinne weiter zu entwickeln. Der Schritt vom theoretischen Wissen zur Umsetzung in Unterrichtshandeln setzt nicht nur voraus, dass eingefleischte Unterrichtsroutinen schrittweise verändert werden müssen, sie setzen auch bewusste Umplanungen des Unterrichtshandeln voraus, bis diese Handlungen zunehmend automatisiert werden und quasi ohne bewusste Kontrolle ablaufen können. Deshalb muss eine Fortentwicklung der Unterrichtskompetenzen folgende Punkte umfassen:

- (1) Eine zusammenfassende und an vielen Beispielen aus dem Fachunterricht erläuterte Darstellung des Erkenntnisstands (z. B. effektive Förderarbeit in Deutsch, Mathematik).
- (2) Verdeutlichung zentraler Aspekte an Videobeispielen.
- (3) Bildung von Lerngemeinschaften, in denen gemeinsam Unterricht geplant und durchgeführt wird, um danach die Planung zu überarbeiten. Danach an anderen Inhalten erneutes Ausprobieren der relevanten Techniken, gemeinsame Unterrichtsplanung und Erprobung usw.

Von der Fahrschule weiß jeder: Es macht einen Riesenunterschied, sich auf die theoretische Fahrprüfung oder auf die praktische Fahrprüfung vorzubereiten. Das zentrale Problem besteht in der Entwicklung einer nach Situationen differenzierenden Handlungskompetenz. In der Lehrerbildung wurde oft von der Annahme ausgegangen, eine theoretische Information über Handlungsmöglichkeiten im Unterricht könnte eine Änderung von Unterrichtsroutinen bewirken. Diese Annahme ist m. E. falsch. Deshalb müsste in Lerngemeinschaften unter systematischer Berücksichtigung des gewählten Schwerpunkts Unterricht geplant, in verschiedenen Klassen durchgeführt, evaluiert und entsprechend der Evaluation überarbeitet werden. Die überarbeitete Version wird dann erneut erprobt, bis die entsprechenden Unterrichtskompetenzen soweit entwickelt sind, dass sie einem größeren Gremium vorgestellt werden können. Die gemachten Erfahrungen werden dann auf Schulkonferenzen oder in weiteren Workshops ausgetauscht.

Ein theoretisches Reflektieren dessen, was unter pädagogischen oder lernpsychologischen Gesichtspunkten im Unterricht zu berücksichtigen ist, reicht somit nicht aus. Zusätzlich wichtig ist die konkrete Übertragung auf fachliche Inhalte und ein systematisches mehrfaches Erproben der Unterrichtsplanungen im Unterricht im Rahmen von Lerngemeinschaften in Schulen (vgl. dazu Fussangel, Rürup & Gräsel 2010; William 2006). Wie bei der Lesson Study, einer japanischen Methode der gemeinschaftlichen Unterrichtsentwicklung, ist nicht nur das gemeinsame Planen von Unterricht unter bestimmten Gesichtspunkten wichtig. Zusätzlich bedeutsam ist die Anwendung und Überprüfung des Gelernten im Unterricht, wobei

andere Lehrer diesen Unterricht sehen, evaluieren und dazu ausführlich Feedback geben sollten. Außerdem sollten aufgrund der genannten Kritikpunkte die Unterrichtsplanung überarbeitet und diese dann erneut erprobt werden.

Ausblick

Eine nachhaltige Bildungsreform muss an der Unterrichtskompetenz von Lehrern ansetzen. Hier gibt es große Unterschiede. Modernes Bildungsmonitoring hilft im besten Falle, die Lehrer zu identifizieren, deren Lehrertrag problematisch ist. Es unterstellt dann einen Automatismus der Weiterbildung und Selbstheilung, der durch keine Forschungen belegt werden kann. Nach meiner Überzeugung sollte man die Erhöhung der Unterrichtskompetenz zur kollektiven Aufgabe einer jeden Schule machen, der für diese Aufgabe auch die notwendigen Unterstützungen gewährt werden müssten. Modernes Bildungsmonitoring wird nach allem, was wir wissen, ohne die Erhöhung der Unterrichtskompetenz von Lehrern keine Erhöhung der Schülerleistungen bewirken. Stattdessen wird noch mehr negativer Stress und Burnout produziert.

Viele Fragen bedürfen noch weiterer Forschung und Entwicklungsarbeit. Insbesondere fehlt Entwicklungsforschung zur Optimierung der inhaltlichen Vermittlung geeigneter Schwerpunkte für die theoretische Lehrerbildung. Auch müsste in verschiedenen Projekten untersucht werden, unter welchen Bedingungen Lerngemeinschaften am besten ihre Unterrichtskompetenz weiterentwickeln können. Ich bin der festen Überzeugung, dass ein konkretes Arbeiten an diesen Projekten eine nachhaltige Bildungsreform eher ermöglicht als das in Deutschland bislang praktizierte „moderne“ Bildungsmonitoring.

Literatur

- Black, P., Harrison, C., Lee, C., Marshall, B., William, D. (2003): *Assessment for Learning. Putting it into practice*. Open University Press, Maidenhead.
- Bruggen, Johan van (2006): *Schulinspektion in den Niederlanden. Metaevaluation und punktuelle Tiefe auf Wunsch der Schule*. In: Buchen, Herbert, Horster, Leonhard und Rolff, Hans-Günther (2006): *Schulinspektion und Schulleitung*. Raabe: Berlin, S. 107- 124..
- Butler, R. (1988): *Enhancing and undermining intrinsic motivation: the effects of task-involving and ego-involving evaluation on interest and performance*. *The British Journal of Educational Psychology*, 58,1-14.
- Carroll, W.M.: *Using Worked Examples as an Instructional Support in the Algebra Classroom*. *Journal of Educational Psychology*, 1994, Vol. 86, No. 3, 360-367.
- Evertson, C.M., Emmer, E.T., Clements, B.S. & Worsham, M. E. (1994³): *Classroom Management for Elementary Teachers*. Boston, Allyn & Bacon.
- Fussangel, K., Rürup, M. & Gräsel, C. (2010): *Lehrerfortbildung als Unterstützungssystem*. In: Altrichter, H. & Maag Mercki, K. (Hrsg.; 2010): *Handbuch Neue Steuerung im Schulsystem*. VS Verlag für Sozialwissenschaften, Wiesbaden, S. 327 - 354.
- Hanushek, E. A (2004): *Some simple Analytics of School Quality*. Working Paper 10229, <http://www.nber.org/papers/w10229>, National Bureau of Economic research.
- Harlen, W. (2005): *Teacher's summative practices and assessment for learning – tensions and synergies*. *The Curriculum Journal*. Vol. 16, No. 2, June, S. 207 – 223.
- Helmke, Andreas (2003): *Unterrichtsqualität – erfassen, bewerten, verbessern*. Seelze.
- Herrmann, U. (2004): *Fördern „Bildungsstandards“ die allgemeine Schulbildung*. Vortrag, über das Internet zugänglich.
- Klieme, E./Avenarius, H./Blum, W./Döbrich, P./Gruber, H./Prenzel, M./Reiss, K./Riquarts, K./Rost, J./Ternorth, H.-E./Vollmer, H.J. (2003): *Zur Entwicklung nationaler Bildungsstandards*. Eine Expertise. Hrsg. v. Bundesministerium für Bildung und Forschung. Bonn: BMBF.
- Kounin (1976) *Techniken der Klassenführung*. Stuttgart.
- Leahy, S., Lyon, C. Thompson, M. & William, D. (2005): *Classroom Assessment: Minute by Minute, Day by Day*. *Educational Leadership*. Vol. 63/3, 19 – 24.
- Maag Merki, K. (2010): *Bildungsstandards, Lernstandserhebungen und Abschlussprüfungen*. In: Altrichter, H. & Maag Mercki, K. (2010): *Handbuch Neue Steuerung im Schulsystem*. VS Verlag für Sozialwissenschaften, Wiesbaden, S. 145 - 169.

- Nichols, S.L., Glass, G.V. & Berliner, D.C. (2005): High-Stakes Testing and Student Achievement: Problems for the No Child Left Behind Act. Forschungsbericht, über Internet abrufbar.
- Pashler, H., Bain, P.M., Bottge, B.A., Graesser, A., Koedinger, K., McDaniel, M. & Metcalfe, J. (2007): Organizing Instruction and Study to Improve Student Learning (NCER 2007-2004). Washington DC: National Center for Education Research, Institute of Educational Sciences, U.S. Department of Education. Retrieved from <http://ncer.ed.gov>.
- Raymond, M. E / Hanushek, E. A. (2003): High Stakes Research. *Education Next*, Summer, S. 49 – 55.
- Rosenshine, B. (2003, August 4). High-Stakes testing: Another analysis. *Education Policy Analysis Archives*, 11(24) p. 4. Retrieved January 7, 2004 from <http://epaa.asu.edu/epaa/v11n24/>
- Schipper, Wilhelm (August 2005): Modul G 4: Lernschwierigkeiten erkennen – verständnisvolles Lernen fördern. Sinus Grundschule, Kiel. (Datei kann im Internet heruntergeladen werden!)
- Stiggins, Richard J. (2004). New Assessment Beliefs for a New School Mission, *Phi Delta Kappan*. September, 22-27.
- Wellenreuther, M. (2005): *Lehren und Lernen – aber wie? Empirisch-experimentelle Forschungen zum Lehren und Lernen im Unterricht*. Grundlagen der Schulpädagogik, Band 50. Schneider Verlag Hohengehren.
- Wellenreuther, M. (2009 a): *Forschungsbasierte Schulpädagogik. Anleitungen zur Nutzung empirischer Forschung für die Schulpraxis*. Schneider: Hohengehren.
- Wellenreuther, Martin (2009 b): Individualisieren – aber wie? *Schulverwaltung Ni.*, 20. Jahrgang, 166 – 169.
- Wellenreuther, Martin (2009 c): Handwerkszeug für erfolgreichen Unterricht. Klassenmanagement ist mehr als Ermahnen und Strafen. *Friedrich Jahresheft XXVII 2009*, S. 45 – 47.
- Wellenreuther, M. (2010): Schulbücher– eine Lernhilfe für Schüler und Lehrer? *Schulverwaltung NRW*, 21. Jahrg., Heft 5, im Druck.
- William, D. (2006): Assessment: Learning communities can use it to engineer a bridge connecting teaching and learning. *Journal of Staff Development*. 27 (1), 16 – 20.
- William, D. (2007): Content, then Process: Teacher Learning Communities in the Service of Formative Assessment. In: D. Reeves (Ed.): *Ahead of the curve: The power of assessment to transform teaching and learning*. Bloomington, IN: Solution Tree, S. 182 – 204.
- Zhu, Xinming & Simon, Herbert A. (1987): Learning mathematics from examples and by doing. *Cognition and instruction*, Vol. 4, No. 3, 137 – 166.