

Achim Brosziewski

Bildungsqualität, statistische Depressionen und das Gedächtnis des Bildungssystems



Pädagogische Hochschule Thurgau
Forschung
Nationalstrasse 19
Postfach
CH-8280 Kreuzlingen
Schweiz

Tel. + 41 (0)71 678 56 56
Fax + 41 (0)71 678 56 57
office@phtg.ch
www.phtg.ch

© PHTG 2005 | Neuauflage in angepasster Form als elektronisches Dokument: Dezember 2005

Inhaltsverzeichnis

Zusammenfassung	4
Einleitung	5
1 Die Vermessung von Bildungsqualität.....	5
2 Statistische Depressionen.....	8
3 Systemgedächtnis und drei Strategien der Bildungsforschung	10
Literatur	16

Zusammenfassung

Der Beitrag skizziert zunächst die historische Entwicklung der Messung von Bildungsqualität in den letzten 40 Jahren. Unter dem Stichwort der "statistischen Depression" stellt er sodann ein Paradox der Kommunikation über Zahlen vor. Jedes "Mehr" verweist notwendigerweise auf ein "Weniger" und umgekehrt. Wird dieses Paradox auch noch mit der Bewertung durch "besser und schlechter" verknüpft, birgt die Kommunikation über Zahlen für das Bildungssystem die Gefahr, sich in endlosen Diskussionsschleifen zu verlieren. Stattdessen könnte eine theoretische Einordnung dazu führen, die Qualitätsmessungen als eine Veränderung des Gedächtnisses des Bildungssystems zu begreifen. Davon ausgehend zeichnet der Beitrag drei Strategien nach, mit denen sich die Bildungsforschung auf solch eine Gedächtnisveränderung einstellen kann: eine Strategie der Faktorisierung von Erfolgsvariablen, eine Strategie der Vermessungskritik sowie eine Strategie der Operationalisierung von Systembeobachtungen.

Einleitung

Seit es Erziehung und Bildung gibt, werden sie von der Sorge um ihre Qualität begleitet. Ein klassisches Beispiel wäre PLATONS Programm zur Elitenbildung ("Wächter der Stadt"), in der er eine unerlässliche Bedingung eines idealen Staates sah (Politeia, 376c-412b). Seit dem späten Mittelalter hatten der Adel und das Bürgertum zu erwägen, welche Hauslehrer, Schulen und Universitäten ihre Söhne am besten auf den künftigen Status vorbereiteten (vgl. FRIED 1986; MOELLER/PATZE/STACKELMANN 1983). Mit der Einrichtung einer in Schulen organisierten Erziehung für alle Bevölkerungsschichten, also seit dem frühen 19. Jahrhundert, trat die Qualifikation der Schullehrer in den Vordergrund einer nunmehr allgemeinen und öffentlichen Sorge um die Bildung der Vielen (vgl. GRUNDER 1996). Nicht viel später, so gegen Ende des 19. Jahrhunderts, sind bereits die ersten Versuche zu verzeichnen, durch Messungen zu prüfen, was denn in Schulen tatsächlich erreicht wird (MADAUS/AIRASIAN/KELLAGHAN 1980, S. 5-8). Die gegenwärtige Diskussion um die Qualität der Bildung und ihre Messung hat mithin traditionsreiche Bezüge. Doch wurde sie seit den 60er Jahren des letzten, des 20. Jahrhunderts auf ein zuvor nicht bekanntes Niveau gebracht. Der nachfolgende Beitrag setzt hier an. Der erste Abschnitt skizziert Entstehung und Form dieser jüngeren Entwicklung der Bildungsqualitätsmessung. Anschliessend wird der Begriff der "statistischen Depression" vorgestellt, der eine sehr grundlegende Problematik in der Kommunikation über Zahlenverhältnisse bezeichnet. Zahlenkommunikation bewegt das Paradox von "mehr ist weniger", das zusammen mit der Bewertung nach "besser und schlechter" in endlose Zirkel führen kann. Der abschliessende dritte Teil führt die These ein, dass sich mit der Verbreitung von Messungen das *Gedächtnis des Bildungssystems* verändert. Es werden drei Strategien nachgezeichnet, mit denen sich die Bildungsforschung auf diese Strukturänderung einstellen kann: eine Strategie der Faktorisierung von Erfolgsvariablen, eine Strategie der Kritik des Messens sowie eine Strategie der Operationalisierung von Systembeobachtungen.

1 Die Vermessung von Bildungsqualität

Für die heutige Lage kann eine Studie aus dem Jahre 1966 als form- und geschichtsbildend angesehen werden – eine Untersuchung, die James COLEMAN und andere (1966) im Auftrag für den US-amerikanischen Kongress über „Equality of Educational Opportunity“ angefertigt hatten und die seither als „COLEMAN-Report“ häufig zitiert wird. Diese Studie, die erstmals landesweite Erhebungen von Schüler-, Schul-, Lehrer- und Bildungspolitikmerkmalen sowie von Schülerleistungen durchführte, kam zu dem damals brisanten Schluss, dass die materielle und personelle Ausstattung der Schulen statistisch gesehen keinen nennenswerten Einfluss auf die Leistungen, die Ambitionen und die Karrierechancen ihrer Schüler hätten. Ausschlaggebend wäre vielmehr der Herkunftsstatus der Schülerinnen und Schüler (sozio-ökonomische Lage und ethnisch-rassische Zugehörigkeiten, S. 21-22 u.ö.). Demnach machte Schule keinen Unterschied für die Chancen der Heranwachsenden aus. Solche Feststellungen wirkten in einer Zeit, die politisch grosse Hoffnungen und Ressourcen auf Bildung als Mittel zu mehr Chancengleichheit setzte, sehr ernüchternd.

Bereits mit ihrem Erscheinen waren diese Ergebnisse starken Einwänden ausgesetzt. Schon die Re-Analyse der im COLEMAN-Report verwendeten Daten gab differenziertere Bilder ab (JENCKS U. A. 1972; MOSTELLER/MOYNIHAN 1972). Hinzu kam eine intensive Kritik der Kriterien, nach denen Schulerfolg bemessen wurde, sowie der Indikatoren und der Messverfahren. Viele Studien folgten, in denen die COLEMAN-Thesen teils bestätigt, insgesamt jedoch stark eingeschränkt und modifiziert wurden (siehe WYATT 1996). Die weitere Entwicklung erhielt Impulse aus vielen anderen Ländern, vor allem aus europäischen und asiatischen. Die Konstruktion der Variablen, die Schulerfolg beschreiben sollen, wurde immer weiter verfeinert, ebenso die Konstruktion der Indikatoren und der Messverfahren. Einen wichtigen Teil bildeten die Schülerleistungstests in Fächern wie Sprache, Mathematik und Naturwissenschaft, so ungefähr in der Form, wie sie durch PISA nunmehr international bekannt sind. Doch werden, weniger bekannt und öffentlichkeitswirksam, zahlreiche andere Grössen erhoben, die sich auch auf Merkmale der Schulen, auf den Ausbildungsstand der Lehrenden und auf die Bildungspolitik selbst beziehen. Eine Übersicht des U.S. Office for Education aus dem Jahre 2000 stellt beispielsweise 13 verschiedene Variablen zusammen, die im Rahmen eines umfassenden Bildungsmonitorings den Zustand des nationalen Bildungssystems abbilden sollen.¹ Eine jüngere Zusammenstellung für den Aufbau des schweizerischen Bildungsmonitorings diskutiert, auf der Grundlage der Entwicklung in vielen anderen europäischen Ländern, insgesamt 18 solcher Variablen.²

Eine zweite Forschungsrichtung zur Bildungsqualität deckt sich weder im Ansatz noch in den Aussageabsichten mit der vorab skizzierten Entwicklung.³ Für sie steht nicht ein nationales oder regionales Bildungssystem im Fokus. Zentraler Gegenstand und Erhebungseinheit ist vielmehr die einzelne Schule. Methodisch korrespondieren mit diesem Ansatz die Fallstudie, die Erschliessung von Schulgeschichten sowie die Ermittlung der Beteiligtenperspektiven: der Lehrer und Lehrerinnen, der Schülerinnen und Schüler, der Schulleitungen und der Eltern. Auch in dieser Richtung kommen Fragebögen und Messinstrumente zum Einsatz, und sie sucht ebenfalls nach generellen Aussagen, die über die Einzelfälle hinausweisen.

Beispielhaft für den deutschsprachigen Raum können die Untersuchungen von Helmut FEND angeführt werden.⁴ Seit den frühen 70er Jahren hat FEND, zunächst unter dem Begriff "Schulklima", in umfangreichen Studien ein vielfältiges Set an Kategorien und fragebogengestützten Instrumenten entwickelt, um die verschiedenen Perspektiven abzubilden, die im System der Schule zusammenkommen. Anhand von Schulbeobachtungen und von Fragebögen erfasst FEND drei Ebenen: die Ebene der Lehrenden, der Schüler und Schülerinnen sowie der Eltern. Jeder dieser drei Ebenen können spezifische Problemdimensionen zugeordnet werden (vgl. FEND 1998, S. 90, Abb. 3.4), wie etwa

¹ Fünf Indikatoren für die Dimension "School Context", vier für die Dimension "Teachers", vier für die Dimension "Classroom" (MAYER/MULLENS/MOORE 2000).

² Dimensionen "Kontext" (3 Indikatoren), "Input" (6), "Prozesse" (3), "Ergebnisse" (3) und "Wirkungen" (3) (STAMM 2003, S. 29). Siehe auch RÜRUP 2003. "Bildung auf einen Blick", verspricht die Organisation for Economic Co-operation and Development (OECD) (2000) mit ihrer Zusammenstellung von Bildungsindikatoren.

³ Vgl. STRIVERS 1985, mit der Unterscheidung von "school effectiveness"- und "school climate"-Forschung (S. 46).

⁴ Für die englischsprachige Tradition siehe nochmals STRIVERS 1985. Zur aktuellen Diskussion siehe SCHÖNIG 2002.

auf der Ebene der Lehrenden: Arbeitszufriedenheit, kollegiale Beziehungen, Berufseinstellung, Wahrnehmung der Schulleitungskompetenz u.a.m. Die Schüler werden unter anderem nach dem Lehrerverhalten befragt, nach dem Umgang mit abweichendem Verhalten, nach ihrer Schulfreude oder Schulverdrossenheit. Den Eltern werden Fragen vorgelegt, in denen sie die schulischen Leistungs- und Disziplinanforderungen, die Förderung ihrer Kinder und Ähnliches mehr einschätzen.

Auch solche eher qualitativen Fragen werden quantitativ operationalisiert. Man gibt zu jeder Dimension bestimmte Aussagen vor, denen die Befragten dann entweder mehr oder weniger oder gar nicht zustimmen. Zur Kompetenz der Schulleitung wurden in einer Untersuchung aus dem Jahr 1977 etwa folgende Statements vorgelegt, die von den Lehrenden auf einer Skala von Zustimmung bis Nichtzustimmung zu beantworten waren (FEND 1998, S. 93):

"Der Schulleiter setzt mit seinem Engagement ein gutes Beispiel für die Lehrer."

"Der Schulleiter tendiert dazu, wichtige Entscheidungen so lange wie möglich hinauszuzögern."

"Der Schulleiter läuft Gefahr, in kritischen Situationen den Überblick zu verlieren."

Für die Analyse ist zunächst interessant, wie sich die verschiedenen Faktoren des Schulklimas untereinander verhalten, also wie etwa Arbeitszufriedenheit mit dem Kollegialverhältnis korreliert oder auch die Schulfreude der Schüler mit dem Berufsverständnis der Lehrenden und vieles andere mehr. Wirklich brisant wird es jedoch erst auf einer zweiten Analyseebene: im Vergleich der verschiedenen Schulen untereinander. Die genannten Faktoren lassen sich je nach Zustimmungswerten auf Skalen von "sehr gut" bis "sehr schlecht" abtragen. Die Kollegialität kann besser oder schlechter sein, ebenso die Schulleitungskompetenz, die Schulfreude, usw. Diese Struktur ermöglicht es, die untersuchten Schulen in eine Rangfolge zu bringen. So unterschied FEND (ebd., S. 118f.) in einer sehr umfangreichsten Studie 63 Schulen anhand ihrer Ränge in zehn Dimensionen und kontrastierte sodann die bestplatzierten mit jenen in den untersten Rängen, um zu ermitteln, welches möglicherweise die spezifischen Differenzen sind, die die Differenz von "guten" und "belasteten" Schulen ausmachen.

Ein zentrales Ergebnis, das sich auch in späteren Studien immer wieder zeigt (FEND 1996), war, dass kein einzelner Faktor identifiziert werden kann, der auf alle wichtigen Qualitätsdimensionen "durchschlagend" wirkte (FEND 1998, S. 143). Zwar ist, um ein Beispiel zu geben, die Kollegialität innerhalb der Lehrerschaft recht einflussreich auf andere Qualitätsdimensionen, etwa das Gefühl der Akzeptanz auf Seiten der Schülerinnen und Schüler. Aber es lässt sich wiederum kein Zusammenhang mit der Leistungsbereitschaft der Schüler feststellen. Erst recht wirkt sich die Kollegiumsqualität nicht auf die Urteile der Eltern über die jeweilige Schule aus. Im englischsprachigen Raum wird die Schule schon seit langem als ein "lose gekoppeltes System" (WEICK 1976) beschrieben, in dem vieles mit vielem anderem unverbunden geschieht. Solch eine lose Kopplung ist auch für die Qualitätsdimensionen und Qualitätsfaktoren der Schulen festzuhalten. Leistungseinstellungen, Schulkulturen, Persönlichkeitschancen und Zertifikatbewertungen mögen miteinander variieren, aber durchgängige Ursache-Wirkungs-Beziehungen zwischen all diesen Faktoren sind nicht zu ermitteln.

Festhalten lässt sich: Es gibt unterschiedliche Wege, Qualitätsdimensionen zu objektivieren und in Messungen zu überführen. Seit den genannten Anfängen in den 60er und 70er Jahren ist ein sehr elaboriertes Feld der Qualitätsmessung im Bildungssystem entstanden, das eine grosse Zahl von Qualitätsindikatoren hervorgebracht hat. Deren theoretische Integration ist hingegen noch nicht gelungen (vgl. NUSSBAUM 2001, S. 3-7; WEISS/ZUTAVERN 2003, S. 7). Auch erfasst die Bildungsqualitätsmessung nicht nur das Schulwesen, sondern ebenso die Hochschulen (öffentlich bekannt: die Rankings) sowie Institutionen der öffentlichen und privatwirtschaftlichen Aus- und Weiterbildung. Bildungseinrichtungen, die nicht vermessen werden und sich selbst nicht vermessen, sind heutzutage kaum mehr zu finden. Was die 90er Jahren diesem Trend noch hinzugefügt haben, ist die Internationalisierung der Qualitätsmessungen und -vergleiche (BOS/POSTLETHWAITE 2001), von denen das "Programme for International Student Assessment" (PISA) nur eines, wenn auch das öffentlich bekannteste Vorgehen darstellt.

2 Statistische Depressionen

Der Ausdruck „statistische Depression“ stammt von Niklas LUHMANN (2000). In einem kleinen Zeitungssessay macht er auf eine Eigenheit von Zahlenwerten und Zahlenvergleichen aufmerksam, die sehr selten beachtet, geschweige denn systematisch analysiert wird. Jede Zahl – egal welche Grösse sie ausdrücken soll – macht nur Sinn im Vergleich zu anderen Zahlen. Eine 20 ist eine 20, weil sie keine 19 und keine 21 ist; auch nicht 20,3 oder 20/100 usw. Zahlen nehmen Positionen in Reihen ein, und erst das macht sie informativ: 10 Jahre – 11 Jahre – 12 Jahre usw. wäre solch eine Reihe, 1 Tor – 2 Tore – 3 Tore eine andere. Oft werden solche Reihen als Kontinuen interpretiert, etwa das Leben von Anfang bis Ende, in der eine Zahl eine bestimmte Position markiert ("12 Jahre alt"). Die unbeachtete Eigenheit, von der hier die Rede sein soll, liegt darin, dass jede Zahl aber nicht nur eine Reihe aufruft, sondern zwei Reihen zugleich: eine Reihe, auf der es mehr wird, und eine zweite Reihe, auf der es weniger wird. Graphisch lässt sich der Unterschied wie folgt veranschaulichen:

Die "20" als Position in einem Kontinuum:

------(19)-----20------(21)-----

Die "20" als Doppelposition in ihren beiden Reihen:

"wird mehr": ----->------(19)----->-----20----->------(21)----->-----

"wird weniger": ----->------(21)----->-----20----->------(19)----->-----

Was in dieser Umstellung zum Ausdruck kommt, ist, dass die 20, obwohl eindeutig sie und keine andere Zahl, nicht dieselbe ist, je nachdem, von welcher Richtung man kommt. Einmal ist sie mehr als alle bisherigen Zahlen, das zweite Mal ist sie weniger als alles Vorherige. Was hier wie reine Gedankenspielererei erscheinen mag, kann praktisch enorme Auswirkungen haben. Rein gedanklich könnte man ja sagen, eine 40h-Arbeitswoche ist eine 40h-Arbeitswoche. Man arbeitet nicht 39 Stunden und nicht 41 Stunden pro Woche. Aber wenn es um eine „Einführung der 40h-Arbeitswoche“ ginge, machte es einen enormen Unterschied, ob in Deutschland oder in der Schweiz darüber diskutiert würde. In der Schweiz käme man von gesetzlichen 42 Stunden her und die 40 wäre ein „Weniger als zuvor“. Das Gegenteil gälte für Deutschland mit Arbeitsregelzeiten von

38,5 Stunden oder weniger. Die 40 wäre ein „Mehr als zuvor“. Zur regional-nationalen kommt noch die Ambivalenz in der historischen Dimension hinzu. Bis zum Wendepunkt 1967 war die 40 in der alten Bundesrepublik ein "Weniger als zuvor", seither wäre sie ein "Mehr", das gegenwärtig wieder aktuell geworden ist. Dieselbe Zahl bedeutet einmal „Mehr“, einmal „Weniger“. In der Eindeutigkeit von Zahlen ist also stets das Paradox der Zahlenreihen versteckt, das die allgemeine Form hat: „Mehr ist weniger“ und „Weniger ist mehr“. Es liessen sich zahlreiche andere Beispiele durchgehen, das Lohneinkommen etwa oder die Arbeitslosenquoten, auch Ranglistenplätze wie jene in der Fussballbundesliga oder in der Nationalliga A.

Nun sind Paradoxien zunächst einmal etwas ganz Normales (PUTNAM 1986; STONE 1997). Vor allem haben sie nicht von sich aus etwas mit Depressionen zu tun. Sie verweisen einfach auf die Notwendigkeit, sich zu entscheiden, die Dinge entweder so herum oder aber andersherum zu betrachten. Ein Depressionspotential wird erst dann gebildet, wenn die Unterscheidung von „mehr und weniger“ durch die Unterscheidung „besser und schlechter“ überdeckt wird. Oft heisst es „Mehr = Besser“, wie etwa beim Bruttosozialprodukt. Zuweilen ist auch „Weniger = Besser“ und „Mehr = Schlechter“, wie bei den Arbeitslosenraten oder bei der Rate der Schulabbrüche. Der Begriff „Qualität“ wäre ohne solche Wertungen gar nicht denkbar. Das Depressionspotential von Zahlen liegt mithin in der Tatsache, dass weder bei Zahlen noch bei Bewertungen die eine Seite ohne die andere Seite zu haben ist. Selbst wenn über Erfolge geredet wird, erneuert sich auf der Ebene des Gedächtnisses die Erinnerung an die Gefahr von Misserfolgen. Wer von „mehr“ redet, macht – ob er will oder nicht – darauf aufmerksam, dass es „weniger“ sein oder werden könnte. Wer von „besser“ redet, macht darauf aufmerksam, dass es „schlechter“ sein oder werden könnte.

Im Bezug auf soziale Systeme (wie das Wirtschaftssystem, das Bildungssystem, die Massenmedien oder die Politik) meint die Rede von Depression nun nicht den Umstand, dass es manche Leute froh oder traurig stimmen mag, wenn sie laufend mit Zahlenvergleichen konfrontiert werden.⁵ Soziale Systeme bestehen nicht in Stimmungen, sondern in Kommunikationen (LUHMANN 1984, S. 191-241). In sozialer Hinsicht steht der Ausdruck „Depression“ für den Umstand, dass bestimmte Kommunikationen und soziale Initiativen unterlassen werden, wenn ihnen durch andere Kommunikationen die Aussicht auf Resonanz und positive Rückmeldungen entzogen wird.⁶ Durch PISA und ähnliche Schülerleistungsmessungen werden *bestimmte* Kategorien von „guten“ und „schlechten“ Leistungen etabliert und die *kommunizierbaren* Sorgen beginnen, sich auf diese Dimensionen zu fokussieren.⁷ Unterrichts- und Schulprogramme, die diese Kategorien nicht berühren, stehen dann zunächst

⁵ Siehe zu intellektuellen Stimmungen rund um den "Zahlenwahn" das gleichnamige Dossier in Schweizer Monatshefte, 84. Jg., 2004, H. 9/10. Nicht zuletzt warnen hier Statistiker und Ökonomen. Aber das Paradox der Doppelreihigkeit von Zahlen und ihrer eindeutigen Zweideutigkeit haben auch sie nicht bemerkt.

⁶ Formaler wäre von kommunikativer Inhibierung zu sprechen (LUHMANN 1984, S. 431, im Anschluss an MEAD 1938, S. 343 u.ö.: "inhibition").

⁷ Dass die PISA-Messungen trotz ihrer allgemein anerkannt hohen Qualität selektiv im Blick auf Leistungsbereiche sind, ist eine zumindest unter Fachleuten gängige Einsicht (MERKENS 2003). Ein Beispiel wäre die Leistung "Lesen", die nach einem *bestimmten Modell* von Lesen gemessen wird (vgl. auch BRAKE 2003).

einmal *ausserhalb* all jener Kommunikation, die sich an den deklarierten Verbesserungen messen lassen will (vgl. KEMPFERT/ROLFF 2000; TILLMANN 2001). Wenn Bilder von „guten Schulen“ und „exzellenten Hochschulen“ zirkulieren und durch Ranglisten jede Einzelschule oder Hochschule permanent daran erinnert wird, dass die nächste Abmessung an diesen Idealen bald bevorsteht, stehen Initiativen, deren Ziele jenseits der Abbildungsraster liegen, vor Formulierungsproblemen. Ob sie dann überhaupt noch vorkommen, und wenn, ob sie sich mit den erfolgsnotwendigen Verständigungen und Ressourcen ausstatten könnten, scheint fraglich.⁸ Dass Messungen, die den Bildungsprozess steuern sollen (monitoring), den Prozess auch unabhängig von den Steuerungsintentionen bereits selektiv verändern, ist im Fall staatlich vorgeschriebener Leistungstests bereits festgestellt. Beispielsweise werden die Lernformen auf die Testkonstruktionen abgestimmt, der Lehrplan wird auf die Testthemengebiete hin verengt, testschwächere Schülerinnen und Schüler werden aus den Testzielgruppen herausgezogen.⁹

Selbst erfolgreiche Massnahmen sind auf Dauer nicht vor sozialen Depressionseffekten gefeit. Wenn die Unterscheidung von besser und schlechter regiert, kann selbst das Erreichen von „guten Effekten“ als ungenügend behandelt werden. „Das Bessere“, so formulierte schon VOLTAIRE unser Problem auf das Treffendste, „ist der Feind des Guten.“¹⁰ Vor allem aber darf das Einpendeln von Erwartungen auf einem gegebenen Niveau nicht unterschätzt werden. Wenn trotz "gezielter" Massnahmen keine Effekte auf den Monitoren ablesbar werden, geraten die Programme in den Strudel der Zweifel an ihrer "Effektivität" – und damit, wie das Beispiel der Wirtschaftsdaten lehrt, eröffnen sich endlos scheinende Zirkel der Diskussion um "Ursachen" und "Wirkungen", die politisch rasch deprimieren (im genannten Sinne der Kommunikationserschwerung; aus Praxiserfahrung siehe BRÜDERLE 2004).

3 Systemgedächtnis und drei Strategien der Bildungsforschung

Im Hinblick auf Leistungstests, Bildungsmonitoring und Schulqualitätsmanagement ist festzuhalten: Die Beobachtung des Bildungssystems wird in vielfältiger Weise, intensiv und extensiv auf die Differenz von "mehr und weniger" eingestellt. Dabei kommt die zuvor beschriebene Paradoxie des Zahlenvergleichs zum Tragen, dass jedes "mehr" auch ein "weniger" mit anzeigt. Es kommt immer nur

Andere Lesemodelle erforderten andere Messformen – und die Frage ist, ob und wie sich solch anderen Modelle unter der PISA-Beobachtung des Unterrichtens noch werden halten lassen.

⁸ WEISS und ZUTAVERN (2003) machen an diesem Punkt ein grosses Forschungsdesiderat aus: Trotz der (zunehmenden) Verbreitung des Rankings gäbe es keine Untersuchungen zu deren Folgen für die Schulen.

⁹ AMREIN/BERLINER 2002, mit der Nennung weiterer nicht-intendierter Folgen des "high-stakes testing": höhere Schulabbruchsquoten, Verschärfung minoritäts- und schichtungsbedingter Bildungsnachteile, gemeinschaftlicher Testbetrug durch Schüler und Lehrer, vermehrte Berufsausstiege in der Lehrerschaft, Absenkung höherer Lernniveaus in Richtung der jeweils definierten Standards (im Sinne von "Standards produzierten Mittelmass"), Preisanpassungen der Immobilien an die publizierten Testergebnisse benachbarter Schulen, ...

¹⁰ Dagegen eine typische Forderung von heute: "From Good to Great" (BARBER 2003, Tony BLAIRS wissenschaftlicher Bildungspolitikberater).

darauf an, wie eng oder wie weit man den Vergleichshorizont in sachlicher, sozialer und zeitlicher Hinsicht anlegt. Selbst der oder die oder das Beste ist immer noch "weniger", als er oder sie oder es in Zukunft sein könnte. Da die Differenz von "mehr und weniger" kommunikativ kaum behandelt werden kann, ohne sie mit "besser und schlechter" zu verknüpfen, ist das Potential für statistische Depressionen angelegt. Dieser Ausdruck bezeichnet hier ja nicht (nur) schlechte Stimmungen, sondern in erster Linie den Entzug von Resonanzchancen für Kommunikationen, Initiativen und Projekte, die sich nicht als "effektiv" innerhalb der ausgelegten Vergleichs- und Bewertungsraster ausweisen können.

Demgegenüber sind einige positive Effekte der Vermessung von Bildungsqualität festzuhalten, auf die das Bildungssystem kaum mehr verzichten können. Erstens unterstützt sie eine Art Entideologisierung der "Diskurse" über Bildung – an den Akademien, in der Intellektualität, in der Politik. Wer heute etwas zu Bildung sagen will, kommt nicht umhin, Zahlenverhältnisse und die von ihnen indizierten Sachverhältnisse zu berücksichtigen. Zweitens sichern Bildungsdaten der Kommunikation über Schule und Bildung eine gewisse Kontinuität im massenmedialen und politischen Wechsel von Themen und Prioritäten. Mit dieser Sichtbarkeit dürfte der Bildungspolitik innerhalb der Politikfelder ein grösseres Gewicht zukommen als bisher. Drittens und schliesslich werden im weitgesteckten zeitlichen, regionalen und internationalen Vergleich sowohl Konsistenzen als auch Inkonsistenzen sichtbar, was Anstrengungen um Theorien und Konzepte stimulieren könnte.

Diese drei Punkte lassen sich in der These zusammenfassen: Das Netz an Bildungsdaten versorgt das Bildungssystem auf weltweiter Ebene mit einem eigenen Gedächtnis, das Registrier- und Sortierfunktionen übernimmt, im Blick auf das, was alles geht und was nicht geht, gehen könnte, gehen sollte.¹¹ Davon ausgehend ergibt sich für die *Bildungsforschung* die Frage, wie sie auf solch eine Veränderung ihres Gegenstandes, auf den Aufbau eines weltweiten Bildungsgedächtnisses, einstellen kann. Soweit zu erkennen und ohne Anspruch auf Vollständigkeit lassen sich drei grundlegende Strategien ausmachen: eine Strategie der *Faktorisierung von Erfolgsvariablen*, eine Strategie der *Vermessungskritik* sowie eine Strategie der *Operationalisierung von Systembeobachtungen*. Abschliessend sollen diese drei Strategien in ihren Konturen nachgezeichnet werden.

1. Faktorisierung von Erfolgsvariablen: Faktorisierung soll hier der Kurzname sein für das Auflösen der gemessenen Daten in Faktoren, die im Schema von Ursache und Wirkungen wieder rekombiniert und geordnet werden.¹² In der Methodologie ist bekannt, dass statistische Werte nicht mit Ursachenbeschreibungen verwechselt werden dürfen (Differenz von Korrelation und Kausalität), und dass die Feststellung von Kausalität eine Frage der Theorie und nicht eine Frage des Messens darstellt. Hier gilt es auszuwählen, welche der gemessenen Grössen als "Erfolge" der Bildung, der Bildungspolitik, der Schulen und / oder des Unterrichts in Schulen gelten sollen, und welche Grössen die Ursachen dieser Erfolge bezeichnen sollen. Mit dieser Strategie wird die Bildungsforschung sowohl auf die gesellschaftliche Umwelt des Bildungssystems als auch auf sich selbst verwiesen: auf

¹¹ Dieser Effekt wird bislang unter dem Ausdruck "Globalisierung" verbucht, so bspw. bei MESSNER 2003.

¹² Beispiele für diese Strategie finden sich im Zusammenhang mit Leistungsmessungen sehr viele, siehe nur die Beiträge in WEINERT 2002 mit zahlreichen Hinweisen auf weitere Literatur. Ein anderes Beispiel wäre die oben referierte Analyse zu den Faktoren der „guten Schule“.

die Umwelt, da die Definition von "Erfolg" dem Bildungssystem nicht selbst überlassen bleibt, sondern im Kontakt mit anderen Kommunikationssystemen wie der Politik, der Wirtschaft, den Massenmedien, den Familien, den bildungsinteressierten Individuen usw. hergestellt wird; auf sich selbst, da Theorien über kausale Abhängigkeiten nicht unabhängig von den Theorietraditionen der Wissenschaft behauptet werden können. In Bezug zur sozialen Umwelt kann sich die Faktorisierung eher "technologisch" oder eher "kritisch" einstellen – "technologisch", wenn sie für die Umwelt Variablen identifiziert, mit denen sich Bildungserfolge "steuern" lassen (OELKERS 2002); "kritisch", wenn sie die Alltagstheorien über Bildungserfolge durch Kausalanalysen zu korrigieren versucht.¹³

Beide Stossrichtungen, die technologische wie die kritische, dürften die Faktorisierung von Ursachen und Wirkungen weiter vorantreiben, zu immer feineren Auflösungen und immer raffinierteren Kausalmodellen führen. Welche Chancen und Probleme in dieser Ausrichtung für die Bildungsforschung liegen könnten, mag ein Seitenblick auf die Geschichte der Volkswirtschaftslehre veranschaulichen, die seit den 60er- und 70er-Jahren des vergangenen Jahrhunderts vergleichbaren Bedingungen ausgesetzt war und mit ähnlichen Tendenzen reagiert hat. Sie konnte über lange Phasen davon profitieren und sich ihrer gesellschaftlichen R) könnten jedoch darauf hinweisen, dass die Möglichkeiten, durch Faktorisierung an gesellschaftliche Relevanzen anzuschliessen, inzwischen ausgereizt sind.

2. *Vermessungskritik*: Messen ist nur möglich, wenn *nicht alles* gemessen wird.¹⁴ Messen ist also notwendig selektiv, woran keine Verfeinerung von Messinstrumenten je etwas ändern könnte. Im Gegenteil: Verfeinerung *ist* Steigerung der Selektivität der Instrumente, Steigerung ihrer Indifferenz gegen Störgrössen. Durch Hinweise auf das Ausgeschlossene kann sich eine *Kritik* der Vermessungen von Bildungsqualität in Gang setzen respektive eine ohnehin bestehende kritische Beobachtung des Bildungssystems verstärken (vgl. BRÜGELMANN 2001; CHOTT 2002; HERRMANN 2003; BRAKE 2003). Die Kritik kann zur Faktorisierung übergehen, wenn sie allein die Relevanz der jeweils aktuell gemessenen Faktoren bezweifelt und bislang ausgeschlossene Grössen zur künftigen Vermessung freilegen will. In diesem Fall kommt es allein zur Verschiebung der Messselektivitäten. Die Kritik kann sich aber auch auf die grundsätzliche Selektivität allen Messens beziehen und in der Orientierung an Quantifikationen eine Fehleinstellung des Bildungssystems konstatieren (so MESSNER 2003). Sie würde sich damit in die lange Tradition der Kritik einer Technisierung und Mathematisierung einreihen, wie sie etwa pointiert und einflussreich von Edmund HUSSERL (1982) formuliert wurde.

¹³ Siehe zu typischen korrekturbedürftigen "Politikerirrtümern" GORARD 1999.

¹⁴ Populärwissenschaftlich ist dieses methodologische Gesetz seit der HEISENBERGSchen Unschärferelation bekannt, die besagt, dass man von elementaren Kleinst"teilchen" nur entweder ihren Lage oder ihre Geschwindigkeit messen kann, nie aber beides zugleich. Die „sozialwissenschaftliche Variante" der HEISENBERGSchen Unschärferelation besagt: Die messtechnische Unbestimmbarkeit steigt mit der sozialen Relevanz des Gemessenen, da höhere Relevanz in Antizipationen von und Reaktionen auf Messungen grössere Veränderungen des zu Messenden mit sich bringt (AMREIN/BERLINER 2002). Beispiel: PISA verändert Unterrichtsstoffe oder Stoffmodelle und misst spätestens in der zweiten Runde bereits eigene Effekte. Zum Nebeneffekt, durch den Ausschlusszwang eher *leicht* messbare Grössen vor schwierigen Grössen zu präferieren, siehe FRÖHLICH 1999.

Hier dürften sich Anschlussmöglichkeiten für die eher philosophischen, geisteswissenschaftlichen, gesellschafts- und kulturkritischen Ansätze der Bildungstheorie bieten. Doch die Frage wäre, ob und wie solch eine Grundsatzkritik auch Folgen für die *Bildungsforschung* zeitigen könnte. Eine nahe liegende Möglichkeit besteht darin, die Kritik in *qualitative* Forschungen umzumünzen und dabei auf die Positionen, Verfahren und Methoden zurückzugreifen, die in der allgemeinen Sozialforschung seit den 1970er-Jahren in einer sehr ähnlichen Diskussionslage entwickelt wurden.¹⁵ Wie die Erfahrungen dort gezeigt haben, kann eine Auseinandersetzung zwischen quantitativer und qualitativer Forschung wechselseitig sehr stimulierend sein, aber auch in die unbefriedigende Situation führen, dass man beginnt, die gegenseitigen „Unverständnisse“ für die Hauptprobleme des Faches und der Theoriebildung zu halten. Vergessen wird dabei, dass das *gemeinsame* Problem in der strukturierten Komplexität des Gegenstandes liegt, die nur auf je unterschiedliche Weise aufgelöst und rekombiniert wird: einmal durch selektives Zählen und starke Generalisierungen der Ergebnisse, das andere Mal durch selektives Interpretieren und starke Spezifikationen in Fallbeschreibungen.

3. *Operationalisierung von Systembeobachtungen*: Eine dritte Möglichkeit bestünde darin, eine Theorie des Systemgedächtnisses für den Fall des Bildungssystems auszuarbeiten und Forschungen auf die Frage auszurichten, ob und wie Änderungen der Gedächtnisstrukturen auch Änderungen in den Formen der *Systemoperationen* nach sich ziehen.¹⁶ Klassischerweise werden Probleme dieser Art mit der Unterscheidung von Theorie und Praxis formuliert; eine Unterscheidung, die in dem beschriebenen Kontext jedoch nicht richtig greift, da die Änderung des Systemgedächtnisses ja gerade darin besteht, von Theorien auf Messungen (und auf Messtheorien) umzustellen. Auch auf der Seite der Praxis kommt es zu Vervielfältigungen; denn die Qualitätsmessungen adressieren ja nicht nur die Praxis des Unterrichts, des Lehrens und des Lernens, sondern auch die Praxen der Schulleitungen, der Schuladministrationen, der Bildungspolitikern und andere mehr. Eine Möglichkeit, hier neu anzusetzen, ohne die Theorie/Praxis-Problematik aus den Augen zu verlieren, bietet die Unterscheidung von Operation und Beobachtung.¹⁷ Systemtheoretisch gefasst, kommen Operationen immer zuerst und sind grundlegend auch insofern, indem die Beobachtungen selbst Operationen sein müssen. Das gilt erst recht für alle Beobachtungen, die als Systemgedächtnis im System wirken sollen. Ein Systemgedächtnis ist nicht allein Erinnerung an vergangene Ereignisse. Man kann darunter vielmehr all jene Beobachtungen verstehen, die aktuelle Ereignisse mit potentiell möglichen Ereignissen *vergleichen* und die Vergleichsergebnisse zur Bestimmung der aktuellen Operationen mit heranziehen. Vergangenheit *und* auch Zukunft bilden aufeinander bezogene Horizonte für die jeweils verfügbaren Vergleichsrahmen. Ein Gedächtnis realisiert daher stets Rückschau und Vorschau

¹⁵ Siehe für einen Überblick über den erreichten Stand in verschiedenen (europäischen) Ländern die Beiträge in der Schweizerischen Zeitschrift für Soziologie 26. Jg. (2000), Nr. 3.

¹⁶ Siehe zu Ansätzen zu einer Theorie des sozialen Gedächtnisses GOODY/WATT 1963; ASSMANN/HARTH 1991; LUHMANN 1996b; FUCHS 1997 sowie ESPOSITO 2002.

¹⁷ Diese Unterscheidung ist bislang vor allem in der soziologischen Systemtheorie von LUHMANN (1984, S. 596ff.; für das Bildungssystem 2002, S. 168ff.) ausgearbeitet worden, an der ich mich hier orientiere, ohne damit behaupten zu wollen, die Systemtheorie sei die einzig mögliche Form, die Differenz von Operation und Beobachtung für die Bildungsforschung fruchtbar zu machen. Ansatzpunkte bieten beispielsweise auch die Entwicklungen zu einer Theorie pädagogischer Kommunikation (u.a. KADE 2003) oder zu einer Theorie pädagogischen Handelns (COMBE/HELSPER 1996; FEND (ohne Jahr)).

zugleich (FOERSTER 1994). Im Bildungssystem übernehmen beispielsweise Schulnoten und Bildungszertifikate diese Funktion (LUHMANN 2002, S. 35). Sie erinnern für sich genommen zwar nicht, was und wie gelernt wurde (erinnern heisst stets auch vergessen), aber immerhin, *dass* gelernt wurde und ungefähr, *wofür* gelernt wurde (zum Beispiel für die Versetzung oder für eine Berufsqualifikation).

Beobachtungen können jedoch dann und nur dann ein System informieren und als dessen Gedächtnis wirken, wenn sie im selben System wie die Operationen durchgeführt werden. Vorab jeder detaillierenden Forschung lässt diese kleine Theorieskizze bereits erkennen, dass die Mehrzahl der Bildungsmessungen von externen Systemen durchgeführt wird, jene Systeme informiert und folglich keinen Anschluss in pädagogischen Kommunikationen findet. Die Wissenschaft baut ihre Instrumente nach eigenen Qualitätskriterien und wertet die Ergebnisse für ihre Theorien und für den Nachweis weiterer Forschungsnotwendigkeiten aus. Die Politik lässt sich vom Finanzbedarf für solche Unternehmungen überzeugen in der Aussicht darauf, dass politische Entscheidungsspielräume durch künftige Sachkenntnisse erweitert und / oder entpolitisiert werden könnten. Die Massenmedien nehmen das von der Wirtschaft, der Kriminalität, vom Wetter, vom Sport, von der Politik (Meinungsumfragen) und von der Wissenschaft her allgemein vertraute Element der Quantifikation gerne auf, um ihre Spiele um Zahlen und Meinungen zu inszenieren (LUHMANN 1996a, S. 62f.). Solche Fremdbeobachtungen bleiben sicherlich nicht ohne Auswirkungen für das Bildungssystem, betreffen jedoch zunächst einmal nur einige seiner Strukturen wie Finanzierungsvolumina, Strukturregelungen oder die „Images“ von Schule und Lehrberuf. Jene hier "Operationalisierung" genannte Strategie könnte ihren Komplex von Forschungsfragen aus dem Grundproblem entwickeln, ob und wenn ja, wie die basalen *Operationen* des Bildungssystems überhaupt erreicht werden.

Die zentralen Operationen des Bildungssystems ereignen sich (bis heute jedenfalls) in der Form des Unterrichts in Schulklassen. Der Unterricht ist – wie immer wieder erinnert wird – weder definitiv noch messtechnisch noch technologisch ganz zu erfassen, geschweige denn zu kontrollieren.¹⁸ Der Unterricht kombiniert Routinen und Zufälle (LUHMANN 2002, S. 109), wobei nicht nur die Lehrerinnen und Lehrer, sondern auch die Schülerinnen und Schüler ihre Routinen und ihre Einfälle beisteuern.¹⁹ Es gibt Routinen des Lehrens und Routinen des Lernens, Routinen der Erziehung und Routinen des Erzogenwerdens, Routinen der Disziplinierung und Routinen der Rettung der Selbstdarstellung. Getragen und ermöglicht durch solche Routinen macht der Zufall des täglichen Kommunikationsgeschehens den Unterricht "spannend", eventuell sogar interessant, für Lehrende wie für Schülerinnen und Schüler. Bildungsqualität im Unterricht kommt zustande, wenn es gelingt, solche Zufälle in Gelegenheiten umzuformen; wenn es gelingt, ein momentanes Zusammenfallen von Unwissen (seitens des Schülers) und Vermittlungsbereitschaft (seitens des Lehrers) in ein Lernen und in ein Lernen des Lernens zu verwandeln. Solche Gelingensmomente sind nicht nur relativ rar (gemes-

¹⁸ Siehe dazu unter anderem die Debatten im Anschluss an die These vom „Technologiedefizit“ der Erziehung in LUHMANN/SCHORR 1982.

¹⁹ Daran ist gerade in Zeiten der Suche nach „guten Methoden“ und „guten Lehrpersonen“ immer wieder ausdrücklich zu erinnern; so auch OELKERS (2004) mit der starken Forderung nach einer "Professionalisierung des Schülerberufs" (S. 12). Wie solch eine Professionalisierung induziert werden könnte, ist offen gelassen. Man muss wohl wieder an Unterricht denken.

sen an den Ansprüchen von Lehrern und Schülern und vielen ihrer Beobachter). Das tiefer liegende Problem ist, dass solche Momente – der Chance, des Ergreifens, des Gelingens und des Verfehlens – nicht abzählbar sind, damit auch nicht aggregierbar und nicht vergleichbar.

Mit dieser Erinnerung an Selbstverständliches ist nicht einer (Re-)Mystifizierung des Unterrichts und der Lehrer-Schüler-Begegnung das Wort geredet. Die hier anstehende Forschungsfrage lautet schlicht, ob und wie ein ausgearbeitetes, mit zahlreichen Datenvergleichen arbeitendes und vorwiegend auf Quantifikationen angewiesenes Systemgedächtnis den Kontakt zu den Grundoperationen der Bildung herstellen kann; und zwar so, dass nicht nur die Grundoperationen beobachtet werden, sondern diese ihrerseits mitbeobachten können, wie sie beobachtet werden. Denn nur so, nur durch Systemschliessung zumindest auf der Ebene des Beobachtens, kann erst erwartbar werden, dass die Qualitätsbeobachtungen die produzierenden Operationen informieren. Abermals wird man nicht mit strikten, sondern nur mit losen Kopplungen rechnen können, was alle technologischen Ansätze in ihre Schranken verweist. Die möglichen Kopplungsformen wären, soweit bislang erkennbar, vor allem in den Bereichen der LehrerInnenausbildung (Professionalisierung) sowie der Schulentwicklung auszumachen (siehe bspw. ROLFF 2002). Ob sich auf diesen, in sich selbst verschlungenen Wegen das veränderte Systemgedächtnis tatsächlich operationalisiert („die Praxis erreicht“), oder ob sich das System, nach einigen „Erregungsfiebern“, nicht einfach gegen die externen Beobachtungen mehr oder weniger „immunisiert“, wären Fragen, die nur empirisch zu beantworten sind.

Literatur

- AMREIN, A. L./BERLINER, D. C. (2002): High-Stakes Testing, Uncertainty, and Student Learning. In: Education Policy Analysis Archives, 10. Jg., H. 18 (ohne Paginierung, online unter <http://epaa.asu.edu/epaa/v10n18/>).
- ASSMANN, A./HARTH, D. (Hrsg.) (1991): Mnemosyne. Frankfurt am Main.
- BARBER, M. (2003): From Good to Great: Large-Scale Education Reform in England. In: OELKERS, J. (ed.): Futures of Education II, S. 293-316.
- BLAUG, M. (1997): Ugly Currents in Modern Economics. In: Options Politiques, 18. Jg., H. 17, S. 3-8.
- BOS, W./POSTLETHWAITE, T. N. (2001): Internationale Schulleistungsforschung. In: WEINERT, F. E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim, S. 251-267.
- BRAKE, A. (2003): Worüber sprechen wir, wenn von PISA die Rede ist? In: Zeitschrift für Soziologie der Erziehung und Sozialisation, 23. Jg., H. 1, S. 24-39.
- BRÜDERLE, R. (2004): Politik statt Zahlenreihen. In: Schweizer Monatshefte, 84. Jg., H. 9/10, S. 18-20.
- BRÜGELMANN, H. (2001): Kontroversen um die Schulleistungsmessung in Deutschland. In: WEINERT, F. E. (Hrsg.): Leistungsmessungen in Schulen. . Weinheim, S. 33-44.
- CHOTT, P. O. (2002): Betrifft: PISA. Analyse der Schülerleistungsstudie 2000. In: SchulVerwaltung. Ausgabe Bayern, 25. Jg., H. 3, S. 84-90.
- COLEMAN U. A. 1966 = COLEMAN, J. S./CAMPBELL, E. Q./HOBSON, C. J./MCPARTLAND, J./MOOD, A. M./WEINFIELD, F. D./YORK, R. L. (1966): Equality of Educational Opportunity. Washington.
- COMBE, A./HELSPER, W. (Hrsg.) (1996): Pädagogische Professionalität: Untersuchungen zum Typus pädagogischen Handelns. Frankfurt am Main.
- ESPOSITO, E. (2002): Soziales Vergessen. Frankfurt am Main.
- FEND, H. (1996): Schulkultur und Schulqualität. In: LESCHINSKY, A. (Hrsg.): Die Institutionalisierung von Lehren und Lernen. Weinheim-Basel, S. 85-97.
- FEND, H. (1998): Qualität im Bildungswesen. Weinheim.
- FEND, H. (ohne Jahr): Neue Theorie der Schule. Unv. Ms. Pädagogisches Institut der Universität Zürich.
- FOERSTER, H. V. (1994): Was ist Gedächtnis, daß es Rückschau *und* Vorschau ermöglicht? In: DERS.: Wissen und Gewissen. Frankfurt am Main, S. 299-336.
- FRIED, J. (Hrsg.) (1986): Schulen und Studium im sozialen Wandel des hohen und späten Mittelalters. Sigmaringen.
- FRÖHLICH, G. (1999): Das Messen des leicht Meßbaren. Output-Indikatoren, Impact-Maße: Artefakte der Szientometrie? In: BECKER, J./GÖHRING, W. (Hrsg.): Kommunikation statt Markt. Sankt Augustin, S. 27-38.
- FREY, B. S. (2000): Was bewirkt die Volkswirtschaftslehre? In: Perspektiven der Wirtschaftspolitik, 1. Jg., H. 1, S. 5-33.
- FUCHS, P. (1997): Wie lernen autopoietische Systeme? In: Soziale Wirklichkeit, 1. Jg., H. 2, S. 119-134.
- GOODY, J./WATT, I. (1963): The Consequences of Literacy. In: GOODY, J. (ed.): Literacy in Traditional Societies. Cambridge, S. 27-68.
- GORARD, S. (1999): Keeping A Sense of Proportion: The 'Politician's Error' in Analyzing School Out-

- comes. In: *British Journal of Educational Studies*, 47. Jg., H. 3, S. 235-246.
- GRUNDER, H. U. (1996): Die gute Lehrerin, der gute Lehrer. In: *Beiträge zur Lehrerbildung*, 14. Jg., H. 1, S. 53-60.
- HERRMANN, U. (2003): "Bildungsstandards". In: *Zeitschrift für Pädagogik*, 49. Jg., H. 5, S. 625-639.
- HUSSERL, E. (1982): *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie*. Hamburg.
- JENCKS U. A. 1972 = JENCKS, C./SMITH, M./ACKLAND, H./BANE, M. J./COHEN, D./GINTIS, H./HEYNS, B./MICHELSON, S. (1972): *Inequality: A Reassessment of the Effects of Family and Schooling in America*. New York.
- KADE, J. (2003): Erziehung als pädagogische Kommunikation. In: LENZEN, D. (Hrsg.): *Irritationen des Erziehungssystems*. Frankfurt am Main, S. 198-231.
- KEMPFERT, G./ROLFF, H.-G. (2000): *Pädagogische Qualitätsentwicklung*. Weinheim-Basel.
- LAWSON, T. (2003): *Reorienting Economics*. London/New York.
- LUHMANN, N. (1984): *Soziale Systeme*. Frankfurt am Main.
- LUHMANN, N. (1996a): *Die Realität der Massenmedien*. Opladen.
- LUHMANN, N. (1996b): Zeit und Gedächtnis. In: *Soziale Systeme*, 2. Jg., H. 2, S. 307-330.
- LUHMANN, N. (2000): Statistische Depressionen. Zahlen in den Massenmedien. In: LUHMANN, N.: *Short Cuts*. Frankfurt am Main, S. 107-112.
- LUHMANN, N. (2002): *Das Erziehungssystem der Gesellschaft*. Frankfurt am Main.
- LUHMANN, N./SCHORR, K. E. (1982): Das Technologiedefizit der Erziehung und die Pädagogik. In: LUHMANN, N./SCHORR, K. E. (Hrsg.): *Zwischen Technologie und Selbstreferenz*. Frankfurt am Main, S. 11-40.
- MADAUS, G. F./AIRASIAN, P. W./KELLAGHAN, T. (1980): *School Effectiveness*. New York.
- MAYER, D. P./MULLENS, J. E./MOORE, M. T. (2000): *Monitoring School Quality: An Indicators Report*. Washington, DC.
- MEAD, G. H. (1938): Passage, Process, and Permanence. In: MEAD, G. H.: *The Philosophy of the Act*. Chicago, S. 321-356.
- MERKENS, H. (2003): PISA – Erfolg als Kombination von Vermarktung und wissenschaftlicher Redlichkeit. In: *Soziologische Revue*, 26. Jg., H. 2, S. 183-194.
- MESSNER, R. (2003): PISA und Allgemeinbildung. In: *Zeitschrift für Pädagogik*, 49. Jg., H. 3, S. 400-412.
- MOELLER, B./PATZE, H./STACKELMANN, K. (Hrsg.) (1983): *Studien zum städtischen Bildungswesen des späten Mittelalters und der frühen Neuzeit*. Göttingen.
- MOSTELLER, F./MOYNIHAN, D. P. (eds.) (1972): *On Equality of Educational Opportunity*. New York.
- NUSSBAUM, P. (2001): *Bildungsindikatorensystem Kanton Zürich*. Zürich.
- OELKERS, J. (2002): *Steuerung durch Forschung*. Ms. Zürich.
- OELKERS, J. (2004): *Bildungsforschung, Schulentwicklung und Weiterbildung*. Ms. Zürich
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD) (2000): *Education at a Glance: OECD Indicators*. Paris.
- PUTNAM, L. L. (1986): Contradictions and Paradoxes in Organizations. In: THAYER, L. (ed.): *Organization - Communication*. Norwood, N.J., S. 151-167.
- ROLFF, H.-G. (2002): Was bringt die vergleichende Leistungsmessung für die pädagogische Arbeit in

- Schulen? In: WEINERT, F. E. (Hrsg.): Leistungsmessungen in Schulen. Weinheim, S. 337-352.
- RÜRUP, M. (2003): Ausländische und internationale Bildungsberichte als Orientierung für die nationale Bildungsberichterstattung in Deutschland. In: AVENARIUS, H./ DITTON, H./DÖBERT, H./KLEMM, K./KLIEME, E./RÜRUP, M./TENORTH, H.-E./WEISHAUPT, H./WEIß, M.: Bildungsberichterstattung für Deutschland: Konzeption. Frankfurt am Main, S. 9-34.
- SCHÖNIG, W. (2002): Organisationskultur der Schule als Schlüsselkonzept der Schulentwicklung. In: Zeitschrift für Pädagogik, 48. Jg., H. 6, S. 815-834.
- STAMM, M. (2003): Konzept "Bildungsmonitoring Schweiz". Aarau.
- STONE, D. A. (1997): Policy Paradox. New York.
- STRIVERS, J. (1985): School Climate. In: REYNOLDS, D. (ed.): Studying School Effectiveness. London, S. 45-57.
- TILLMANN, K.-J. (2001): Leistungsvergleichsstudien und Qualitätsentwicklung. In: Neue Sammlung, 41. Jg., H. 3, S. 359-368.
- VANBERG, V. J. (2004): Mathematikmanie und die Krise der Ökonomik. In: Schweizer Monatshefte, 84. Jg., H. 9/10, S. 21-24.
- WEICK, K. E. (1976): Educational Organizations as Loosely Coupled Systems. In: Administrative Science Quarterly, 21. Jg., S. 1-19.
- WEINERT, F. E. (Hrsg.) (2002): Leistungsmessungen in Schulen. Weinheim.
- WEISS, J./ZUTAVERN, M. (2003): PISA-ergänzende Perspektiven. In: Schweizerische Zeitschrift für Bildungswissenschaften, 25. Jg., H. 1, S. 5-10.
- WYATT, T. (1996): School Effectiveness Research. In: Issues in Educational Research, 6. Jg., H. 1, S. 79-112.

Die Reihe „Materialien zur Bildungsforschung“ der Pädagogischen Hochschule Thurgau wird fortgesetzt. Bisher sind erschienen:

- Nr. 1 Trachsler, E. (2004). *Konsequenter Umbau der Schulaufsicht in der Schweiz. Schulautonomie und Qualitätssteuerung am Beispiel der Kantone Thurgau, Zürich und Aargau.*
- Nr. 2 Trachsler, E. (2004). *Der Einfluss von Schulleitungen auf das Weiterlernen von Lehrpersonen. Vier Koordinaten auf dem Weg zur Frage: Inwiefern gelingt es den Schulleitungen, das Weiterlernen der Lehrpersonen zu fördern?*
- Nr. 3 Brosziewski, A. (2004). *Die Pädagogische Hochschule Thurgau als Adresse (in) der Bildungspolitik.*
- Nr. 4 Trachsler, E. (2004). *Grundsätzliche Überlegungen zur Schülerbeurteilung und zur Schaffung neuer Zeugnisse – eine konstruktive Kritik.*
- Nr. 5 Brosziewski, A. (2005). *Bildungsqualität, statistische Depressionen und das Gedächtnis des Bildungssystems.*
- Nr. 6 Trachsler, E. & Brosziewski, A. (2005). *Schulentwicklung zwischen Konzepten und Konkretisierungen.* Plenumsreferat anlässlich der Akademie 2005 EDK-Ost vom 27. Mai 2005 in der Kartause Ittingen (erweiterte Version).

