

Bellmann, Johannes

Bildungsforschung und Bildungspolitik im Zeitalter 'Neuer Steuerung'

Zeitschrift für Pädagogik 52 (2006) 4, S. 487-504

urn:nbn:de:0111-opus-44682

in Kooperation mit / in cooperation with:

BELTZ

<http://www.beltz.de>

Nutzungsbedingungen / conditions of use

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.

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.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.
By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS
Deutsches Institut für Internationale Pädagogische Forschung (DIPF)
Mitglied der Leibniz-Gemeinschaft
Informationszentrum (IZ) Bildung
Schloßstr. 29, D-60486 Frankfurt am Main
E-Mail: pedocs@dipf.de
Internet: www.pedocs.de

Inhaltsverzeichnis

Essay

Johannes Bellmann

Bildungsforschung und Bildungspolitik im Zeitalter ‚Neuer Steuerung‘ 487

Thementeil: Pädagogik und die Formierung nationaler Kulturen

Daniel Tröhler

Pädagogik und die Formierung nationaler Kulturen.
Einführung in den Thementeil 505

Rita Casale

Die italienische Erziehung des bourgeois gentilhomme 508

Fritz Osterwalder

Condillacs Rose – Erfahrung als pädagogisches Konzept zwischen
radikalem Sensualismus und religiöser Innerlichkeit 522

Daniel Tröhler

Lehrerbildung, Nation und pädagogische Historiographie.
Die ‚Geschichten der Pädagogik‘ in Frankreich und Deutschland nach 1871 540

Allgemeiner Teil

Dominik Gyseler

Problemfall Neuropädagogik 555

Gunther Graßhoff/Davina Höblich/Bernhard Stelmaszyk/Heiner Ullrich

Klassenlehrer-Schüler-Beziehungen als biografische Passungsverhältnisse.
Fallstudien zum Verhältnis von Lehrer-Schüler-Interaktionen und
Selbstverständnis der Lehrerschaft an Waldorfschulen 571

<i>Stefan Weyers</i> Pacta sunt servanda? Das kindliche Verständnis von Verträgen am Beispiel des Tausches und der Leihe	591
--	-----

Besprechungen

<i>Wolfgang Beywl</i> Harm Kuper: Evaluation im Bildungssystem. Eine Einführung	611
--	-----

Dokumentation

Pädagogische Neuerscheinungen	614
-------------------------------------	-----

Johannes Bellmann

Bildungsforschung und Bildungspolitik im Zeitalter ‚Neuer Steuerung‘¹

1. Die Kontroverse um Bildungspolitik und Wissenschaftspolitik im Zeitalter Neuer Steuerung

Das Verhältnis von Bildungsforschung und Bildungspolitik im Zeitalter Neuer Steuerung steht im Zentrum aktueller Kontroversen in der Erziehungswissenschaft. Genau genommen handelt es sich um zwei in sich komplexe und miteinander verknüpfte Kontroversen: In der *bildungspolitischen* Kontroverse geht es um neue Steuerungsinstrumente im Bildungssystem, insbesondere Output- und Wettbewerbssteuerung. In der *wissenschaftspolitischen* Kontroverse geht es darum, wie sich Bildungsforschung und Erziehungswissenschaft im Kontext Neuer Steuerung positionieren und entwickeln.

Zunächst zur *wissenschaftspolitischen* Seite: Viele erinnern sich an die zugespitzten Thesen Hans Weilers auf dem Münchener Kongress der DGfE. „Die Erziehungswissenschaft kommt ihrer Pflicht, bildungspolitische Entscheidungen empirisch zu erhellen und kritisch zu begleiten, in einem kaum angemessenen Umfang nach“ (Weiler 2003, S. 184). Weiler zufolge liegt dem von ihm diagnostizierten Defizit eine „tiefe Entfremdung zwischen den zwei Kulturen von Erziehungswissenschaft und Bildungspolitik zugrunde, die nicht auf Deutschland beschränkt ist, die aber hierzulande besonders stark ausgeprägt zu sein scheint“ (ebd.).

Was Hans Weiler etwas pauschal über „die Erziehungswissenschaft“ sagt, gilt gegenwärtig nicht für die empirische Bildungsforschung. Ihr Verhältnis zur Politik ist nicht von „tiefer Entfremdung“ geprägt, sondern von einem lebhaften wechselseitigen Interesse. In der „Stellungnahme der DFG zur strukturellen Stärkung der empirischen Bildungsforschung“ wird bemerkt, dass die Expertise empirischer Bildungsforschung „im Gefolge der internationalen Schulleistungsstudien zunehmend nachgefragt“ werde, was als Ausweis dafür gesehen wird, dass sich die „Bildungsforschung nach langen Jahren der Abstinenz wieder als ernst genommener Partner der Bildungspolitik etabliert hat“ (DFG 2001, S. 2).

Wie kam es zu dieser neuen Annäherung? Auffällig ist, dass der Wandel der bildungspolitischen Steuerungsinstrumente als ein möglicher Hintergrund für die gesteigerte Nachfrage nach empirischer Bildungsforschung nicht erwähnt wird. Wenn man aber die neue Nachfrage nach empirischer Bildungsforschung in diesem steuerungspolitischen Paradigmenwechsel (vgl. Döbert 2003, S. 301) verortet, so spricht vieles dafür, dass wir gegenwärtig nicht nur eine vorübergehende Konjunktur im Verhältnis von Bildungspolitik und Bildungsforschung erleben. Im Zeitalter Neuer Steuerung entsteht

1 Der Text ist die leicht überarbeitete Fassung meines am 8. Februar 2006 an der Humboldt-Universität zu Berlin gehaltenen Habilitationsvortrags.

vielmehr ein kontinuierlicher Bedarf an Bildungsforschung, insbesondere in Gestalt von Systemmonitoring und Programmevaluation. Die gegenwärtig nachgefragte wissenschaftliche Expertise unterscheidet sich dabei von wissenschaftlicher Politikberatung alter Manier dadurch, dass sie politische Steuerung nicht nur informiert und begleitet, sondern selbst schon ein konstitutives Moment Neuer Steuerung darstellt.

Das führt zur *bildungspolitischen* Seite, nämlich der Kontroverse um Neue Steuerung. Hier zeigt sich eine zunehmend polarisierte Debatte, in der es immer schwerer wird, eine gemeinsame Konsensbasis auszumachen. Schon die nahe liegende Einschätzung, dass die generelle Frage nach Steuerung und Steuerbarkeit der Pädagogik und Erziehungswissenschaft nicht fremd, sondern in ihrer grundsätzlichen Veränderungsperspektive durchaus verankert ist (vgl. Roth/Friedrich 1975, S. 30f.), scheint gegenwärtig nicht mehr konsensfähig. Freilich wird Neue Steuerung heute vielfach als funktionssystemübergreifend und international anwendbares Erfolgsmodell angepriesen, worüber die Einsicht in historische und teilsystemspezifische Grenzen der Übertragbarkeit entsprechender Steuerungsmodelle in den Hintergrund tritt. Wo diese Einsichten in Recht und Grenzen Neuer Steuerung sich auflösen, bleiben unversöhnliche Positionen zurück.

Auf beiden Seiten dieser polarisierten Debatte zeigt sich zudem eine Art Dramatisierung des Augenblicks, als gelte es jetzt eine Grundsatzentscheidung über Wohl und Wehe des Bildungssystems herbeizuführen. Jetzt sei das Zeitfenster offen, durch nachhaltige und irreversible Reformen endlich internationale Entwicklungen nachzuvollziehen, erklären die entschlossenen Modernisierer (vgl. Oelkers 2005a, 2005b). Umgekehrt heißt es nicht minder dramatisch, Erziehungswissenschaft und Pädagogik stünden an einem Scheidepunkt. Was bevorstehe sei die „Verwandlung der Pädagogik in angewandte, empirische Evaluationswissenschaft, [...] die Transformation von Schule, Universität und Wissenschaft in Dienstleistungsbetriebe“ (Koch 2004, S. 55). Man müsse seine warnende Stimme erheben, bevor es zu spät sei.

Auffällig ist hierbei, dass es im Streit um Neue Steuerung nicht nur um eine Reform der Bildungspraxis, sondern zugleich um eine Reform der Erziehungswissenschaft zu gehen scheint. Bildungsreform und Wissenschaftsreform werden von beiden Seiten als zwei eng miteinander koordinierte Projekte angesehen, die sich wechselseitig voraussetzen. Beschreibt man beide Positionen im Vokabular des jeweiligen Gegners, so ergibt sich ein finsternes Bild: Auf der einen Seite stehen technokratische Modernisierer, die das Bildungswesen einer neoliberalen Ökonomisierung ausliefern. Ihnen stehen reaktionäre Anhänger eines deutschen Bildungsdenkens gegenüber, die an einem ineffizienten und ungerechten System festhalten, um eigene Privilegien zu sichern. So neigt man mitunter eher zur Diffamierung der Gegenseite, statt deren Anliegen ernsthaft aufzugreifen. Sollten beide Positionen ein zutreffendes Portrait der jeweiligen Gegenseite zeichnen, so wüsste man nicht, welche Alternative das kleinere Übel darstellt. Woody Allen hat eine solche Situation einmal so beschrieben: „More than any other time in history, mankind faces a crossroads. One path leads to despair and utter hopelessness. The other, to total extinction. Let us pray we have the wisdom to choose correctly“² (Allen 1981, S. 81).

2 Das Zitat verdanke ich Denis C. Phillips (vgl. Phillips 2005).

Im Folgenden möchte ich mich nicht weiter auf die polarisierte deutsche Debatte beziehen. Die konkreten Beispiele, an denen ich das veränderte Verhältnis von Bildungspolitik und Bildungsforschung untersuchen will, stammen aus den USA. Im Einzelnen geht es um zwei Gesetzesinitiativen der Bush-Regierung. Die eine zielt auf eine Reform in der Bildungspraxis, die andere zielt auf eine genau darauf abgestimmte Reform erziehungswissenschaftlicher Forschung.

Das Beispiel USA ist aus meiner Sicht aus verschiedenen Gründen geeignet, um die veränderte Konstellation von Bildungsforschung und Bildungspolitik im Kontext Neuer Steuerung zu studieren. Zunächst spricht der Erfahrungsvorsprung für das Beispiel USA, weil dort bereits über einen längeren Zeitraum neue Steuerungsinstrumente im Bildungssystem erprobt werden. Auch hinsichtlich nicht-intendierter Nebenfolgen und perverser Effekte scheint man einen Erfahrungsvorsprung zu haben, was wiederum die kritische Evaluation der Reformen selbst weiter vorangetrieben hat als andernorts. Das Beispiel USA ist darüber hinaus auch deshalb interessant, weil Einzelaspekte Neuer Steuerung, die hierzulande noch recht unverbunden nebeneinander stehen, dort durch Gesetzespakete der Bundesregierung systematisch miteinander verzahnt wurden. Ich nutze die Beispiele nicht, um ihre Übertragbarkeit auf die deutsche Situation zu suggerieren, sondern um einige Problemzonen im gegenwärtigen Verhältnis von Bildungsforschung und Bildungspolitik zu identifizieren, die eine über die untersuchten Fälle hinausgehende Bedeutung haben könnten. Doch bevor ich zu meinen Beispielen komme, will ich sehr holzschnittartig voranschicken, was ich unter Neuer Steuerung verstehe.

2. Neue Steuerung im Bildungssystem

Das neue Steuerungsmodell im Bildungssystem beruht im Kern auf zwei Steuerungsinstrumenten, die inzwischen eine große internationale Verbreitung gefunden haben. Auf der einen Seite spricht man von „standards-based reform“, womit gegenwärtig eine Form von *Outputsteuerung* durch die Setzung von Bildungsstandards und die externe Evaluation von Schulleistungen gemeint ist. Auf der anderen Seite spricht man von *Wettbewerbssteuerung* oder sog. „choice policies“, worunter die Etablierung von Quasimärkten im Bildungssystem durch Dezentralisierung, Schulautonomie und freie Schulwahl verstanden wird, ggf. verstärkt durch die Umstellung von der Angebots- zu einer Nachfragefinanzierung. Während die Bildungspolitik sich bislang vor allem auf zentrale Steuerung durch Inputs wie ökonomische Ressourcen, detaillierte Lehrpläne und die Professionalisierung der Lehrerschaft konzentrierte, rückt im neuen Steuerungsmodell „governing by goals and results“ in den Mittelpunkt (vgl. Lundgren 2001).

Diese Beschreibung stellt sicherlich eine grobe Vereinfachung dar, denn innerhalb dieses allgemeinen Trends hin zu Output- und Wettbewerbssteuerung gibt es bedeutende Differenzen. Testkulturen und Rückmeldesysteme unterscheiden sich z.B. darin, ob sie Evaluationsergebnisse mit hohen Sanktionen oder Gratifikationen verbinden (*high-stakes testing*) oder nicht, ob sie Leistungsmessungen mit Selektionen und zentralen Prüfungen verbinden oder nicht, ob sie eher an der Evaluation von Programmen

und Maßnahmen oder von Personen und Institutionen orientiert sind. Outputsteuerung ist folglich nicht gleich Outputsteuerung. Dasselbe gilt für Wettbewerbssteuerung. Entscheidend ist hier beispielsweise, ob und inwiefern nicht einfach Wettbewerb induziert wird, sondern welche flankierenden Maßnahmen hinzukommen, die sicherstellen, dass beim Wahlverhalten der Anbieter- und der Nachfrageseite tatsächlich Qualitätsgesichtspunkte im Mittelpunkt stehen und ohnehin benachteiligte Gruppen durch das neue Steuerungssystem nicht schlechter gestellt werden.

Erst allmählich wird die *systematische* Verknüpfung beider Steuerungsinstrumente deutlich, die sich im Zuge ihrer Etablierung zunehmend wechselseitig voraussetzen, ergänzen und verstärken. So werden im Zuge von Outputsteuerung zuallererst Unterschiede in der Leistung von Schulen sichtbar, die vorher, zumindest in objektivierter Form, nicht sichtbar waren; systematische Evaluationen schaffen zuallererst die Voraussetzungen dafür, dass Eltern eine begründete Wahl zwischen Anbietern treffen können. Zugleich kann Outputsteuerung aber auch die (legitime) Erwartung wecken, tatsächlich von den besseren Angeboten Gebrauch machen zu können.

Die umgekehrte Lesart des Wechselverhältnisses lautet folgendermaßen: Wenn Wettbewerbssteuerung nicht zu Wildwuchs führen soll, muss durch Standards und regelmäßige externe Evaluationen eine gewisse Vergleichbarkeit der Anbieter im Rahmen objektivierbarer Erwartungen gewährleistet werden. Zugleich soll der Wettbewerb im Zusammenspiel kontinuierlicher Leistungsvergleiche selbst schon als Instrument der Qualitätsentwicklung im Bildungssystem fungieren.

Historisch gesehen gehören beide Steuerungsinstrumente nicht von Anfang zusammen. Für die USA sehen viele Beobachter eine Wende in der zweiten Hälfte der 1990er Jahre. Die Wende vollzog sich in drei Schritten, die auf eine zunehmend enge Verzahnung beider Steuerungsinstrumente hinausliefen: Zunächst wurden Testergebnisse nicht mehr allein an Schulen zum Zwecke interner Qualitätsentwicklung rückgemeldet, sondern zunehmend einer breiteren Öffentlichkeit zur Verfügung gestellt. Der zweite Schritt war die zunehmende Verknüpfung von Testergebnissen mit Belohnungen und Sanktionen. Der dritte Schritt schließlich war die verstärkte Förderung freier Schulwahl als zusätzliches Anreizsystem für Schulen (vgl. Betebenner/Howe/Foster 2005). Damit bin ich bei meinem ersten Reformbeispiel angelangt.

3. Bildungspolitik: Das Gesetzespaket „No Child Left Behind“

Anfang 2002 unterzeichnete George W. Bush ein umfangreiches Gesetzespaket mit dem Titel „No Child Left Behind“. Bei der politischen Begründung für diese Reforminitiative stand nicht das durchschnittliche Abschneiden der USA bei internationalen Leistungsvergleichsstudien im Vordergrund. PISA hat in der amerikanischen Öffentlichkeit und sogar in den Erziehungswissenschaften deutlich weniger Beachtung gefunden als in Deutschland. Längst vor PISA finden sich zahlreiche, z.T. dramatisch zugespitzte Krisendiagnosen im Blick auf den Zustand des US-amerikanischen Bildungssystems (vgl. Cizek 1999). Beachtung in Öffentlichkeit, Politik und Bildungsforschung finden vor al-

lem die ungleiche Ausstattung von Schulen in reichen und armen Schuldistrikten, die hohe Leistungsstreuung zwischen Schulen sowie die hohe Abhängigkeit der Schulleistungen von sozialer Herkunft und ethnischer Zugehörigkeit. Diesen immer wieder beklagten und seit langem bekannten Defiziten stehen hohe Erwartungen an Chancengleichheit im Bildungssystem gegenüber, so dass die Notwendigkeit von Reformen zu meist eher unter Berufung auf Gerechtigkeitsgesichtspunkte als unter Berufung auf Effizienzgesichtspunkte begründet wird.

Dies zeigt auch der Titel der nun unternommenen Reforminitiative „No Child Left Behind“ (NCLB). Das übergeordnete Ziel der Initiative lautet „Promoting educational excellence for all Americans“. Auf der Homepage des U.S. Department of Education wird genannt, was im Einzelnen erreicht werden soll: eine stärkere Rechenschaftspflicht und Verantwortung der einzelnen Schulen und Schuldistrikte für die erzielten Schulleistungen, mehr Handlungsspielraum auf lokaler Ebene, eine verstärkte Orientierung an wissenschaftlich erprobten Unterrichtsmethoden und schließlich größere Möglichkeiten freier Schulwahl.³

Um Schulen und Schuldistrikte für die erzielten Leistungen stärker in die Verantwortung zu nehmen, werden die Bundesstaaten verpflichtet, Bildungsstandards zu setzen und ab dem Schuljahr 2005/2006 jährlich für alle Schüler Tests in Mathematik und Lesen der Jahrgänge 3 bis 8 durchzuführen. Ab 2007/2008 kommen verpflichtende Leistungserhebungen in Naturwissenschaften hinzu. Bis 2014 sollen in allen Bundesstaaten die jeweiligen Mindeststandards erreicht werden, auch von üblicherweise benachteiligten Gruppen, deren Leistungen zusätzlich gesondert erfasst werden müssen. Schulen müssen einen jährlichen Fortschritt (AYP – adequate yearly progress) auf dem Weg zur Erfüllung der Mindeststandards nachweisen und über ihre Ergebnisse in sog. „report cards“ öffentlich Bericht erstatten.⁴ Gelingt es Schulen nicht, entsprechende Fortschritte nachzuweisen, erwartet sie ein Maßnahmenbündel von Unterstützungen und Sanktionen. Sobald eine Schule zweimal hintereinander ihr Jahressoll nicht erfüllt, muss den Eltern zum neuen Schuljahr eine Ausstiegsoption eröffnet werden, das heißt der Wechsel auf eine frei gewählte andere öffentliche Schule oder auf eine *charter school* des Schuldistrikts. In einzelnen Bundesstaaten (z.B. Florida) werden solche Ausstiegsoptionen durch Bildungsgutscheine unterstützt. Mit seiner Fokussierung auf Outputsteuerung und die Stärkung von Wettbewerb verknüpft NCLB systematisch die bedeuten zentralen Steuerungsinstrumente der gegenwärtigen Reformära.

3 Die amtliche Präsentation des Gesetzespakets findet sich unter www.ed.gov/nclb/. Generell gilt: Im föderalen Bildungssystem der USA kann die Bundesregierung nur indirekt auf die Bildungspolitik der Bundesstaaten Einfluss nehmen, vor allem indem sie, wie bei NCLB, die Vergabe föderaler Mittel an bestimmte Kriterien bindet. Da diese Mittel 7-10 % des öffentlichen Bildungsbudgets der Bundesstaaten ausmachen, bestand ein hoher Anreiz – wenn nicht gar ein Zwang – den von der Bundesregierung definierten ‚Vertragsbedingungen‘ zuzustimmen (vgl. McColl 2005).

4 Vgl. als Beispiel Washington State, <http://reportcard.ospi.k12.wa.us/>.

3.1 Zwischenbilanz der Reform

Erste Zwischenbilanzen der Reform (vgl. CEP 2005, NWEA 2005) zeigen insgesamt ein gemischtes Ergebnis: Trotz steigender Testleistungen vermelden Schuldistrikte und Bundesstaaten Schwierigkeiten bei der Implementierung des Gesetzes und der fristgemäßen Erfüllung der Anforderungen, woraufhin die Bundesregierung begonnen hat, einzelnen Staaten Sonderkonditionen wie beispielsweise längere Fristen zur Erfüllung der gesetzten Ziele einzuräumen. Die Komplexität einer Gesetzesinitiative wie NCLB und die zeitgleich dazu stattfindenden von dieser Initiative unabhängigen Reformbemühungen der Einzelstaaten machen eine Gesamtbewertung der Effekte äußerst schwierig. Auch wenn aus Sicht der Bildungsforschung eine solche Bilanz von NCLB fast unmöglich erscheint, hat die Bildungspolitik doch ein Interesse daran, den Erfolg ihrer Gesetzesinitiative unter Beweis zu stellen. Ich möchte nur einen Aspekt in den Mittelpunkt stellen, nämlich die Frage, ob Erfolg und Misserfolg der neuen Steuerungsinstrumente eindeutig festgestellt werden können.

Im April 2005 erschien eine große Studie, die die Testergebnisse von 320.000 Schülern in 23 Bundesstaaten in den Jahren vor und nach NCLB erfasst (vgl. NWEA 2005). Demnach gibt es zwar leichte Zugewinne der Testergebnisse von Klassen (*cross-sectional*), aber innerhalb der Klassen (*longitudinal*) machen Schüler geringere Fortschritte als vor In-Kraft-Treten des Gesetzes. Darüber hinaus zeigt sich ein Unterschied zwischen dem Fortschrittstempo weißer und nicht-weißer Schüler. Zwar zeigen sich auch bei den meisten nicht-weißen Schülern Fortschritte, aber geringere als bei weißen Schülern, was – entgegen dem erklärten Ziel des Gesetzespakets – die Leistungsdifferenzen weiter vergrößern würde.⁵

Ungeachtet dieser erwartungswidrigen Ergebnisse zieht man in Washington eine positive Zwischenbilanz. Bildungsministerin Margaret Spellings sagte im April 2005 vor einer nationalen Versammlung von Schulleitern: „Already we can see the law is making a difference. Across the country, test scores are rising, and the achievement gap is beginning to close“ (Spellings 2005).

Die Erfolgsbilanz des Gesetzespakets wird also offensichtlich sehr unterschiedlich eingeschätzt. Steigen die gemessenen Testleistungen, so deuten Bildungspolitik, Bildungsforschung und Bildungspraxis solche Zugewinne nach systemspezifischen Kriterien.

3.2 Systemspezifische Deutungen von Testergebnissen

Die *Politik* deutet Zugewinne zunächst einmal als Erfolg, den sie den eigenen Reforminitiativen zurechnet. „Die Reformen greifen.“ Damit sind wichtige politische Ziele bereits erreicht, nämlich Handlungsfähigkeit unter Beweis zu stellen und öffentlich Erfol-

5 Erarbeitet wurde die Studie von der „Northwest Evaluation Association“, einer großen non-profit Evaluationsagentur (vgl. Winter 2005).

ge vorweisen zu können. Einzelne Bundesstaaten sind hier vorangegangen. So konnte Texas nach der Einführung eines landesweiten sanktionsbewehrten Evaluationssystems in kurzer Zeit dramatische Anstiege von Testleistungen vermelden, die zu einer politischen Erfolgsgeschichte umgemünzt wurden. Der Versuch, das von der Bildungsforschung inzwischen gründlich untersuchte „Texas miracle in education“ (Haney 2000) gewissermaßen auf nationaler Ebene zu wiederholen, war ein wichtiges Motiv der Bush-Regierung bei der Erarbeitung von NCLB (vgl. Nicholas/Glass/Berliner 2006, S. 4).

In der *Bildungspraxis* wird man Zugewinne ebenfalls begrüßen, aber versuchen, sie mit der Wirksamkeit der eigenen pädagogischen Technologien in Verbindung zu bringen. „Der Unterricht zeigt Wirkung“. Das äußerst schwierige Problem, ob und inwieweit steigende Testleistungen tatsächlich dem Handeln der Akteure zuzurechnen sind, muss dabei nicht bis ins letzte geklärt werden, zumindest nicht, solange die Ergebnisse zufrieden stellen. Grundsätzlich definiert Outputsteuerung für Schulen neue Spielregeln der Beschaffung von Legitimation, die – wie andernorts auch – nicht mehr pauschal, sondern leistungsbezogen zugeteilt wird. Unabhängig von tatsächlichen Qualitätsgewinnen kann schon das Akzeptieren und die Umsetzung der neuen Spielregeln für eine Bildungsinstitution vorteilhaft sein.

Aus Sicht der *Bildungsforschung* dagegen werden gemessene Zugewinne sehr häufig zum Anlass weiterführender Forschungsfragen. Was bedeuten diese Zugewinne? Sind sie tatsächlich Ausdruck einer gesteigerten pädagogischen Produktivität und besserer Schulleistungen? Daniel Koretz (2003, 2005) von der *Harvard Graduate School of Education* beschäftigt sich seit längerer Zeit mit dem bekannten Phänomen, dass Testergebnisse nach Einführung eines neuen sanktionsbewehrten Evaluationsverfahrens zunächst abfallen, dann deutlich ansteigen, um sich nach wenigen Jahren auf einem ähnlich hohen Niveau zu stabilisieren, auf dem auch das vorher zum Einsatz gekommene Evaluationsverfahren angelangt war. Hinter diesem Phänomen können, Koretz zufolge, sehr unterschiedliche Verhaltensweisen der Akteure im pädagogischen Feld stecken, worunter sowohl erwünschte als auch unerwünschte Reaktionsformen fallen.

Eindeutig erwünscht wäre es beispielsweise, wenn Lehrer bessere Arbeit leisten, indem sie etwa auf effektivere Unterrichtsmethoden zurückgreifen. Eindeutig unerwünscht wären dagegen alle Formen des Betrugs, etwa indem vorab an Testitems geübt wird, Hilfestellungen während des Tests gegeben werden oder falsche Antworten korrigiert werden. Zwischen diesen beiden Extremen gibt es eine Grauzone von Vorbereitungen auf den Test, die weniger eindeutig zu bewerten sind. „Teaching to the test“ kann dabei keineswegs pauschal als unerwünscht bezeichnet werden. Es ist vielmehr ein unscharfer Sammelbegriff für unterschiedliche Verhaltensweisen, die im Ergebnis zu steigenden Testleistungen führen. In diese Grauzone fallen z. B. die Umverteilung von Unterrichtszeit innerhalb und zwischen Leistungsdomänen (*reallocation*), die inhaltliche Ausrichtung des Unterrichts auf den Test (*alignment*) sowie die Ausrichtung des Unterrichts am Format von Testaufgaben (*coaching*). Alle drei Varianten können ambivalente Effekte für die Validität gesteigener Testleistungen haben. Dabei ist die curriculare Validität der verwendeten Tests noch keine hinreichende Bedingung für die Validität von

Testleistungen als Indikator für Schülerleistungen (vgl. Koretz 2005). Das Resümee lautet daher: „Absent additional information, one cannot tell from aggregate data what mix of these responses occurred, and one therefore cannot judge how meaningful the gains are“ (Koretz 2003, S. 764). Als Bildungsforscher kann sich Koretz daher nur wundern, wenn Zugewinne in Testleistungen von Politik und Öffentlichkeit als unzweideutige Indikatoren für gestiegene Schülerleistungen interpretiert werden. Interessant ist auch sein Hinweis, dass die Politik erfahrungsgemäß nur wenig Interesse zeigt, Studien zu lancieren, die die Inflation von Testleistungen genauer untersuchen. Hier zeigt sich ein bedeutender Unterschied zum wissenschaftlichen Diskurs, für den die systematische Erforschung von Gegenevidenzen und rivalisierenden Erklärungen ein konstitutives Merkmal ist (vgl. Howe 2002).

Wenn Outputsteuerung wie bei NCLB mit Unterstützungen und Sanktionen verbunden wird, tritt fast zwangsläufig eine gewisse Spannung in der Verknüpfung von Bildungspolitik und Bildungsforschung zu Tage: Während es aus Sicht der Bildungsforschung durchaus wünschenswert sein kann, verschiedene konkurrierende Testinstrumente zum Einsatz zu bringen, um dadurch Referenzmaße für die Validität einzelner Tests zu erhalten, ist aus Sicht der Politik *eine* Datengrundlage erforderlich, nach der kollektiv bindende Entscheidungen über Unterstützungen oder Sanktionen getroffen werden. Da unterschiedliche Tests jedoch in aller Regel unterschiedliche Daten hervorbringen, entscheidet sich die Politik in der Regel für *ein* Testinstrument und beauftragt *eine* Instanz mit der Durchführung des Tests. Mehrere „report cards“, die beispielsweise ein und der gleichen Schule unterschiedliche Ergebnisse bescheinigen, wären als Grundlage für politisch legitime Entscheidungen unbrauchbar. So aber entsteht im Rahmen von Outputsteuerung eine Differenz zwischen dem wissenschaftlichen Wert standardisierter Leistungstests als Instrument der Gewinnung von Wissen und dem politischen Wert solcher Tests im Rahmen der Ausübung von Macht.

Diese Andeutungen mögen genügen, um zu zeigen, dass der Erfolg neuer Steuerungsinstrumente keineswegs so eindeutig bestimmbar ist, wie Protagonisten der gegenwärtigen Reform glauben machen. Zwar stehen im Unterschied zu den bisweilen hochfliegenden Bildungszielen früherer Reformphasen heute leichter operationalisierbare Ziele im Mittelpunkt der Aufmerksamkeit, aber dies ändert nichts daran, dass der Erfolg einer Reform immer auch nach systemspezifischen Kriterien bewertet wird und schon aus wissenschaftsinterner Sicht gestiegene Testleistungen kein unzweideutiger Hinweis auf den Erfolg einer Reform darstellen.

3.3 *Scientifically based research*

Was im Zusammenhang meines Themas an NCLB besonders interessiert, ist der im Gesetzestext an rund 120 Stellen zu findende Hinweis auf „scientifically based research“. Generell wird von Reformstrategien der Bundesstaaten verlangt, dass sie auf Unterrichtsmethoden und -programme zurückgreifen, deren Effektivität durch wissenschaftliche Forschung hinreichend belegt ist (vgl. NCLB 2002, 1473). Unterstützungen für

Schulen oder Schulbezirke, die ihren jährlichen Fortschritt nicht erreichen, werden davon abhängig gemacht, dass Schulen auf entsprechende wissenschaftlich geprüfte Unterrichtsmethoden zurückgreifen (vgl. ebd., 1476).

Das Gesetz enthält auch eine Definition, was unter “scientifically based research” verstanden werden soll. Unter den einzelnen Definitionsmerkmalen fällt besonders die Betonung experimenteller oder quasi-experimenteller Forschungsdesigns auf. Eine besondere Präferenz wird dabei “random assignment experiments” gegeben (vgl. ebd., 1965), also randomisierten kontrollierten Studien, wie sie besonders in der Medizin und Pharmakologie, aber auch in der Agrarwissenschaft als sog. “Goldstandard” der Forschung gelten.⁶ Besonderes Augenmerk gilt der Wiederholbarkeit der in experimentellen Studien nachgewiesenen Effekte, um generalisierbare Schlussfolgerungen aus ihren Ergebnissen ziehen zu können. Während also das Gesetzespaket NCLB bereits Definitionen erziehungswissenschaftlicher Forschung enthält, an denen sich bildungspolitische und pädagogische Reformprogramme zu orientieren haben, gehen jüngste Gesetzesinitiativen einen Schritt weiter: Sie zielen nicht nur auf eine forschungsgestützte Bildungsreform und Bildungspraxis, sondern auf eine Reform erziehungswissenschaftlicher Forschung selbst.

4. Wissenschaftspolitik

Wie jede Reform beginnt auch die Reform der Erziehungswissenschaften mit einer Defizitdiagnose. Seit den 1990er-Jahren mehren sich in den USA die Klagen über den Zustand und den üblen Ruf erziehungswissenschaftlicher Forschung. Vorgetragen werden diese Klagen von hochrangigen Bildungspolitikern (vgl. Castle, zit.n. SRE 2002, S. 28), aber auch von einzelnen prominenten Erziehungswissenschaftlern selbst (vgl. Kaestle 1993) und sogar dem nationalen Wissenschaftsrat (SRE 2002). Der gemeinsame Tenor lautet: Die Erziehungswissenschaft folge jeder Theoriemode, ohne verlässliche Ergebnisse zu erzeugen, bzw. sei z.T. selbst nicht mehr als eine Mode, die sich als Wissenschaft ausbeuge (vgl. Castle, zit.n. COMET 2002). Angesichts der steigenden Erwartungen und Herausforderung, mit denen sich die Bildungsforschung gegenwärtig konfrontiert sieht, wird dieser Zustand immer weniger als hinnehmbar empfunden. Im Auftrag des Nationalen Wissenschaftsrats formuliert eine Gruppe prominenter Bildungsforscher die Herausforderung so:

„In today’s rapidly changing economic and technological environment, schooling cannot be improved by relying on folk wisdom about how students learn and how schools should be organized. No one would think of designing a rocket to the moon or wiping out a widespread disease by relying on untested hunches; likewise, one cannot expect to improve education without research” (SRE 2002, S. 12).

6 Der Terminus lautet „randomized controlled trials“ (RCT), randomisierte kontrollierte Studien. Von „kontrollierten Studien“ spricht man, weil eine Experimental- und eine Kontrollgruppe gebildet werden; Randomisierung bedeutet die zufällige Verteilung der Versuchsteilnehmer auf Experimental- und Kontrollgruppe (vgl. Lind 2005).

Politische Konsequenz aus dieser Defizitdiagnose sind in den USA neuerdings Gesetzesinitiativen der Bundesregierung, die auf eine Reform der bestehenden Bildungsforschung abzielen. Dabei bedient sich die Bundesregierung in einem föderal organisierten System vor allem zwei Mitteln: der Gründung bzw. Wiederbelebung eines nationalen Instituts der Qualitätsentwicklung im Bildungswesen und der Neuregelung der Vergabekriterien für nationale Forschungsgelder.⁷ Besondere Aufmerksamkeit verdient ein 2002 verabschiedetes Gesetzespaket mit dem Titel „Education Sciences Reform Act“ (H.R. 3801), das allgemein als wissenschaftspolitisches Komplementärstück zu NCLB gesehen wird. Oberstes Ziel des Gesetzes ist es zu definieren, was als erziehungswissenschaftliche Forschung angesehen wird, die Anspruch auf Förderung durch zusätzlich bereitgestellte Gelder der Bundesregierung erheben kann.

Die Politik erwartet von der Reform erziehungswissenschaftlicher Forschung unmittelbare Effekte sowohl für die Verbesserung der Bildungspraxis als auch für die Verbesserung der Bildungspolitik. Erklärtes Ziel ist die Ausrichtung der Forschung an Methoden- und Programmevaluation. Die Frage „what works“ soll im Mittelpunkt stehen, wobei die Ziele meist im Voraus fest zu stehen scheinen, bzw. von so hoher Plausibilität sind, dass die elliptische Formulierung ‚finding out what works‘ zumeist ohne eine Diskussion der Ziele oder eine Diskussion bestimmter Kombinationen von Mitteln und Zielen auskommt.

4.1 *Analogie zur Medizin*

Interessant an den Verlautbarungen der Bundesregierung ist eine immer wieder gezo-gene Analogie zwischen Erziehungswissenschaft und Medizin. Insbesondere von der Orientierung am Methodenideal randomisierter kontrollierter Studien verspricht man sich, die Erziehungswissenschaft nun endlich wie die Medizin auf den Weg eines nachhaltigen Fortschritts bringen zukönnen. „It is time to bring education science up to the standards of scientific inquiry that has given us the most advanced medical system in the world“ (Castle, zit.n. COMET 2002). Deutlicher noch wird ein anderer hochrangiger Beamter im Bildungsministerium: „We are close to a point where the right investment in the right structure could get us close to a tipping point, where education becomes an evidence-based field. Medicine only got to that point in the last 75 years.“ (Whitehurst, zit.n. COMET 2002)⁸

7 Zwei Gesetzespakete sind hervorzuheben: Der „Scientifically Based Education Research, Statistics, Evaluation, and Information Act“ von 2000 (H.R. 4875), mit dem das nationale „Office of Research and Improvement“ (OERI) wiederbelebt wurde, und der „Education Sciences Reform Act“ (H.R. 3801) von 2002, der – an der Stelle von OERI – ein nationales Institute for Education Studies (IES) gründete sowie die Vergabe von Forschungsgeldern neu regelt. Zur Wissenschaftspolitik der Bundesregierung und den Initiativen der Scientific Community, darauf Einfluss zu nehmen vgl. im Detail (Eisenhart/Towne 2003).

8 Grover J. Whitehurst ist Assistant Secretary for Educational Research and Improvement.

Schon jetzt werden mehr als 80% aller Forschungsgelder des U.S. Department of Education für randomisierte kontrollierte Studien ausgegeben (vgl. Phillips 2005, S. 584). Das neu gegründete Institut für Qualitätsentwicklung (*Institute for Education Sciences*) finanziert gegenwärtig den Aufbau einer umfangreichen Datenbank des experimentell gesicherten Wissens über bewährte Unterrichtsmethoden und -programme. Zu diesem Zweck wird geprüft, inwieweit neuere empirische Studien strengen methodologischen Standards genügen. Die Studien werden sodann mit einem Qualitätssiegel versehen und im Internet aufgelistet.⁹

4.2 Kritische Diskussion in der Scientific Community

In der *scientific community* haben die Gesetzesinitiativen der Bundesregierung und die dahinter stehenden Erwartungen inzwischen zu einer lebhaften kritischen Diskussion geführt. Ein Markstein dieser Debatte ist eine Veröffentlichung mit dem Titel „Scientific Research in Education“, die 2002 im Auftrag des Nationalen Wissenschaftsrates entstand. Ziel dieser Initiative war es, im Streit zwischen der Wissenschaftspolitik der Bundesregierung einerseits und einer zunehmend beunruhigten Forschergemeinschaft andererseits zu schlichten. Zu diesem Zweck wurde eine Gruppe hochrangiger und angesehener Fachvertreter aus unterschiedlichen Disziplinen angefragt, die eine breite Konsensbasis im Streit um erziehungswissenschaftliche Forschung herstellen sollten. Auf den ersten Blick erkennt man in dem bemerkenswerten Dokument der Expertengruppe den Versuch, abgewogene und differenzierte Aussagen über die Natur wissenschaftlicher Forschung in der Erziehungswissenschaft zu formulieren (vgl. SRE 2002, XII).

Im Vergleich mit den Gesetzesinitiativen der Bundesregierung zeichnet die Expertengruppe ein vielschichtigeres Bild erziehungswissenschaftlicher Forschung. Betont wird, dass sich die Qualität wissenschaftlicher Forschung nicht an der Befolgung einer bestimmten Methode festmachen lässt. Ausdrücklich plädiert man für eine Pluralität der Forschungsmethoden in der Erziehungswissenschaft. Letzter Maßstab für die Qualität wissenschaftlicher Forschung seien die selbstregulativen Normen der Forschergemeinschaft selbst. Statt einer Orientierung an bestimmten Methoden versucht man, Leitprinzipien zu formulieren, die aller wissenschaftlichen Forschung zugrunde liegen. Dennoch hat auch die Publikation von *Scientific Research in Education* die Debatte um das Selbstverständnis und die Zukunft erziehungswissenschaftlicher Forschung nicht beenden können. Im Gegenteil wurde das als breite Konsensbasis gedachte Dokument inzwischen selbst zum Gegenstand einer lebhaften und scharfen Debatte.¹⁰

Diskussionsbedürftig erscheint zunächst das als alternativlos präsentierte einheitswissenschaftliche Modell, das dem gesamten Dokument zugrunde liegt. Im Kern, so

9 Vgl. <http://www.whatworks.ed.gov/>

10 Verschiedene Zeitschriften haben inzwischen Schwerpunktheft zu dieser Debatte veröffentlicht: Vgl. insbesondere *Qualitative Inquiry* 10, No. 1; *Teachers College Record* 107, No. 1; *Educational Theory* 55, No. 3. Die Debatte spielte auch auf dem Kongress der AERA von 2003 eine wichtige Rolle. Vgl. hierzu diverse Hefte von *Educational Researcher*.

hatte die Expertengruppe erklärt, sei wissenschaftliche Forschung in allen Gebieten gleich (vgl. SRE 2002, S. 2). Die Problematik dieses einheitswissenschaftlichen Ansatzes tritt insbesondere dann zutage, wenn er zum Ausgangspunkt gewagter Analogien zwischen medizinischer und erziehungswissenschaftlicher Forschung dient, die auch in den Verlautbarungen der Bundesregierung eine wichtige Rolle spielen. Auch die Medizin, so liest man im Dokument der Expertengruppe, war noch vor nicht allzu langer Zeit ein Feld, das weitgehend nicht auf Forschung, sondern auf praktischer Erfahrung beruhte. Wenn im Bildungsbereich vergleichbare Fortschritte wie in der modernen Medizin möglich werden sollen, dann müsse die Kluft zwischen Wissensformen der pädagogischen Praxis und der Erziehungswissenschaften geschlossen werden (vgl. SRE 2002, S. 14f). Randomisierte kontrollierte Studien und die mit ihnen verknüpften quantitativen Korrelationsmethoden werden in diesem Zusammenhang zwar nicht als alternativlos angesehen, aber doch als das zuverlässigste Forschungsdesign zur Ermittlung kausaler Zusammenhänge (vgl. ebd., S. 109f.). Qualitative und interpretative Ansätze werden zwar gewürdigt, ihnen wird aber eher eine unterstützende Funktion zugewiesen oder sie werden auf Situationen beschränkt, in denen randomisierte kontrollierte Studien nicht durchführbar sind. Eng mit diesem Methodenideal verknüpft ist die Forderung nach Wiederholbarkeit und Generalisierbarkeit, der interpretative sozialwissenschaftliche Ansätze häufig nicht genügen. Gerade von einem an der Medizin orientierten Forschungsmodell verspricht sich die Expertengruppe effektive Steuerungstechnologien für Pädagogen und Politiker: „Educators today require new knowledge to reengineer schools in effective ways“ (ebd., S. 12). Schließlich stellt sich die Frage, ob das proklamierte Vertrauen in die selbstregulativen Normen der Forschergemeinschaft nicht schon dadurch relativiert wird, dass Standards für „gute Forschung“ gewissermaßen stellvertretend für die *scientific community* im Auftrag des nationalen Wissenschaftsrats durch ein Expertengremium formuliert werden.

Unverkennbar handelt es sich also bei „Scientific Research in Education“ nicht bloß um eine akademische wissenschaftstheoretische und methodologische Abhandlung, sondern um eine überaus ambivalente Stellungnahme in einem Richtungsstreit um die Zukunft der Erziehungswissenschaft in einem veränderten bildungs- und wissenschaftspolitischen Kontext. Dabei geht es nicht zuletzt auch um die Definitionsmacht für ‚gute Forschung‘ im Zeitalter Neuer Steuerung.

5. Kritische Auswertung

Ich verlasse nun meine beiden Beispiele aus der US-amerikanischen Reformdebatte, um am Ende vier Problemzonen im gegenwärtigen Verhältnis von Bildungspolitik und Bildungsforschung zu beleuchten, die m.E. eine über die untersuchten Beispiele hinausweisende Bedeutung haben. Die Auseinandersetzung mit diesen Problemzonen könnte daher auch für die aktuellen Kontroversen in der deutschen Erziehungswissenschaft weiterführende Impulse geben.

5.1 Performative Effekte Neuer Steuerung

Im Kontext Neuer Steuerung, darauf weisen die behandelten Beispiele hin, scheint sich ein verändertes Verhältnis von wissenschaftlichem Wissen und politischer Macht zu etablieren. Das traditionelle Modell der Politikberatung ging davon aus, dass wissenschaftliches Wissen bestenfalls eine Informations- und Entscheidungsgrundlage für politische Steuerung darstellt. Diesem Modell zufolge geht die eigentliche Steuerungswirkung von politischen Maßnahmen aus, die auf der Grundlage wissenschaftlicher Expertise getroffen werden. In der häufig verwendeten Redewendung „Allein vom Wiegen wird die Sau nicht fett“ kommt diese Denkweise zum Ausdruck, wonach Qualitätskontrollen nur die Voraussetzung für die eigentlichen Maßnahmen der Qualitätsentwicklung seien.

Im Zeitalter Neuer Steuerung dagegen sind Kontrollen selbst schon ein Moment der Steuerung und nicht bloß ihre Voraussetzung. Die angesprochene Inflation von Testleistungen wirft ein erhellendes Licht auf diese *performativen Effekte* Neuer Steuerung. Schon die Tatsache, dass flächendeckende Qualitätskontrollen durchgeführt werden, kann eine Lenkungswirkung haben, auch unabhängig davon, welche konkreten Maßnahmen und Interventionen zur Qualitätsentwicklung eingeleitet werden. Bezogen auf die deutsche Situation brachte der Chef von McKinsey Deutschland, Jürgen Kluge, den performativen Effekt von Qualitätskontrollen pointiert zum Ausdruck: „Pisa hat bewiesen: Sobald Sie etwas messen, wird es besser“ (Kluge 2005, S. 96).

Gerade die Politik profitiert von den performativen Effekten neuer Steuerung. In Kombination mit *choice policies* kann mehr und mehr darauf gesetzt werden, dass der Wettbewerb für die Entwicklung von Qualität sorgt, während sich die Politik auf Standardsetzung und Qualitätskontrollen konzentriert. In dieser Neubestimmung öffentlicher Aufgaben verändern sich auch die politischen Inklusionsstrategien: Mit der Erfassung aller Schüler durch regelmäßige Qualitätskontrollen kann die Politik Verantwortung für das gesamte Bildungswesen signalisieren und sich zugleich von wohlfahrtsstaatlich begründeten Inklusionsstrategien mehr und mehr zurückziehen.

5.2 Reflexivität

Für die Bildungsforschung bedeuten die angesprochenen performativen Effekte Neuer Steuerung die Herausforderung, eine größere kritisch-reflexive Aufmerksamkeit für die Mitkonstruktion ihres eigenen Forschungsgegenstands zu entwickeln. Reinhard Pekrun hat unlängst zu internationalen Leistungsvergleichsstudien angemerkt: „In teils impliziter, teils expliziter Weise hat die internationale und nationale Bildungsdiskussion vielerorts begonnen, sich an den Kompetenzdefinitionen zu orientieren, die durch die jeweiligen Aufgabenserien dieser Studien bereitgestellt werden. Sinnvoll schiene es mir“, so Pekrun weiter, „solche Rückwirkungen von Evaluationsstudien empirisch zu erfassen und bildungstheoretisch angemessen zu bewerten“ (Pekrun 2002, S. 122). Die hier angemahnte Aufmerksamkeit für die Formierung des Gegenstands der Bildungsforschung

durch die Forschung selbst bedeutet zugleich eine größere Aufmerksamkeit für die *politische* Dimension der Bildungsforschung in ihrer Rolle als aktive Mitspielerin im Reformprozess. Solche Reflexivität ist freilich ein Qualitätsmerkmal der gesamten Forschungslandschaft einer Disziplin und kein exklusives Merkmal einzelner Teildisziplinen.

Dennoch fällt auf, dass gegenwärtig das Interesse nach Generalisierbarkeit und Übertragbarkeit experimentell erforschter Effekte zu Lasten derjenigen Teildisziplinen zu gehen scheint, die gerade ein reflexives Wissen über die Grenzen von Generalisierbarkeit und Übertragbarkeit entwickelt haben. Ich nenne nur die Historische Bildungsforschung, die eher ein Wissen über das entwickelt hat, was unter welchen Voraussetzungen und Bedingungen *nicht* funktioniert, ohne damit schon über ein Wissen zu verfügen, was in Zukunft funktionieren wird (vgl. Popkewitz 2004). Zu nennen wäre weiter eine Form der International Vergleichenden Erziehungswissenschaft, die gerade nicht an der Übertragbarkeit von Erfolgsmodellen interessiert ist, sondern an der Erforschung kulturspezifischer Voraussetzungen und Anverwandlungen globaler Modelle (vgl. Schriewer 1994). Zu nennen wäre auch eine Form der Erziehungs- und Bildungsphilosophie, die in der Reflexion der theoretischen Voraussetzungen und der Machteffekte wissenschaftlicher Forschung ein Bewusstsein für die gegenstandskonstitutive Seite einer praktischen Wissenschaft entwickelt hat (vgl. Biesta/Burbules 2003). Jedenfalls scheint die in den USA sich abzeichnende Tendenz, gerade reflexive Forschungstraditionen in der Erziehungswissenschaft im Zeichen eines bestimmten Methodenideals zu marginalisieren, der politischen Selbstaufklärung der Bildungsforschung eher entgegenzuwirken.

5.3 Experimentelle und interpretative Forschung

In der hierdurch entfachten wissenschaftspolitischen Kontroverse droht ein unproduktiver Rückfall in alte Lagerkämpfe. Unterschiedliche Forschungsparadigmen – quantitative vs. qualitative, experimentelle vs. interpretative – werden gegeneinander ausgespielt, statt sie konstruktiv miteinander zu verknüpfen (vgl. Howe 2004). Diese Lagerkämpfe befördern eine Arbeitsteilung, die sich gerade im Kontext Neuer Steuerung als äußerst problematisch erweist. So zeichnet sich – zumindest in der Wissenschaftspolitik der US-Regierung – eine einseitige Privilegierung einer bestimmten Art von Bildungsforschung ab, die die Frage, was funktioniert, unabhängig davon meint beantworten zu können, welche Reformmaßnahme von wem wie interpretiert wird. Die Arbeitsteilung zwischen einer Forschung, die technologisches Steuerungswissen bereitstellt, und einer Forschung, die unterschiedliche Interpretationen sozialer Wirklichkeit durch die Akteure untersucht, führt aber zu systematischen Fehleinschätzungen von Steuerung und Steuerbarkeit.

Gerade Erziehungswissenschaftler könnten wissen, dass jede Steuerung sozialer Beziehungen mit dem Sachverhalt der *doppelten Kontingenzen* rechnen muss und jede Erforschung der sozialen Welt an eine *doppelte Hermeneutik* gebunden ist. Eine Bildungsfor-

schung, die für die Rekursivität im Verhältnis zu ihrem Gegenstand kein methodisches Instrumentarium mehr hätte, könnte am Ende erwartungswidrige Reformfolgen nur als nicht-intendierte Nebenfolgen konzipieren, die bestenfalls ein gezieltes Gegensteuern verlangen. Die nicht-intendierten Nebenfolgen neuer Steuerungsinstrumente sind aber nicht allein Folge technischer Fehler im Steuerungsmodell, sondern ganz wesentlich auch Folge davon, dass die sich im Rahmen Neuer Steuerung etablierende Ordnung von den Akteuren sehr unterschiedlich ‚entziffert‘ und ‚genutzt‘ wird (vgl. Zymek 2000). Die Allgegenwart eines solchen kreativen Umgangs mit Reformen könnte für die Bildungsforschung Anlass genug sein, die Frage nach dem, was in der Erziehungswirklichkeit funktioniert, in enger Verbindung mit der Frage nach den differenten Deutungen der Reform durch die Akteure zu beantworten. Hierbei wird man über bloße Akzeptanzforschung hinausgehen müssen. Sicher scheint, dass durch eine einseitige Präferenz für randomisierte kontrollierte Studien und statistische Korrelationsmethoden die komplexen Erfolgsbedingungen von Reformen nicht hinreichend erfasst werden können.

5.4 Umdefinition des Politischen

Die empirische Wende in der Bildungspolitik, die in den USA unter dem Terminus „evidence based politics“ zugleich gefordert und vollzogen wird, führt aber nicht nur zu einer problematischen Umgewichtung innerhalb der Bildungsforschung, sondern auch zu einer problematischen Neudefinition des Politischen. Suggestiert wird, Politik bestehe im Kern darin, noch unbekannte Lösungen für bekannte Probleme zu finden. Tatsächlich aber dürfte dieser Fall eher ein seltener Grenzfall politischen Handelns darstellen, der auf eine kleine Zahl überschaubarer Probleme beschränkt ist. In der Regel aber hat es die Politik eher mit hochgradig komplexen Problemen zu tun, und die Schwierigkeit besteht nicht so sehr darin, überhaupt einen Lösungsansatz zu finden, sondern für eine der möglichen und andernorts häufig bereits realisierten Lösungen Mehrheiten und Legitimität zu gewinnen (vgl. Lindblom 1959/1995). So gesehen besteht die Wirklichkeit nicht so sehr aus lauter Problemen, sondern aus lauter Problemlösungen, die z.T. konkurrieren. Durch die Umdefinition des politischen Kampfes um legitime Entscheidungen in eine technologische Problematik des Findens einer funktionierenden Lösung wird gewissermaßen der demokratische Prozess unterlaufen. Im Streit um die Vielfalt etablierter Problemlösungen ist aber Wissenschaft nicht oberste Richterin, sondern nur eine Stimme unter anderen. Gerade insofern neue Steuerungsinstrumente ein ganzes Bündel von z.T. konkurrierenden Zielen öffentlicher Bildung gleichzeitig betreffen und diese Effekte von unterschiedlichen Akteuren unterschiedlich bewertet und gewichtet werden, kommt es für eine demokratische Öffentlichkeit vor allem darauf an, dass alle Stimmen gehört werden (vgl. Howe 2004). Welche Rolle Bildungsforschung dabei spielt, ist durchaus offen: Sie kann dazu beitragen, unterschiedlichen Akteuren und unterschiedlichen Bewertungskriterien Gehör zu verschaffen, sie kann aber auch bei der problematischen Umdefinition des Politischen mehr oder weniger offen kollaborieren.

Vielleicht ist aber auch diese Alternative zu dramatisch formuliert. Es sollte nur deutlich gemacht werden, dass Bildungsforschung im Kontext Neuer Steuerung zu einem konstitutiven Moment der Politik wird, ohne dass damit schon entschieden wäre, wie sie ihre neue Rolle ausfüllt, mit welchem Grad an kritischer Reflexivität nach innen und Transparenz nach außen. Da sehe ich eine ganze Bandbreite attraktiver Möglichkeiten, jenseits von Woody Allens Alternative zwischen „utter despair“ und „total extinction“.

Literatur

- Allen, W. (1981): *My Speech to Graduates*. In: Ders.: *Side Effects*. New York: Ballantine Books.
- Betebenner, D.W./Howe, K.R./Foster, S.S. (2005): *On School Choice and Test-Based Accountability*. In: *Education Policy Analysis Archives* 13, S. 1-18. Verfügbar unter: <http://epaa.asu.edu/epaa/v13n41/>
- Biesta, G./Burbules, N. (2003): *Pragmatism and Educational Research*. Lanham: Rowman and Littlefield.
- CEP (2005): *Center on Education Policy: From The Capital to the Classroom. Year 3 of the No Child Left Behind Act*. March 2005. Verfügbar unter: http://www.cep-dc.org/pubs/nclby3/press/cep-nclby3_21Mar2005.pdf.
- Cizek, G.J. (1999): *Give Us This Day Our Daily Dread: Manufacturing Crises in Education*. In: *Phi Delta Kappan* 80, No. 10, S. 737-743. Verfügbar unter: <http://www.pdkintl.org/kappan/kciz9906.htm>.
- COMET (2002): *California Online Mathematics Education Times* 3, No. 10, 12 March 2002. Verfügbar unter: http://csmmp.ucop.edu/cmp/comet/2002/03_12_2002.html.
- DFG (2001): *Stellungnahme zur strukturellen Stärkung der empirischen Bildungsforschung*. Bonn. Verfügbar unter: http://www.dfg.de/aktuelles_presse/reden_stellungnahmen/download/empirische_bildungsforschung_st.pdf.
- Döbert, H. (2003): *Neue Steuerungsmodelle von Schulsystemen in Europa?* In: Ders. u.a. (Hrsg.): *Bildung vor neuen Herausforderungen. Historische Bezüge – Rechtliche Aspekte – Steuerungsfragen – Internationale Perspektiven*. Neuwied: Luchterhand.
- Eisenhart, M./Towne, L. (2003): *Contestation and Change in National Policy on „Scientifically Based“ Education Research*. In: *Educational Researcher* 32, No. 7, 31-38.
- Haney, W. (2000): *The Texas miracle in education*. *Education Policy Analysis Archives* 8. Verfügbar unter: <http://epaa.asu.edu/epaa/v8n41/>.
- Howe, K.R. (2002): *Free Market Free-for All*. In: *Education Week*, April 10, 2002. Verfügbar unter: <http://www.asu.edu/educ/epsl/EPRU/articles/EPRU-0204-32OWI.doc>.
- Howe, K.R. (2004): *A Critique of Experimentalism*. In: *Qualitative Inquiry* 10, S. 42-61.
- H.R. 3801. *Education Sciences Reform Act of 2002*. Verfügbar unter: <http://thomas.loc.gov>.
- H.R. 4875. *Scientifically Based Research in Education, Statistics, Evaluation, and Information Act of 2000*. Verfügbar unter: <http://thomas.loc.gov>.
- Kaestle, C. (1993): *The Awful Reputation of Educational Research*. In: *Educational Researcher* 22, S. 23-31.
- Kluge, J. (2005): *Mehr Staat! McKinsey will die Kindergärten umkrepeln. Die Milliarden dafür soll der Bund zahlen. Ein Gespräch mit dem Deutschland-Chef der Unternehmensberatung*. In: *Die Zeit*, Nr. 44, S. 96, 27.10. 2005. Verfügbar unter: http://www.zeit.de/2005/44/C-Interview-F_9arderung?page=all.
- Koch, L. (2004): *Normative Empirie*. In: Heitger, M. u.a.: *Kritik der Evaluation von Schulen und Universitäten*. Würzburg: Ergon, S. 39-55.

- Koretz, D.M. (2003): Limitations in the Use of Achievement Tests as Measures of Educators' Productivity. In: *The Journal of Human Resources* 37, S. 752-777.
- Koretz, D. (2005): Alignment, High Stakes, and the Inflation of Test Scores. CSE Report 655. Center for the Study of Evaluation (CSE), UCLA 2005. Verfügbar unter: <http://www.cse.ucla.edu/>.
- Lind, G. (2005): Evaluation von Bildung – Eine systematische Einführung in die entwicklungsorientierte Evaluation. Verfügbar unter: http://www.uni-konstanz.de/ag-moral/evaluation/k_eva_theorie.htm.
- Lindblom, Ch.E. (1959/1995): The ‚Science‘ of Muddling Through. In: Theodoulou, S./Can, M. (Hrsg.): *Public Policy: The Essential Readings*. Englewood Cliffs: Prentice Hall, S. 113-127.
- Lundgren, U.P. (2001): Political Governing of Education and Decentralization. Keynote Speech at the ATEE Annual Conference, 28. August 2001. Verfügbar unter: www1.lhs.se/atee/proceedings/Ulf_P_Key_note.doc.
- McCull, A. (2005): Tough Call: Is No Child Left Behind Constitutional? In: *Phi Delta Kappan* 86, No. 8, S. 604-610.
- Nicolas, S.L./Glass, G.V/Berliner, D.C. (2006): High-stakes testing and student achievement: Does accountability pressure increase student learning? In: *Education Policy Analysis Archives* 14, S. 1-172. Verfügbar unter <http://epaa.asu.edu/epaa/v14n1/>
- NCLB (2002): No Child Left Behind Act of 2001. Public Law 107-110, 107th Congress. January 8th 2002. Verfügbar unter: <http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf>.
- NWEA (2005): Northwest Evaluation Association: The Impact of the No Child Left Behind Act On Student Achievement and Growth: 2005 Edition. Verfügbar unter: http://www.greatlakes-center.org/pdf/NCLBImpact_2005_Study.pdf.
- Oelkers, J. (2005a): Bildungsstandards, Tests und Schulentwicklung: Eine bildungspolitische Option. Vortrag am IQB der Humboldt Universität Berlin am 11. Oktober 2005. Verfügbar unter: http://www.paed.unizh.ch/ap/downloads/oelkers/Vortraege/210_IQBvortrag.pdf.
- Oelkers, J. (2005b): Der Handlungsspielraum der Bildungspolitik. Eine wissenschaftliche Perspektive. Vortrag auf der nationalen PISA-Tagung am 25. November 2005 im Bundesamt für Statistik. Verfügbar unter: http://www.paed.unizh.ch/ap/downloads/oelkers/Vortraege/221_Neuenburg.pdf.
- Pekrun, R. (2002): Vergleichende Evaluationsstudien zu Schülerleistungen: Konsequenzen für zukünftige Bildungsforschung. In: *Zeitschrift für Pädagogik* 48, S. 111-128.
- Phillips, D.C. (2005): The Contested Nature of Empirical Educational Research (and Why Philosophy of Education Offers Little Help). In: *Journal of Philosophy of Education* 39, S. 577-597.
- Popkewitz, Th.S. (2004): Is the National Research Council Committee's Report on Scientific Research in Education Scientific? On Trusting the Manifesto. In: *Qualitative Inquiry* 10, S. 62-78.
- Roth, H./Friedrich, D. (Hrsg.) (1975): *Bildungsforschung. Probleme – Perspektiven – Prioritäten*. 2. Bde. Stuttgart: Klett.
- Schriewer, J. (1994): Internationalisierung der Pädagogik und Vergleichende Erziehungswissenschaft. In: Müller, D.K. (Hrsg.): *Pädagogik – Erziehungswissenschaft – Bildung. Eine Einführung in das Studium*. Köln/Weimar/Wien: Böhlau, S. 427-462.
- SRE (2002): *Scientific Research in Education*. Committee on Scientific Principles for Education Research, edited by R.J. Shavelson and L. Towne. Washington: National Academy Press.
- Spellings, M. (2005): A Commitment to Doing What Works. Prepared Remarks for Secretary Spellings at the Convention of the National Association of Elementary School Principals. Verfügbar unter: <http://www.ed.gov/news/speeches/2005/04/04182005.html>.
- Weiler, H.N. (2003): Bildungsforschung und Bildungsreform – Von den Defiziten der deutschen Erziehungswissenschaft. In: Gogolin, I./Tippelt, R. (Hrsg.): *Innovation durch Bildung*. Beiträge zum 18. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft. Opladen: Leske & Budrich, S. 181-203.

- Winter, G. (2005): Study Finds Shortcoming in Law on Education. In: New York Times, April 13, 2005, S. 15.
- Zymek, B. (2000): Re-Partikularisierung des Bildungssystems? Historische Anmerkungen zu aktuellen Strategien der Schulreform. In: Schlömerkemper, J. (Hrsg.): Differenzen. Über die politische und pädagogische Bedeutung von Ungleichheiten im Bildungswesen (= Die Deutsche Schule, Beiheft 6), S. 6-20.

[Die Verfügbarkeit aller Internet-Ressourcen wurde am 28. Februar zuletzt überprüft.]

Anschrift des Autors:

Prof. Dr. Johannes Bellmann, Pädagogische Hochschule Freiburg, Institut für Erziehungswissenschaft I, Kunzenweg 21, 79117 Freiburg.