

Lange, Hermann

Qualitätssicherung in Schulen

Die Deutsche Schule 91 (1999) 2, S. 144-159



Quellenangabe/ Reference:

Lange, Hermann: Qualitätssicherung in Schulen - In: Die Deutsche Schule 91 (1999) 2, S. 144-159 -

URN: urn:nbn:de:0111-pedocs-309201 - DOI: 10.25656/01:30920

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-309201>

<https://doi.org/10.25656/01:30920>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Digitalisiert

Mitglied der


Leibniz-Gemeinschaft

Die Deutsche Schule

Zeitschrift für Erziehungswissenschaft,
Bildungspolitik und pädagogische Praxis

91. Jahrgang 1999 / Heft 2

Offensive Pädagogik:

134

Jochen Schweitzer

Keine Angst vor PISA!

Perspektiven einer neuen Bildungsoffensive

Als Antwort auf die mäßigen Leistungen, die bei TIMSS für deutsche SchülerInnen ermittelt wurden, haben die Kultusminister beschlossen, sich an der internationalen Leistungsstudie PISA zu beteiligen. Was sich dahinter verbirgt und welche Chancen für die Entwicklung des Bildungswesens damit verbunden sind, wird von einem Insider erläutert.

Hermann Lange

144

Qualitätssicherung in Schulen

In Zeiten des knappen Geldes müssen die verfügbaren Mittel effektiv(er) verwendet werden. Die größere Autonomie der Schulen soll dazu beitragen. Erforderlich ist aber zugleich, daß die Kriterien der "guten Schule" geklärt werden und daß Verfahren zur Sicherung bzw. Prüfung dieser Maßstäbe entwickelt und angewendet werden. Dies ist nicht ohne Risiken. Die Projekte TIMSS, LAU und PISA werden in ihren Absätzen beschrieben.

Volker Hagemeyer

160

Was wurde bei TIMSS erhoben?

Über die empirische Basis einer aufregenden Studie

Wer sich die bei TIMSS verwendeten Items – soweit sie verfügbar sind – ansieht, wird etliche problematische Aspekte finden: Aufgaben sind mißverständlich, die 'richtigen' Antworten unklar, Übersetzungen sind nicht angemessen, nicht alle Aufgaben passen zu dem in Deutschland üblichen Unterricht. Es ist deshalb unzulässig, aus den Daten auf Lernfortschritte oder Unterschiede zwischen Schulformen zu schließen. Die erforderliche Diskussion wird erst möglich sein, wenn die Testaufgaben und die Messverfahren öffentlich zugänglich sind.

Zum Umgang mit den Ergebnissen von TIMSS

Selbst wenn TIMSS in Details zu kritisieren ist, sollte die produktive Herausforderung genutzt werden, die mit einem genaueren Blick auf Leistung und Lernzuwächse möglich wird.

Bildung und Leistung gehören zusammen!

Auf den durch TIMSS ausgelösten „Schock“ sollten die Reformer nicht mit Abwehrreaktionen antworten. In einer offenen Gesellschaft sind nun einmal emanzipatorische und sozialelektive Momente eng miteinander verbunden. Wichtig ist jetzt mehr denn je empirischer Forschung, die vor allem die Qualität des Unterrichts in den Blick nimmt.

TIMSS – kulturell interpretiert

Die öffentliche Diskussion über TIMSS basiert auf einer Auswertung, die auf einfache Ländervergleiche verengt ist: Es werden höchstens mikrodidaktische Analysen (Videostudien) vorgenommen. Vergleicht man die teilnehmenden Länder auch nach der Ausprägung kultureller Dimensionen, dann werden Unterschiede deutlich, die zu der Folgerung führen, daß jede Nation in der Bildungspolitik ihren eigenen Weg finden muß.

TIMSS-Differenzen

Die Leistungen in den Naturwissenschaften und der Mathematik in Deutschland und der Schweiz

Dass den SchülerInnen in Deutschland und der Schweiz bei TIMSS unterschiedliche Leistungen bescheinigt werden, kann verständlich werden, wenn man sich die Test-Aufgaben genauer ansieht. Die für Deutschland ermittelten Defizite liegen besonders im Bereich der schwierigeren Aufgaben des „Anwendens und Problemlösens“. Sie werden mit Leistungsdruck und -kontrolle kaum zu beheben sein.

Der Ruf nach Leistung

Der Ruf nach „mehr Leistung“ in der Schule ist unüberhörbar. Bedeutet für Schulen die Wiederentdeckung des Leistungsprinzips den Verzicht auf das Lernziel Bildung? Was heißt Leistung? Welche Ansprüche sind an einen verantwortbaren Begriff schulischer Leistung zu stellen?

Schulen im Vergleich

Probleme des Ranking und Chancen eines Monitoring

Wenn man die „Leistungen“ von Schulen „fair“ miteinander vergleichen will, müssen die jeweiligen Lernausgangslagen und die sozialen Lernbedingungen berücksichtigt werden. Wie kann man die Lernwirksamkeit von Schulen dokumentieren und vergleichend beurteilen? Und welche Ergebnisse sind dabei zu erwarten? – In Verfahren des „Schulmonitoring“ kommen die Lernentwicklungsverläufe einzelner Schulen in den Blick. So kann Evaluation eine pädagogische Funktion haben.

Lernzeit und Konzentration**Grundschulkinder in offenen Lernsituationen**

Die Diskussion über offene Lernsituationen kann sich bisher kaum auf empirische Untersuchungen beziehen, in denen die Auswirkungen erkennbar werden. Wie nutzen Kinder mit unterschiedlichen Konzentrationsleistungen die Lernzeit, die ihnen in offenen Lernsituationen zur Verfügung steht? Die Ergebnisse der Studie verdeutlichen, wie wichtig eine differenzierte Sicht auf das Lernverhalten von Kindern in offenen Lernsituationen ist.

Gerhard H. Duismann, Wolf Martin, Rolf Oberliesen und Joseph Pangalos: 246

Lernszenarien

Eine Leitidee für die Gestaltung von Lernprozessen am Beispiel der Technischen Bildung

Technische Bildung erhält in der technisierten Welt eine immer größere Bedeutung. In dieser Studie wird versucht, durch die Zentrierung auf Schlüsselprobleme unterschiedliche Handlungsniveaus zu identifizieren und dadurch allen Lernenden die drängenden Problemen der technisierten Welt zugänglich zu machen. Ein Beispiel aus der Lehrerbildung macht die beispielhaft deutlich.

Neuerscheinungen:

- Kurt Hahn: Reform mit Augenmaß. (HGH)
- Ulrike Mietzner: Enteignung der Subjekte – Lehrer und Schulen in der DDR. (HGH)
- Carl-Heinz Evers: Zwischen-Fälle. (D.W.)
- Helmut Fend: Qualität im Bildungswesen. (HGH)
- Andreas Hinz u.a.: Die Integrative Grundschule im sozialen Brennpunkt. Und: Die Entwicklung der Kinder in der Integrativen Grundschule. (Deppe-Wolfinger)
- Sigrid Blömeke (Hg.): Reform der Lehrerbildung? (Bracht)
- Anton Haas: Unterrichtsplanung im Alltag. (Lißmann)
- Annette Pöler-Klassen: Soziales Lernen. (Wrobel)
- Felix Mattmüller und Josef Fragner: Integration als Projekt der Gleichwertigkeit. (Schnurer)
- Dagmar Schulz: Zum Leistungsprinzip in der DDR. (JöS)
- Ursula Carle und Astrid Kaiser (Hg.): Rechte der Kinder. (JöS)
- Guido und Michael Grandt: Waldorf Connection. (JöS)

Hinweis: Inhaltsverzeichnisse früherer Hefte und eine Liste der rezensierten Publikationen sind im Internet über <http://www.gew.de> zu finden.

„Die Deutsche Schule“ (Schools in Germany)

Content of Volume 91, 1999, No. 2

Pedagogy on the offensive:

Jochen Schweitzer

Don't worry about PISA!

Perspectives on a new educational offensive

As an answer to the mediocre results established by TIMMS in regard to German students, the ministers of education have resolved to participate in the international study of educational achievement PISA. What does this imply for the further development of the educational system?

Hermann Lange

Quality assessment at schools

It is necessary to clarify the criteria for quality assessment, to develop and apply instruments and procedures of evaluation. That is not without risks. The projects TIMMS, LAU and PISA are described.

Volker Hagemester

What has been measured by TIMMS?

On the empirical basis of an entangling study

A closer look at the items used in TIMMS – as far as they are available – reveals several problematical aspects. It is therefore inadmissible to conclude from these data to the amount of learning progress or differences between school forms. The necessary debate will become possible only if the public has access to test assignments and the research procedures.

Thomas Bethge

On dealing with the results of TIMMS

Even if some of the details of TIMMS are to be criticized, there is a productive challenge consisting in taking a closer look at achievement and learning progress.

Hajo Sygusch

'Bildung' and Achievement belong together

The shock released by TIMMS should not lead reformers to answer with defensive reactions. It is a fact of open societies that emancipatory and social-selective moments are closely intertwined. Empirical research is now more important than ever, especially focussed on the quality of instruction.

Matthias v. Saldern

A Cultural-oriented Interpretation of TIMMS

The public debate on TIMMS is based on an evaluation narrowed down to simple comparisons of countries: At the most microdidactical analyses (videostudies) are carried out. However, if the participating countries are also compared in regard to cultural dimensions, differences will become visible that lead to the conclusion that every nation has to find its own way in educational politics.

Erich Ramseier

TIMMS Differences

Educational Achievement in the Natural Sciences and Mathematics in Switzerland and Germany

The differences in educational achievement of students in Switzerland and in Germany become understandable if a closer look is taken at the tests. The deficits ascertained for Germany are located in the area of difficult assignments of "application and problem-solving". These deficits will not be rectified by exerting pressure and control.

Jürgen Linke

The call for achievement

The call for "better performance" in schools can't be overheard. Does the rediscovery of the achievement-principle in schools imply the renunciation of Bildung as a learning goal? What does achievement (attainment, performance) mean? What kind of claims are laid to a responsible concept of educational achievement?

Karl-Heinz Arnold

Comparison of Schools

Problems of Ranking and Chances of Monitoring

How can the effectivity of schools be documented and comparatively evaluated? What kind of results are to be expected? – In Procedures of school monitoring the courses of learning development of individual schools become evident. In this way, evaluation can have a pedagogical function.

Frank Lipowsky

Learning time and capabilities to concentrate

Primary school pupils in open education

There is a lack of relevant empirical studies dealing with pupils' behaviour in open education. How do pupils with different capabilities to concentrate use the learning time in open learning arrangements? The results show the importance of differentiation when it comes to analysing and understanding pupils' behaviour..

Gerhard H. Duismann, Wolf Martin, Rolf Oberliesen and Joseph Pangalos

Learning-sceneries

A leading idea for the shaping of learning processes

Technical education gains in importance in a technical world. This study attempts to identify different levels of action by focusing on key problems. This focus, illustrated by an example of teacher education, is meant to provide access to urgent problems of the technical world for all learners.

Qualitätssicherung in Schulen

Es ist immer gut, die Realitäten so in den Blick zu nehmen, wie sie wirklich sind. In der Bildungspolitik besteht dazu heute mehr denn je Anlaß. Vor ungefähr 30 Jahren hatte *Heinrich Roth* die *empirische Wendung* der pädagogischen Wissenschaften gefordert. Sie gilt es, so könnte man auch sagen, endlich zu vollziehen, und zwar auch in der Bildungspolitik und im Schulalltag.

Das im Augenblick im Vordergrund der Wahrnehmung vieler Beobachter stehende Phänomen ist das der Mittelknappheit im Bildungswesen. Verdrängt wird dabei, daß die gegenwärtige Situation jedenfalls quantitativ durch ein hohes Niveau der Mittelausstattung deutscher Schulen gekennzeichnet ist. Die *Phase des großen Wachstums* der Mittel für das Bildungswesen fiel in die zwanzig Jahre zwischen 1965 und 1985, als demographisch bedingte Zuwächse der Schülerzahlen mit nachhaltigen strukturellen Veränderungen der Bildungsbeteiligung und einer beachtlichen Verbesserung der Ausstattungsparameter (Klassengröße, Studentafeln, vielfältige Förderprogramme etc.) zusammentrafen.

Sie hatten für die Schulen eine *enorme Ausweitung der Zahl der Lehrerstellen* zur Folge. Entsprechend hoch war die Zahl von Neueinstellungen junger Lehrkräfte. Die äußere Verbesserung der Situation ist durch die Konsolidierungspolitik der letzten Jahre nicht entscheidend zurückgeführt worden. Zwar sind z.B. die Schüler-Lehrer-Relationen heute durchweg ungünstiger als Anfang der 90er Jahre. Verglichen mit den Werten 1970 oder 1975 bewegen sie sich aber immer noch auf einem sehr hohen Niveau (vgl. BMW 1988/89 und 1997/98). Auch im internationalen Vergleich ist das Niveau der Bildungsausgaben in Deutschland, bei allen Schwierigkeiten solcher Vergleiche, so gering nicht. Das gilt nicht zuletzt auch für die Lehrergehälter (vgl. Sachverständigenrat 1998, S. 23 ff).

Die zahlreichen Neueinstellungen in den 70er Jahren haben die *gegenwärtige strukturelle Schieflage in der altersmäßigen Zusammensetzung der Lehrerschaft* bewirkt. Sie wird gemeinhin als „Überalterung“ diskutiert. Dieser Begriff trifft nicht, was gemeint ist. Mit 50 Jahren – dies ist etwa ist das gegenwärtige Durchschnittsalter der Lehrerinnen und Lehrer in Hamburg – sollte man eigentlich auf der Höhe seiner beruflichen Leistungsfähigkeit sein. Problematisch aber ist, daß ein großer Teil der Lehrerinnen und Lehrer sich in diesem Alter befindet. Man durchlebt sozusagen gemeinsam seine Midlife-Crisis und muß auf das Anregungs- und Lastverteilungspotential verzichten, das aus einer größeren altersmäßigen Durchmischung eines Personalkörpers erwächst. Die strukturelle Schieflage hat sich durch die Stellenstreichungen im Rahmen der Konsolidierungspolitik bei wachsenden

Schülerzahlen tendenziell weiter verstärkt. Dies ist vielleicht das gravierendste Erschwernis der Qualitätssicherung, wo diese tiefgreifende Veränderungen gewohnter Verhaltensweisen voraussetzt.

Es ist unter Fachleuten unstrittig, *daß das gegenwärtige Niveau der Mittelausstattung auf lange Zeit nicht wesentlich zu verändern sein wird*. Eine Situation, in der man auf alle tatsächlichen oder vermeintlichen Probleme im Schulwesen mit einem Mittelzuwachs reagieren konnte, wird es nicht mehr geben, und zwar aus einem doppelten Grunde:

(1) *Wir können es uns nicht leisten*: Das bereits jetzt erreichte Niveau der öffentlichen Verschuldung, die anhaltende Notwendigkeit hoher Transferleistungen für die neuen Länder, die enormen Kosten der sozialen Absicherung der Arbeitslosen und die drohende Versorgungslawine machen eine Steigerung des Bildungsbudgets – jenseits aller Wünschbarkeit und inhaltlichen Begründbarkeit – unmöglich (vgl. Sachverständigenrat 1998).

(2) *Die Bereitschaft der Politik*, erneut in eine nachhaltige Verbesserung der Ressourcenausstattung der Schulen zu investieren, z.B. Klassenstärken weiter zu verringern, Stundentafeln weiter zu verbessern, neue Förderprogramme aufzulegen oder die Arbeitszeit der Lehrer zu reduzieren, *ist nicht vorhanden*, weil und solange die Frage nach den Erträgen einer derartigen Verbesserung, wenn überhaupt, bisher eher diffus beantwortet ist.

Die Gleichung „Mehr Geld = besserer Ertrag“ gilt nicht mehr (Weiß 1997). Wie viele andere öffentlichen Bereiche auch, stehen die Schulen und die Bildungsressorts unter hohem Kosten- und Legitimationsdruck. Dabei stellen sich die aufgeworfenen Fragen gewissermaßen von beiden Enden her. Einerseits geht es darum, wie trotz finanzieller Restriktionen ein Optimum an Leistungen der Schulen im Interesse von Schülerinnen und Schülern gesichert werden kann. Andererseits wird aber auch gefragt, was die Ergebnisse der Arbeit von Schulen angesichts des hohen Mittelaufwandes für das Bildungswesen in den öffentlichen Haushalten sind. Damit tritt die Frage nach den „Erträgen“ der Arbeit von Schulen zunehmend in den Vordergrund. Sie ist nicht mehr ohne weiteres als Ausdruck von Unverstand und banausischem betriebswirtschaftlichen Denken zu denunzieren. Es werden Nachweise dafür verlangt, daß in der Sache auch wirklich erreicht wird, wofür Mittel gefordert und bewilligt wurden. Pädagogik und Betriebswirtschaft schließen einander prinzipiell ebensowenig mehr aus wie Ökonomie und Ökologie.

Dabei deckt sich die *Wahrnehmung der Situation seitens der Lehrerinnen und Lehrer* keineswegs mit der *Wahrnehmung ihrer Umwelt* (Eltern, Arbeitgeber, „Politik“). Erstere sehen die Lösung vieler Probleme noch immer in einem Ressourcenzuwachs, den niemand zu bewilligen bereit und in der Lage ist. Über diesen Befund sollte auch die verhältnismäßig leicht zu erzeugende Bereitschaft zu öffentlichen Demonstrationen für mehr Mittel zugunsten der Bildung nicht hinwegtäuschen. Als Steuerzahler angesprochen, sind die Demonstranten keineswegs bereit, mehr Geld für die Bildung aus ihrem Portemonnaie abzugeben. Sie setzen auf Umverteilungen innerhalb des öffentlichen Sektors, für die es aber keine Spielräume gibt, und am Ende letztlich ebenfalls auf die Aktivierung von Leistungsreserven innerhalb des Bildungswesens selbst.

Die Diskrepanz der Situationswahrnehmung ist keineswegs ein isoliertes deutsches Phänomen: Die jüngste Ausgabe der amerikanischen Zeitschrift „Education Week“ berichtet z.B. über Untersuchungen in Amerika (Olson 1999), wonach z.B.

zwischen 70 und 80 % der Arbeitgeber und der Eltern die Möglichkeit einer Beendigung des Kontraktes des Schulleiters für den Fall, daß eine Schule keine gute Arbeit macht, für eine richtige Idee halten, während zwei Drittel der Lehrerinnen und Lehrer dies für eine schlechte Idee halten.

zwischen 60 und 70 % der Arbeitgeber und der Eltern die externe Überprüfung solcher Schulen befürworten, während sich 68 % der Lehrerinnen und Lehrer gegen diese Möglichkeit aussprechen.

zwischen 50 und 60 % der Arbeitgeber und der Eltern für eine leistungsmäßige Differenzierung der Lehrerbesehung plädieren, während 76 % der Lehrerinnen und Lehrer dies ablehnen.

Diese Wahrnehmungsdiskrepanz ist beunruhigend. Veränderungen lassen sich nur mit, nicht gegen wesentlich Beteiligte erreichen. Man muß also etwas tun, um die Lehrerschaft „mitzunehmen“ (die freilich nicht die einzigen Betroffenen sind), und zwar unabhängig davon, ob man ihre Position in der Sache teilt oder nicht.

1. Der Paradigmenwechsel der Organisations- und Steuerungsphilosophie: das „Neue Steuerungsmodell“

Die Begriffe „Evaluation“, „Qualitätssicherung“, „Erfolgskontrolle“ sind zur Zeit in aller Munde. Dies gilt für Schulen und Hochschulen wie für andere Bereiche öffentlicher Verwaltung gleichermaßen. Hierin zeigt sich ein *genereller Paradigmenwechsel der Organisations- und Steuerungsphilosophie des öffentlichen Bereichs*. Dieser wird allenthalben unter dem Begriff „Neues Steuerungsmodell“ diskutiert. Dieses Modell fand, am niederländischen Vorbild orientiert, zunächst Eingang in die deutsche Kommunalverwaltung¹, gilt inzwischen als Orientierung aber weit über den Bereich der Kommunalverwaltung hinaus. Im Kern geht es um den Übergang von input-orientierten zu output-orientierten Steuerungsverfahren: Wurden bisher den Organisationseinheiten Ressourcen mit detaillierten haushaltsrechtlichen Vorgaben für die Zwecke ihrer Verwendung zur Verfügung gestellt, wirksame Verfahren einer systematischen Ergebnisfeststellung und -bewertung jedoch kaum entwickelt, so sollen die Dinge künftig gewissermaßen vom Kopf auf die Füße gestellt werden. Betont wird die Eigenverantwortung der Organisationseinheiten für den Einsatz ihrer Mittel unter zunehmender Lockerung formeller Bindungen. Kehrseite dieser Entwicklung ist die Verpflichtung, systematischer und genauer als bisher Rechenschaft über Arbeitsergebnisse abzulegen.

Diese Entwicklungen beschränken sich nicht auf Deutschland. Ihre *Ursachen* sind vielfältig. Sie liegen zum einen auf der Ebene genereller organisationstheoretischer Überlegungen, die die Ablösung traditioneller physikalisch-mechanistischer Steuerungsmodelle für Großorganisationen durch

1 Vgl. Kommunale Gemeinschaftsstelle 1993 und 1995

biologisch-adaptive Steuerungsverfahren nahelegen (vgl. Lange 1995). Sie haben ihre Ursache zweitens – wie bereits erwähnt – in der Haushaltssituation, die die traditionelle Lösung von Problemen durch einen Zuwachs an Haushaltsmitteln verbietet. Was immer erreicht werden soll, muß zunächst durch die *Steigerung des Wirkungsgrades der vorhandenen Mittel* erreicht werden. Für den Bildungsbereich schließlich kommt drittens hinzu, daß seine Ergebnisse vor dem Hintergrund der Bedeutung von Ausbildung und Qualifizierung im weltweiten Wettbewerb gesellschaftlich zunehmend einer kritischen Revision unterzogen werden. Letzteres hat – wie z.B. die TIMS-Studien zeigen – durchaus Gründe. Ihre Ergebnisse können nicht einfach als Ausdruck altväterlicher und überholter Bildungsziele weg-gewischt werden.

Der Weg, der zu gehen ist, wird dornig sein. Zu bewältigen ist ein Wandel, der in unserer politischen Kultur bisher kaum vorbereitet ist. Qualitätssicherung ist nicht das Bemühen um plakative Selbstbestätigung auf Feldern, auf denen dies gefahrlos möglich ist. Sie ist vor allem die Suche nach Mißerfolgen und nach Schwächen, die es zu beseitigen gilt. Wenn es einen Effizienzunterschied zwischen öffentlichen und privaten Organisationen gibt, so besteht dieser nicht darin, daß die einen permanent Fehler machen und die anderen nicht. Man tut sich im öffentlichen Bereich nur ungleich schwerer, Fehler schnell genug zu erkennen und zu beseitigen, fehlgeschlagene Experimente konsequent genug zu beenden und den damit verbundenen Aufwand abzuschreiben. Die Gründe hierfür liegen vor allem in der Struktur des öffentlichen Diskurses über Politik. Wer Fehler und Mißerfolge offenbart, wird dafür selten belohnt. Er bekommt vielmehr zu hören, was er erwartet: Man habe es immer schon gewußt, er (oder sie) sei seiner (ihrer) Aufgabe nicht gewachsen.... Der Versagensvorwurf an die Gegenseite als Stereotyp der politischen Argumentation, was immer passiert und was immer auch die Gründe dafür sein mögen, ist so etwas wie ein Pawlow'scher Reflex von Politikern. Er wird gefördert von einer Publizistik, die sich auch in ihren seriösen Sektoren den Zwängen des allgemeinen Amüsierbetriebes – Neil Postman (1985) läßt grüßen – oft genug nicht entziehen kann und daraus die Dramaturgie ihrer Berichterstattung ableitet. Dies führt auf Seiten der Verantwortlichen leicht zu einer Rechtfertigungspolitik, die sich schwer tut, Fehlschläge oder Fehler einzugestehen und notwendige Veränderungen entschlossen einzuleiten.

2. Qualitätssicherung und Evaluation im Kontext von „Schulautonomie“

Die in den Entwicklungen der allgemeinen Organisationsphilosophie angelegte *Tendenz einer größeren Autonomie der Subsysteme gilt auch für die Schulen*. Auch für sie wird bekanntlich die Forderung nach einer größeren Autonomie („Schulautonomie“) erhoben. Diese Autonomie ist notwendig. Freilich haben sowohl ihre Befürworter unter den wackeren Pädagogen als auch ihre Gegner unter den ebenso wackeren Politikern und Schulverwaltern häufig eigenartige Vorstellungen über ihre Ziele und Konsequenzen. Die einen sehen in der Schulautonomie endlich die Möglichkeit, sich in eine selbstbestimmte Welt des Traums von einer unbegrenzten pädagogischen Freiheit zu flüchten, die anderen fürchten gerade dieses und sehen das Chaos ausbrechen. Beides ist unzutreffend. Es geht in Wahrheit um die Rückge-

winnung von Steuerung in einem Bereich, der sich inzwischen der Steuerungsmöglichkeit mit den traditionellen Mitteln eines engmaschigen Netzes von Vorschriften weitgehend entzogen hat. Dies gilt es näher zu erläutern.

(1) Autonomie heißt nicht, daß sich jeder seine eigenen Gesetze geben kann. Sie ist *kein Freibrief* für Beliebigkeit und keine Rechtfertigung für Weggucken. Freiheit und Verantwortung gehören zusammen (vgl. Liket 1993). Die Ziele des Schulwesens entspringen nicht dem subjektiven Empfinden des einzelnen Pädagogen oder der einzelnen Schule. Sie sind mit ihrem Auftrag vermacht, über den letztlich politisch entschieden wird und der sich in vielfältigen Formen schulgesetzlicher Normierung und curricularer Konkretisierung niederschlägt. „Politisch entschieden“ heißt nicht, daß die verantwortlichen Entscheidungsträger dezisionistisch im luftleeren Raum agieren. Sie bringen vielmehr auf den Begriff, was sich als Ergebnis eines komplizierten Verständigungsprozesses in Wissenschaft und Gesellschaft herauskristallisiert und im Zuge der damit bezeichneten Entwicklungen – z.B. als Reaktion auch auf Entwicklungen jenseits des eigenen nationalen Horizonts – jeweils wieder verändert.

(2) *Qualität läßt sich nicht befehlen*. Dies gilt allemal dann, wenn es nicht um oberflächliche Veränderungen, sondern um einen tiefgreifenden Wandel der Organisationskultur geht. Er kann ohne Akzeptanz seitens der wesentlichen Akteure nicht gelingen. Mögen die Ziele, die es zu erreichen gilt, auch vorgegeben sein. Wie sie unter den je spezifischen Bedingungen einer Schule am besten erreicht werden können, kann sie selbst am ehesten beurteilen. Was sich an der einen Stelle nach Schulumwelt und Kontextbedingungen als sinnvoll erweist, muß nicht an anderer Stelle gleichermaßen wirksam sein. Deshalb ist es erforderlich, daß die jeweilige Schule ein eigenes Interesse entwickelt, ihren Auftrag möglichst gut zu erfüllen. Sie muß hierfür auch neue Wege gehen können, ohne immer andere fragen zu müssen. Deshalb benötigt sie Gestaltungsfreiheit, aus der sie Identität und Selbstvertrauen gewinnt. Dies macht den Kern ihrer Autonomie aus. Ein Instrument, mit dem die einzelne Schule auf den Begriff bringen kann, was sie will, ist das Schulprogramm. Es erfüllt seinen Zweck allerdings nur, wenn es sich in umfassender Weise den Aufgaben der Schule, das heißt dem Lernen von Schülerinnen und Schülern stellt, die darauf bezogenen Aktivitäten präzise beschreibt und geeignete Verfahren einer Kontrolle der Zielerreichung angibt. Ein Schulprogramm, das diese Mitte ausspart und sich unter der Überschrift „Profil“ lediglich den Rändern zuwendet, verfehlt seinen Zweck.

(3) *Ob man erreicht, was man erreichen will*, muß systematisch mit dafür geeigneten und belastbaren Methoden geklärt werden. Man muß sich den Ergebnissen dieses Klärungsprozesses stellen und Folgerungen daraus ziehen. Daraus folgt die Notwendigkeit von Evaluation. Diese hat eine doppelte Funktion: Bemühen um Verbesserung („improvement“) und Rechenschaftslegung („accountability“). Es gibt mithin interne und externe Funktionen, die auch Konsequenzen für die Verfahrensgestaltung haben. Freilich garantiert die Evaluation als solche noch nicht die Qualität des angestrebten Ergebnisses. Sie ist nur ein Mittel im Prozeß der Qualitätssicherung. Dabei kann schon die Tatsache, daß die Beteiligten systematisch

nach den Ergebnissen ihrer Arbeit fragen, den Prozeß der Qualitätssteigerung befördern.

- Das Bemühen um *Evaluation ist zunächst Sache der betroffenen Schule selbst*. „Interne Evaluation“ ist deshalb ein zentraler Bestandteil der Schulautonomie. Dabei wird aus der Sicht der betroffenen Schule der Aspekt der Verbesserung im Vordergrund stehen. Die Evaluation muß sich mit dem Kernauftrag der Schule, dem Lernen von Schülerinnen und Schülern in allen relevanten Feldern auseinandersetzen und sich damit auch auf die extern gesetzten Leistungsziele beziehen. Sie stellt hohe Anforderungen an die Bereitschaft zu kritischer Selbstreflexion. Dabei ist es verständlich, daß man bestrebt ist, Stärken und Schwächen nicht auf dem offenen Markt zu klären. Man muß auch einmal bei sich bleiben, um sich gegenseitig in die Karten sehen und auch unangenehme Wahrheiten sagen zu können. Interne Evaluation ist keine bequeme Sache. Deshalb fragen auch die Skeptiker: Wie ernst wird man sie nehmen? Die Informationsquellen für den Evaluationsprozeß können vielfältiger Art sein. Neben gezielten Untersuchungen unter Einsatz von Testinstrumenten spielt z.B. das Urteil von Eltern und Schülern eine große Rolle.

- Die *Evaluation hat aber auch externe Funktionen*; sie bedarf zudem zur Klärung ihrer Maßstäbe auch der Außensicht. Als Rechenschaftslegung verstanden, gehen die Ergebnisse nicht nur die jeweilige Schule an. Dies ist kein Plädoyer dafür, daß sich eine „Schulaufsicht“ als Kontrollinstanz weiterhin in alles einmischt. Sie muß sich aber sehr wohl um den Prozeß der Evaluation und seine Ergebnisse kümmern. Dem steht nicht entgegen, daß die Schulaufsicht, um die Funktion eines systembezogenen „Controlling“ sachgerecht erfüllen zu können, ihrerseits qualifiziert werden muß. Zudem ist im Rahmen eines Evaluationsprozesses die Bereitschaft, sich auch dem Blick von außen zu stellen, unverzichtbar. Man bedarf eines Referenzrahmens, für den auch der Vergleich mit anderen notwendig ist, und man muß versuchen, kritische Distanz zu sich selbst zu gewinnen, die es ermöglicht, die eigenen Stärken und Schwächen in vorurteilsloser Weise zu erkennen. Deshalb haben die Vergleichsuntersuchungen, von denen im folgenden die Rede sein wird, im Rahmen von Evaluationsprozessen eine eigene Bedeutung. Sie zu organisieren, übersteigt die Möglichkeit der einzelnen Schule und bringt damit notwendigerweise externe Akteure ins Spiel². Im übrigen verfolgen breiter angelegte Vergleichsuntersuchungen eigene legitime Zwecke und Erkenntnisinteressen der staatlichen Instanzen, nämlich z.B. einen Überblick über die Situation der Schulen im eigenen Verantwort-

2 Allerdings wäre es durchaus einen Gedanken wert, die Schulen selbst an der Steuerung übergreifender Vergleichsuntersuchungen und Evaluationsprozesse stärker zu beteiligen, als dies bisher der Fall ist. Hier könnte das Vorbild der Hochschulevaluation in den Niederlanden oder durch die niedersächsische zentrale Evaluationsagentur in Hannover durchaus Anregungen geben: Träger der Evaluation sind Agenturen, die von den Hochschulen getragen werden und von ihnen als Teil des Hochschulsystems begriffen werden, auch wenn der konkrete Evaluationsprozeß sich für die jeweilige Hochschule als ein externer darstellt. Zugleich gibt es Rückkopplungen zu den jeweiligen Hochschulministerien, die die Funktion der Rechenschaftslegung gegenüber den staatlichen Instanzen erfüllen.

tungsbereich zu gewinnen und sie vor dem Hintergrund der Ergebnisse anderer (z.B. im Vergleich zu anderen Bundesländern oder im internationalen Vergleich) bewerten zu können. Dies kann für die Entwicklung neuer curricularer Konzeptionen oder die Lehreraus- und -fortbildung von Bedeutung sein.

Interne und externe Evaluation gehören mithin zusammen. Genau an dieser Stelle gibt es derzeit freilich erbitterte Diskussionen. Viele Anhänger des Konzepts der Schulautonomie halten es für eine Sünde wider den heiligen Geist pädagogischer Unschuld, neben interner gleichzeitig auch externe Evaluation zu betreiben. Ihnen sei gesagt: Schulautonomie ist kein Unternehmen zur Steigerung der eigenen Selbstzufriedenheit. Die Verantwortung für die Ergebnisse der Arbeit besteht nicht nur gegenüber sich selbst. Auch eine „autonome Schule“ bleibt ein Unternehmen in öffentlicher Verantwortung.

(4) Ob man bei einer Evaluation die richtigen Antworten bekommt, hängt davon ab, *ob man die richtigen Fragen stellt*. Eine wesentliche Wirkung der gegenwärtigen Bemühungen um Schulevaluation ist deshalb, daß sie dazu zwingt, sich genau darauf zu besinnen, welche Fragen man stellen muß: Was sind die maßgebenden Ziele des jeweiligen Handelns, welche Maßnahmen ergreift man, um die Ziele zu erreichen und woran kann man erkennen, ob man sie auch erreicht? Die Notwendigkeit zu klären, ob man erreicht, was man sich vorgenommen hat, stellt sich unabhängig davon, welches die spezifischen Ziele jeweils sind. Sie gilt sowohl für den, der geneigt ist, eher traditionelle fachliche Lernziele in den Mittelpunkt der Aufgaben von Schulen zu stellen, als auch für den, der Schlüsselqualifikationen, Kritik- und Urteilsfähigkeit oder soziale Kompetenz stärker akzentuieren möchte. Diese Hinweise bringen zugleich die Überzeugung zum Ausdruck, daß – welche Ziele auch immer man verfolgt – diese sich in feststellbaren Effekten manifestieren müssen, welche in der Regel wiederum überprüfbar sind. Die gegenteilige Behauptung mangelnder Messbarkeit pädagogischer Erfolge ist oft nichts anderes als ein Beleg für unklare Zielsetzungen. Die Erreichung diffuser Ziele läßt sich in der Tat nur schwer messen. Diese Aussage will nicht leugnen, daß das Entwickeln geeigneter, intelligenter, aussagefähiger und wirklich anwendbarer Indikatoren und Meßinstrumente eine schwierige Sache und noch keineswegs überall gelungen ist. Zu akzeptieren sind nur zieladäquate Meßinstrumente und Indikatoren, die auch besagen, was sie besagen sollen. Die Forderung lautet: „Zu messen ist, was wichtig ist“. Ihre Umkehrung in: „Wichtig ist, was (möglichst einfach) zu messen ist“, wäre fatal.

(5) Damit ist die Frage aufgerufen, die in der Diskussion um „Qualitätssicherung“ in Schulen zwangsläufig auftauchen muß, *was nämlich eine „gute Schule“ ausmacht* und worin sich die „Qualität von Schule“ ausdrückt. Dies ist – wie jeder weiß – ein abendfüllendes Thema, das schon Generationen von Pädagogen beschäftigt hat. Sie ist in Wahrheit in pauschaler Form nicht zu beantworten. Deshalb ist in den vorausgegangenen Überlegungen mit Bedacht von „Zielen“ und „Klärung der Zielerreichung“ und nicht in allgemeiner Form von „Qualität“ gesprochen worden. Schule verfolgt ein Bündel unterschiedlicher Lernziele: fachliche und überfachliche, kognitive und metakognitive, soziale, emotionale, persönlichkeitsbezogene usw. Zwi-

schen ihnen bestehen positive wie negative Rückkopplungen im Sinne eines – wie zu vermuten ist – nichtlinearen Systemzusammenhangs. Angesichts der Mehrdimensionalität des Zielsystems, der Vielzahl von Interaktionen zwischen den Beteiligten und der Unmöglichkeit, die Anfangszustände des Systems genau zu beschreiben, ist auch – durchaus chaostheoretisch gedacht – eine Voraussage darüber, wie sich das Gesamtsystem bei dem Versuch der Optimierung zugunsten einzelner Ziele bezüglich der anderen entwickeln wird, nur in Grenzen möglich. Allerdings wäre es auch unsinnig, ausschließlich in Entgegensetzungen zu denken und anzunehmen, daß die bessere Erreichung eines der Ziele nur negative Rückwirkungen auf das Erreichen anderer Ziele hat. Es bestehen auch positive Zusammenhänge: Überfachliche Inhalte werden an fachlichen Inhalten gelernt. Auch soziales Lernen ist in bestimmtem Umfang wissensbasiert. Lernen gelingt in der Schule in der Regel nur, wenn Schülerinnen und Schüler durch wechselseitige Rücksichtnahme wenn nicht gar Unterstützung eine geeignete Lernatmosphäre entstehen lassen und so zugleich eine soziale Leistung erbringen. Die Attitüde des Raisonnements über alles und jedes ohne vorherige sorgfältige Klärung von Fakten und ohne die Formulierung geeigneter Urteilsmaßstäbe ist noch keine Kritik- und Urteilsfähigkeit. Selbstorganisiertes Lernen setzt ebenso wie Team- und Kommunikationsfähigkeit die Beherrschung von Sprache (eigener und fremder) wie auch die Fähigkeit voraus, in mathematischen Modellen denken zu können. Deshalb wäre es unsinnig, z.B. fachliche gegen soziale Ziele aufzurechnen und auf Kompensation des nichtgelingenden einen durch das gelingende andere zu hoffen. Man könnte die böse Überraschung erleben, daß beides nicht gelingt.

(6) Auch wenn die Zusammenhänge schwierig sind: Man muß auch im Bildungswesen *die Ziele, die man verfolgen will, explizieren*, die der Zielerreichung dienenden Maßnahmen beschreiben und in geeigneter Form überprüfen, ob man seine Ziele erreicht. Da man indessen, ganz praktisch, nicht stets alles messen kann, was man messen könnte, muß man also auswählen und sich entscheiden, welchem Ausschnitt des Zielbündels man im gegebenen Zeitpunkt seine besondere Aufmerksamkeit zuwenden will. Dazu gehört natürlich auch, daß man auf der Metaebene – in der Evaluation der Evaluation – die Nebenwirkungen des Evaluationsprozesses zu klären versucht. Dies schließt ein, daß man zu unterschiedlichen Zeitpunkten unterschiedliche Gegenstände in den Blick nimmt. Bei den derzeitigen Ansätzen zur empirischen Klärung der Leistung von Schulen stehen insbesondere die sprachliche Kompetenz der Schülerinnen und Schüler (Muttersprache, Fremdsprache), ihre mathematische Kompetenz und ein Grundverständnis naturwissenschaftlicher Prozesse im Vordergrund³. Die Untersuchungen zielen mithin auf zentrale Kompetenzbereiche, die in vielfältigen Zu-

3 In Vorbereitung ist ferner eine internationale Untersuchung zur politischen und historischen Bildung von Schülerinnen und Schülern (Civic Education Study) durch die IEA, die als Forschungsvorhaben von der Deutschen Forschungsgemeinschaft gefördert wird. Zur Frage der historischen und politischen Bildung auch Angvik, M., von Borries, B. (Hrsg.), *Youth and History: A Competitive European Survey on Historical Consciousness and Political Attitudes among Adolescents*, Edition Körber-Stiftung, 1997.

sammenhängen außer- und überfachlicher Art ihre Bedeutung haben und Voraussetzungen selbstorganisierten Lernens sind. Ihnen läßt sich das Standardargument des Abfragens rasch veraltender Wissensbestände nicht entgegenhalten. Nicht umsonst hat auch die Kultusministerkonferenz in ihren Vereinbarungen zur Sekundarstufe II gerade diese Kompetenzbereiche besonders betont (vgl. Abschlußbericht 1995). Das heißt – um es zu wiederholen – nicht, daß sich die Aufgabe der Schule in der Vermittlung dieser Kompetenzen erschöpft und daß die empirischen Untersuchungen sich hierauf beschränken dürften.

3. Was ist neu an der „empirischen Wende“?

Methoden und Instrumente, die zur Klärung des Erreichens der jeweils verfolgten Ziele eingesetzt werden, sind nicht neu. Neu ist indessen, *daß diese Instrumente und Methoden in breiter Form genutzt werden sollen*, um die Ergebnisse des konkreten Handelns systematisch zu überprüfen. Dies wird große Anstrengungen erfordern, und zwar nicht nur weil es immer schwer ist, einen vielfältigen Einflüssen und Interessen unterliegenden Großbereich überhaupt in Bewegung zu bringen. Vielmehr sprechen auch verbreitete Gewohnheiten des gegenwärtigen Wissenschaftsbetriebes dagegen. Zwar ist die Wendung der Erziehungswissenschaften zu einer (auch) empirisch-sozialwissenschaftlichen Orientierung bereits seit langem verkündet, insgesamt aber noch kaum vollzogen. Die Frontlinie der Auseinandersetzung um Eignung und Nebenwirkungen der Instrumente verläuft deshalb heute bemerkenswerterweise nicht zwischen „A-Ländern“ (SPD-regiert) und „B-Ländern“ (CDU-regiert), sondern quer durch die Profession selbst. Insofern geht es nicht nur um die empirische Wende in der Bildungspolitik, sondern noch immer auch um die empirische Wende der pädagogischen Wissenschaften. Das Bemühen, Folgerungen hieraus zu ziehen, stößt freilich auch auf die Schwierigkeit, daß die erziehungswissenschaftlichen Fakultäten vielfach nicht über die Stellen (Lehrstühle, Institute) verfügen, die sich kompetent mit den quantitativen Methoden empirischer Forschung befassen könnten. Daraus resultieren Defizite in der Lehre, die die Rezeption der Ergebnisse empirischer Untersuchungen durch die schulische Praxis erschweren.

Neu ist auch eine *Haltung, die das Dogma von der Unfehlbarkeit des Pädagogen aufhebt*, solange er nur im rechten Geist handelt. Man muß zugeben lernen und zugeben können, daß auch gute Absichten fehlschlagen können, daß nicht jedes gute Ziel auch erreicht wird und daß die Gründe hierfür nicht bereits mit der heute vielfach üblichen pädagogischen Verelendungstheorie (gesellschaftliche Entwicklungen, fehlende Mittel etc.) abschließend beschrieben sind. Das Nichterreichen von Zielen, das bisher als Folge nicht veränderbarer Umstände allzu leicht aus dem Verantwortungsbereich der handelnden Personen herausgeschoben wird, als individuelles Scheitern hinter verschlossenen Türen verborgen bleibt, von der einzelnen Lehrkraft als persönlich belastendes Schicksal verarbeitet werden muß oder durch Gelingensgemeinschaften und Verbrüderungen (der Lehrer mit den Schülern, der Kollegen untereinander und mit den Schulleitungen) verschleiert wird, kann offen kommuniziert werden. Ehrlichkeit gegen sich selbst kann freilich schmerzlich sein. Sie ist ohne Frage risikoreich.

Die Veränderungen werden Ängste und Widerstand auslösen. Sie können aber auch als Chance der (individuellen wie gemeinsamen) Befreiung und Erleichterung begriffen und als Möglichkeit genutzt werden, das Bildungswesen wieder in das Zentrum öffentlicher Aufmerksamkeit und öffentlichen Interesses zu rücken.

4. Leistungsmessung und Leistungsvergleiche, national und international: TIMSS, LAU und PISA

Gegenwärtig haben auf nationaler wie auf internationaler Ebene insbesondere vergleichende Untersuchungen zum Leistungsstand von Schülerinnen und Schülern Konjunktur. Das bekannteste und derzeit viel diskutierte Beispiel einer internationalen Vergleichsuntersuchung ist die „*Dritte Internationale Mathematik- und Naturwissenschaftsstudie*“ (The Third International Mathematics and Science Study – TIMSS), die unter dem Dach der International Association for the Evaluation of Educational Achievement (IEA) durchgeführt wurde. Die Ergebnisse dieser Untersuchung dürfen als bekannt vorausgesetzt werden, so daß sie hier nicht im einzelnen referiert werden müssen (vgl. Baumert/Lehmann 1997, Baumert/Bos/Watermann 1998).

Als ein wichtiges Beispiel für derartige Untersuchungen in Deutschland ist ferner die „*Lernausgangslagenuntersuchung*“ (LAU) zu nennen, mit deren Hilfe in Hamburg flächendeckend der Leistungsstand der Schülerinnen und Schüler am Beginn der Sekundarstufe I, d.h. letztlich am Ende der Grundschule, erfaßt wurde. Untersucht wurden dabei Kenntnisse, Fertigkeiten und Fähigkeiten in den Bereichen Sprachverständnis, Leseverständnis, Rechtschreibung, Mathematik und Informationsentnahme aus Karten, Tabellen und Diagrammen, die Fähigkeit zu freier und kreativer Text- und Bildgestaltung unabhängig von der Rechtschreibkompetenz, die Fähigkeit zu schlußfolgerndem Denken, insbesondere auch die Fähigkeit, komplexe Beziehungen in neuartigen Situationen wahrzunehmen, und zwar weitgehend unabhängig von dem Grad der Beherrschung der deutschen Sprache. Erfaßt wurden zugleich die unterschiedlichen familiären Hintergründe und Schülermerkmale wie Staatsangehörigkeit, Schullaufbahn, Besuch einer Vorschulklasse etc. sowie die Einstellungen der Kinder zum schulischen Lernen.

Die *Ergebnisse* der Untersuchung lassen sich – sehr vereinfacht – in folgenden Thesen zusammenfassen (vgl. Lehmann/Peek/Gänsfuß 1996):

(1.) Es gibt deutliche regionale Unterschiede im Leistungsniveau der Grundschulen. Ein großer Teil der beobachteten Varianz erklärt sich noch immer aus soziostrukturellen Merkmalen (insbesondere dem durchschnittlichen Buchbestand in den Elternhäusern, dem Schulabschluß der Eltern und dem Ausländeranteil), die das Umfeld einer Schule bestimmen. Dies gilt nicht in gleicher Weise für die Fähigkeit zum nicht sprachgebundenen schlußfolgernden Denken und schulbezogene Einstellungen von Schülerinnen und Schülern.

(2.) Vor diesem Hintergrund gibt es Schulen, die – gemessen an den Bedingungen ihres sozialen Umfeldes – erwartungswidrig gut, und solche, die erwartungswidrig schlecht abschneiden.

(3.) Bildet man Durchschnittswerte für alle Hamburger Schulen, so deuten sich möglicherweise auch für deren Leistungsniveau insgesamt Probleme im Verhältnis zu dem Niveau der Schulen in anderen Ländern (auch in anderen Großstädten) an.

Inzwischen werden im Rahmen der OECD unter dem Kürzel *PISA* („*Programme for International Student Assessment*“) weitere Vergleichsuntersuchungen vorbereitet, an denen sich ungefähr 30 Staaten, darunter auch die Bundesrepublik Deutschland, beteiligen werden. Projektnehmer auf internationaler Ebene ist das Australian Council of Educational Research (ACER)⁴. Die Durchführung der Untersuchungen in Deutschland ist einem Konsortium unter Führung des Max-Planck-Instituts für Bildungsforschung in Berlin übertragen worden. Gegenstand der Untersuchung ist ein Vergleich der Leistungen von 15-jährigen Schülerinnen und Schülern in den Bereichen Leseverständnis, Mathematik, Naturwissenschaften, wobei neben international vereinbarten Untersuchungsansätzen⁵ bestimmte Ergänzungen in den einzelnen beteiligten Ländern (nationale Optionen) vorgesehen sind. Im Rahmen der nationalen Optionen soll in Deutschland z.B. geklärt werden, welche aussichtsreichen pädagogischen Interventionsmöglichkeiten zur Verbesserung von Leseverständnis und Lesekompetenz sich identifizieren lassen.

Dabei geht es sowohl um die *pädagogisch-psychologische Basis von Fördermaßnahmen* (bereichsspezifisches Vorwissen, motivationale Faktoren wie Aufmerksamkeit und thematisches Interesse, Arbeitsgedächtnis und Verfügbarkeit von Lese- und Lernstrategien) als auch um die Situation von Risikogruppen („schwache Leser“ und Leser nicht-deutscher Muttersprache), um außerschulische Stützsysteme beim Erwerb der Lesekompetenz (Leseanlässe und Lesegelegenheiten, häusliche Lesekultur und Regeln der Mediennutzung, Wertschätzung des Lesens in den jeweiligen peer-groups), um die Rolle des Lesens im Sprach- und Literaturunterricht sowie seine Bedeutung und Nutzung als Instrument selbständigen Wissenserwerbs. Untersucht werden sollen ferner *fächerübergreifende Kompetenzen* (cross curricular competencies), so z.B. die Fähigkeit, extracurriculare Projekte zu planen und zu steuern, allein oder gemeinsam mit anderen technische Probleme zu lösen oder Lösungsstrategien durch Analogiebildung auf neue Aufgabenstellungen zu übertragen. Unter dem Aspekt „soziale Kompetenz“ soll die Fähigkeit untersucht werden, soziale Informationen zu dekodieren und zu interpretieren sowie ein Repertoire variabler Handlungsentwürfe für gängige Interaktionssituationen zu entwickeln. Die Untersuchungen sollen sowohl die sozialen Kontexte wie auch Aspekte der Schulqualität etc. einschließen.

Die Felderprobung der erforderlichen Tests wird bis zum Sommer 1999 durchgeführt werden. Die Hauptuntersuchung soll im Jahre 2000 erfolgen, der Bericht im Jahre 2001 vorliegen.

4 Beteiligt sind ferner das Netherlands Institute of Educational Management Cito, der Service de Pédagogie Expérimentale, Université de Liège (SPE), und Westat Inc., Rockville (USA), als Datenverarbeitungszentrum.

5 Die „Frameworks“ der geplanten Untersuchungen sind nach dem jeweils aktuellen Stand im Internet verfügbar (<http://www.pisa.oecd.org>).

Inzwischen hat die Kultusministerkonferenz beschlossen, die Untersuchungen über die für den internationalen Vergleich erforderliche begrenzte Stichprobe hinaus so anzulegen, daß *auch Leistungsvergleiche innerhalb der Bundesrepublik* möglich werden⁶. Ebenfalls beschlossen wurde von der Kultusministerkonferenz, eine weitere Vergleichsuntersuchung zur Fremdsprachenkompetenz der Schülerinnen und Schüler an deutschen Schulen durchzuführen. Diese Untersuchung ist inzwischen ausgeschrieben worden⁷.

5. Was weiß man und was tut man, wenn man untersucht hat?

Vorbereitung und Durchsetzung von empirischen Untersuchungen zum Leistungsstand von Schulen erfordern einen hohen Aufwand. Dieser steht in einem eigenartigen Gegensatz zu dem vergleichsweise geringen Aufwand, den man sich bisher noch leistet, um eine tragfähige Konzeption für die *strategische Nutzung der Ergebnisse solcher Untersuchungen* zu entwickeln. Zwar hat die Rückmeldung der Untersuchungsergebnisse der Lernausgangslagenuntersuchung in Hamburg an die Grundschulen und an die weiterführenden Schulen einen intensiven Diskussionsprozeß in Gang gesetzt. Er bezieht sich auf die Klassenbildung, die pädagogische Arbeit innerhalb der Klassen und die Entwicklung von spezifischen Fördermaßnahmen, die Zusammenarbeit der Lehrkräfte innerhalb einer Schule ebenso wie auf die Zusammenarbeit von Grundschule und weiterführenden Schulen, Aktivitäten der Schulberatung, der Lehrerfortbildung und der Standardsicherung durch Vergleichsarbeiten (vgl. Behörde ... 1998). Insgesamt wäre es jedoch falsch zu glauben, daß auf diese Weise allein bereits alle notwendigen *Impulse zur Veränderung der Situation* ausgelöst werden können. Jedes der oben mitgeteilten Ergebnisse dieser Untersuchung muß Folgerungen und weitere Klärungen auf unterschiedlichen Ebenen nach sich ziehen, die vielfach nicht von einer einzelnen Schule allein bewirkt werden können.

Die Aussagen in These (2) legen den Schluß auf *Ursachenfaktoren im Lehrerhandeln und im Konzept der einzelnen Schule* nahe. Um herauszufinden, was man tun sollte, wäre es vor allem interessant zu klären, was die erwartungswidrig guten Schulen anders machen als andere und ob bzw. unter welchen Bedingungen ihre Vorgehensweisen auf andere Schulen übertragbar sind.

Die Aussagen zu These (3) legen Vermutungen nahe, die sich in gleicher Weise auf die Situation aller Hamburger Schulen beziehen. Wenn man auf dieser Ebene handeln will, muß man klären, was andere Länder in gleicher Situation anders machen als Hamburg (curriculare Vorgaben? Struktur, Einstellungen und Qualifikation der Lehrerschaft? Steuerungsverfahren? oder ...?).

Der Umgang mit den Folgerungen aus These (1) ist fraglos am schwierigsten. Sie bestätigt, was längst bekannt ist, aber zeitweilig im Glauben an die Wirksamkeit vielfältiger Förderprogramme und den mit ihnen verbunde-

6 Beschluß der Amtschefskonferenz der KMK vom 17./18. September 1998.

7 Beschluß der Amtschefskonferenz der KMK vom 4./5. Februar 1999.

nen finanziellen Input zu sehr verdrängt wurde: Unterschiedliche Lernleistungen haben ihre Ursache auch in den sozialen Bedingungen, unter denen Schüler aufwachsen. Soweit sich dies daraus erklärt, daß Lernen eben nicht nur in der Schule stattfindet, sondern in starkem Maße auch von außerschulischen Stützsystemen abhängt, drängt sich die Frage auf, ob man dies resignierend auf sich beruhen lassen darf oder ob es nicht doch außerschulisch ansetzende wirksame Strategien gibt, soziale Unterschiede besser als bisher zu kompensieren. Entzaubert ist freilich der Glaube, daß man bereits durch eine Addition vielfältiger Förderprogramme ohne weiteres Probleme löst.

Bei alledem lassen sich viele der damit aufgeworfenen Fragen zuverlässig anhand von Querschnittsuntersuchungen noch nicht beantworten. Deshalb plant Hamburg *Folgeuntersuchungen für die Klassenstufen 7 und 9*. Sie werden gleichfalls flächendeckend angelegt sein. Die Tests für die Klassenstufe 7 sind abgeschlossen, aber noch nicht ausgewertet. Die Untersuchungen werden von der Frage geleitet, ob sich die Unterschiede in den Leistungen am Ende der Grundschule im weiteren Verlauf ausgleichen, ob sie unverändert bleiben oder ob sie sich vielleicht sogar verstärken. Erst auf dieser Basis ließen sich z.B. Entscheidungen über die weitere Entwicklung der Grundschulkonzeption treffen. Denn immerhin ist denkbar, daß auch bei schwächeren Leistungen zu einem bestimmten Zeitpunkt die Basis für positive Entwicklungen in der Folgezeit bereits gelegt ist.

Es zeigt sich mithin: Leistungsvergleiche erlauben *eine allgemeine Standortbestimmung* („bench-marking“) und geben erste Hinweise auf Handlungsnotwendigkeiten. Sie liefern jedoch nur begrenzt Aufschluß über mögliche Ursachenzusammenhänge und erst recht nicht bereits didaktische Konzepte zur Beseitigung erkannter Defizite. Dabei sind Aussagen stets nur in den Grenzen des gewählten Untersuchungsansatzes möglich. Dies darf indessen keine Rechtfertigung dafür sein, unbequeme Untersuchungsergebnisse mit dem Argument der begrenzten Reichweite von Untersuchungen vorschnell zu relativieren oder Untersuchungen nur dort anzusetzen, wo man sich für unangreifbar hält.

Vergleichsuntersuchungen sind nach alledem *nur der Einstieg in ein Feld, dem weitere Schritte in Form gezielter Untersuchungen folgen müssen*. Je nachdem, auf welcher Ebene man ansetzen will und welche Fragestellungen dabei leitend sein sollen, wird auch die Frage unterschiedlich zu beantworten sein, wie breit (flächendeckend oder stichprobenartig, bestimmte Gruppen), in welcher Zeitabfolge (einmalig, regelmäßig, anlaßbezogen) und auf welchen Veranlassungsebenen (Einzelschule zum Zwecke der Selbstevaluation, Steuerungswissen der Schuladministration) weitere Untersuchungen anzulegen sind. Dabei ist es eine offene Frage, ob gelegentliche Untersuchungen zu unterschiedlichen Fragestellungen ausreichen oder ob es einer regelmäßigen Berichterstattung über die Leistungsstände von Schülerinnen und Schülern bedarf. Auch wenn hinter dieser Frage ein erhebliches Maß an Sprengstoff stecken könnte, wird sie in absehbarer Zeit zu beantworten sein.

Die Folgerungen aus bereits gewonnenen Erkenntnissen und aus weiteren Untersuchungen werden das Handeln vieler Akteure erfordern. Sie beziehen

sich auf Änderungen in der Lehrerausbildung ebenso wie auf verstärkte Anstrengungen der Lehrerfortbildung, auf curriculare Entwicklungen und die Entwicklung didaktischer Konzepte⁸ wie auf Bemühungen im Rahmen der Schulentwicklung und des Schulmanagements.

6. Risiken und Nebenwirkungen ausgeschlossen?

Untersuchungen, wie sie derzeit durchgeführt werden, sollen Steuerungswissen erzeugen. Sie dienen nicht einer Überprüfung der individuellen Leistung einzelner Schüler oder einzelner Lehrer. Zentrale Tests *sollen weder die geltenden Formen der Leistungsbewertung* von Schülerinnen und Schülern (Noten, Zeugnisse) *noch der Beurteilung des einzelnen Lehrers ersetzen*. Sie sind weder die Grundlage der individuellen Entscheidung über Berechtigungen zum Besuch weiterführender Bildungseinrichtungen noch ein Element der Personalbeurteilung und Karriereplanung. Insofern ist die Situation in Deutschland der Situation in England oder in den Vereinigten Staaten nicht vergleichbar, wo Tests bereits in großer Breite zur regelmäßigen Leistungsbewertung im Vergleich einzelner Schulen untereinander („Ranking“) wie auf der individuellen Ebene der Schülerinnen und Schüler Verwendung finden. Eine solche Entwicklung zuzulassen, kann auch kein vernünftiges schulpolitisches Ziel sein. Trotzdem muß man sich darüber klar sein, daß auch die in Deutschland eingeleiteten Entwicklungen nicht frei von Risiken und Nebenwirkungen sind:

(1) Es können sich *Rückwirkungen für Curricula und Unterrichtsmaterialien* ergeben, die Ziel und Mittel umdrehen. Statt zu testen, ob man erreicht, was man erreichen will, versucht man zu erreichen, was sich auf eine einfache Weise testen läßt.

(2) Eine zunehmende Bedeutung von Testergebnissen für die Wertschätzung der einzelnen Schule im Urteil der Eltern kann Schulen zu *unsinnigen Reaktionen* veranlassen, z.B. Tests in extensiver Weise zu trainieren oder Schüler, von denen ein negativen Einfluß auf das Testergebnis erwartet wird, vom Test fernzuhalten.

(3) Die Funktion der Rechenschaftslegung („accountability“) kann die der Verbesserung („improvement“) auf Dauer überlagern und statt zu einer Qualitätssteigerung *zum Festhalten am (scheinbar) Bewährten verleiten*. Diese Gefahr wäre insbesondere dann gegeben, wenn sich die Meßverfahren in der Fixierung auf einmal gewählte Parameter gegenüber der kritischen Reflexion komplexer Zusammenhänge unterschiedlicher Dimensionen von Leistung und Qualität verselbständigen.

(4) *Standardisierung und Vereinheitlichung können die Reduzierung von Vielfalt bewirken*, die gerade das Kennzeichen eines föderalen Systems sein

8 Vgl. in diesem Zusammenhang als Reaktion auf die Ergebnisse von TIMSS etwa die Erklärung der Fachverbände DMV, GDM und MNU vom 21. Mai 1998, die die Notwendigkeit einer konzeptionellen Veränderung des Mathematikunterrichts in Deutschland unterstreicht, sowie das Gutachten von Baumert et. al. für die Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (BLK) zur Vorbereitung des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“, Materialien zur Bildungsforschung, Heft 60.

soll. Auch wenn es unsinnig ist, unterschiedliche Bildungsziele gegeneinander auszuspielen, so kommt man doch nicht daran vorbei, daß sich nicht alle Ziele (zumal angesichts der Restriktionen im Zeitbudget von Schülerinnen und Schülern) gleichzeitig optimal realisieren lassen. Man muß sich also entscheiden. Diese Entscheidung kann unterschiedlich ausfallen. Man muß sie in ihrer Unterschiedlichkeit gelten lassen, solange sie sich innerhalb einer Bandbreite hält, die Durchlässigkeit und Vergleichbarkeit nicht ernstlich gefährdet.

(5) Das Leistungsniveau der Schülerinnen und Schüler einer Schule hängt stark von deren sozialen *Kontextbedingungen* ab. Ein „*ranking*“ von Schulen, das diesen Aspekt vernachlässigt, ist nicht zulässig. Dennoch kann niemand garantieren, daß es angesichts der zu erwartenden plakativen Darstellung von Untersuchungsergebnissen zu hinreichend differenzierten Urteilen kommt. Eine Schule in objektiv ungünstiger Situation wird vermutlich auf Dauer eine ungünstige Wettbewerbssituation haben. Demotivation von Lehrkräften und Schülern könnte die Folge sein. Es ist offen, ob sich dieser Gefahr bereits dadurch ausreichend begegnen ließe, daß Gegenstand von Untersuchungen nicht der Leistungsstand zu einem bestimmten Zeitpunkt, sondern der Leistungszuwachs innerhalb eines bestimmten Zeitraums (value added approach) wird.

Die Gefahr von Fehlentwicklungen in dem damit angedeuteten Sinne ist durchaus real. Sie darf dennoch kein Grund sein, den Kopf in den Sand zu stecken und darauf zu verzichten, Urteilsmaßstäbe auch aus den Vergleichen mit anderen abzuleiten. Aber man muß sich der Möglichkeit einer mißbräuchlichen oder unkundigen Verwendung von Untersuchungsergebnissen sehr bewußt sein. Auch wenn niemand daran denkt, derartigen Entwicklungen Vorschub zu leisten, kann man nicht sicher sein, daß alle den Versuchen auf Dauer widerstehen können. Die Faszination einer scheinbaren Objektivität, die sich ergibt, wenn man in einem allzu simplen Urteil die vielfältigen Aspekte vernachlässigt, die bei der Interpretation von Untersuchungsergebnissen bedacht sein wollen, ist nicht gering. Welche Dynamik den eingeleiteten Entwicklungen innewohnt, läßt sich deshalb derzeit kaum abschätzen. *Wachsamkeit ist unerlässlich.*

Literatur

- Abschlußbericht der von der Kultusministerkonferenz eingesetzten Expertenkommission zur „Weiterentwicklung der Prinzipien der gymnasialen Oberstufe und des Abiturs“, Bonn, 1995
- Angvik, M. / von Borries, B. (Hg.): Youth and History: A Competitive European Survey on Historical Consciousness and Political Attitudes among Adolescents, Hamburg, 1997
- Baumert, J. / Bos, W. / Watermann, R.: TIMSS/III – Schülerleistungen in Mathematik und den Naturwissenschaften am Ende der Sekundarstufe II im internationalen Vergleich. Zusammenfassung deskriptiver Ergebnisse, Berlin, 1998
- Baumert, J. / Lehmann, R. / Lehrke, M. / Schmitz, B. / Clausen, M. / Hosenfeld, I. / Köller, O. / Neubrand, J.: TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im Vergleich. Deskriptive Befunde, Opladen, 1997
- Behörde für Schule, Jugend und Berufsbildung Hamburg: Standardisierte Leistungsmessung als Instrument der Qualitätssicherung im Bildungswesen. Materialien für die Referendarinnen und Referendare, Hamburg, 1998

- Der Bundesminister für Bildung und Wissenschaft: Grund- und Strukturdaten 1988/89
- Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (bmb+f): Grund- und Strukturdaten 1997/98
- Gutachten zur Vorbereitung des Programms „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“ der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung, Materialien zur Bildungsforschung, Heft 60, Bonn, 1997
- Kommunale Gemeinschaftsstelle (KGSt), Köln: Das Neue Steuerungsmodell (Begründung, Konturen, Umsetzung), Bericht Nr. 5/1993
- Kommunale Gemeinschaftsstelle (KGSt), Köln: Qualitätsmanagement, Bericht Nr. 6/1995
- Lange, H.: Schulautonomie. Entscheidungsprobleme aus politisch-administrativer Sicht, in: Zeitschrift für Pädagogik. 1995, S. 21 ff.
- Lehmann, R. / Peek, R. / Gänsfuß, R.: Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen. Bericht über die Untersuchung im September 1996, Hamburg, 1997
- Liket, T. M. E.: Freiheit und Verantwortung. Das niederländische Modell des Bildungswesens, Gütersloh, 1993
- Olson, L.: Shining a Spotlight on Results. In: Quality Counts '99: Rewarding Results, Punishing Failure, Education Week, Volume XVIII, Number 17, January 1999
- Postman, N.: Amusing Ourselves to Death. Public Discourse in the Age of Show Business, New York, 1985
- Sachverständigenrat Bildung bei der Hans-Böckler-Stiftung: Für ein verändertes System der Bildungsfinanzierung, 1998
- Weiß, M.: Mehr Ressourcen = mehr Qualität? in: Böttcher, W. / Weishaupt, H. / Weiß, M. (Hg.): Wege zu einer neuen Bildungsökonomie. Pädagogik und Ökonomie auf der Suche nach Ressourcen und Finanzierungskonzepten, Weinheim und München, 1997
- Hermann Lange*, Staatsrat in der Behörde für Schule, Jugend und Berufsbildung, Hamburg
- Anschrift: Behörde für Schule, Jugend und Berufsbildung, Hamburger Str. 31, 22083 Hamburg