

Besserer Unterricht dank Bildungsstandards und Kompetenzmodellen?*

Walter Herzog

Für eine klare Beantwortung der Frage im Titel meines Referats ist es vermutlich noch zu früh. Ob Bildungsstandards und Kompetenzmodelle zur Verbesserung des Unterrichts beitragen oder nicht, lässt sich zurzeit wohl nicht abschließend sagen. In Ländern wie Deutschland, Österreich und der Schweiz sind wir eben erst dabei, die standardbasierte Schulreform umzusetzen.

Allerdings sind wir nicht die Ersten, die mit dieser Art von Reform experimentieren. In den USA gibt es seit 1965 und intensiviert seit 1983 Versuche, das Schulsystem über Standards zu steuern (vgl. Gamoran 2007, p. 80ff.; Mabry 2004, p. 50). Auch wenn die Situation in den USA nicht direkt mit unserer eigenen Situation vergleichbar ist, gibt es genügend Parallelen und Übereinstimmungen, um die Fragestellung meines Referats anhand einiger Studien aus den USA anzugehen.

Um die Differenz zwischen den Ländern nicht einzuebnen, werde ich mich allerdings nicht ausschließlich auf empirische Studien stützen, sondern eine Gesamtschau der standardbasierten Bildungsreform versuchen. Wenn wir etwas besser verstehen, was die theoretischen Grundlagen der Reform sind, werden wir auch eher in der Lage sein zu beurteilen, was mit Bildungsstandards und Kompetenzmodellen auf uns zukommt.

* Referat an der Tagung „Bildungsstandards und Kompetenzmodelle“ der Kommissionen *Schulforschung/Didaktik* und *Professionsforschung und Lehrerbildung* der Deutschen Gesellschaft für Erziehungswissenschaft in Heidelberg vom 25.-27. März 2009.

Ich beginne mit der Frage, wodurch sich die standardbasierte Reform von bisherigen, ähnlichen Ansätzen der Schulreform unterscheidet (1). Danach werde ich die Logik herausarbeiten, die der standardbasierten Schulreform zugrunde liegt (2). Im dritten und vierten Schritt werde ich der Umsetzung der Logik im Rahmen von *High-Stakes Tests* (3) und im Rahmen von Kompetenzmodellen (4) nachgehen. Schließlich werde ich mit einem kurzen Fazit schließen (5).

(1) Beginnen wir also mit der Frage, wo der Ansatzpunkt der standardbasierten Schulreform liegt.

Ein Schulsystem kann man nach vier Ebenen differenzieren:

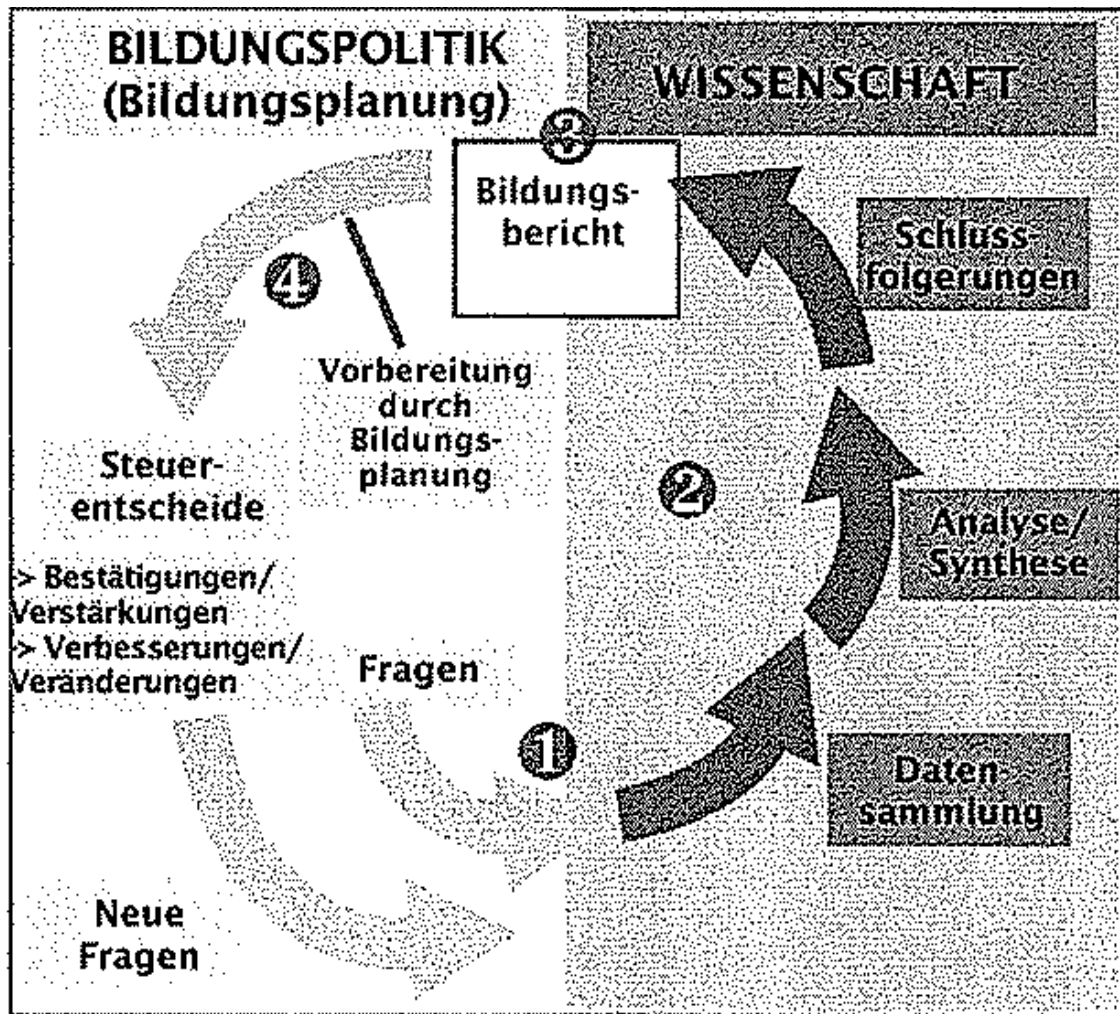
Individuen: Hier finden wir die einzelnen Schülerinnen und Schüler sowie die Lehrkräfte.

Schulklassen: Der Unterricht findet im Normalfall im Rahmen einer administrativen Zusammenführung von Individuen zu Gruppen statt.

Einzelschulen: Die Schulklassen fügen sich zu Schulen als bürokratischen Organisationseinheiten zusammen.

Schulsystem: Das Schulsystem umfasst die Gesamtheit der Einzelschulen sowie den administrativen Apparat, der die Schulen „steuert“.

Bildungsstandards betreffen zunächst einmal die Systemebene des Schulsystems. Es sind „Steuerungsinstrumente“ zuhanden der Politik und der Bildungsverwaltung. Damit gehören sie in den größeren Kontext des Bildungsmonitoring. Das kann das folgende Schema zeigen; es stammt von der EDK, der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren (dem Analogon zur deutschen Kultusministerkonferenz).



EDK 2006, p. 30

Zwar ist hier nicht von Bildungsstandards die Rede, doch gibt es genügend Hinweise, die zeigen, dass die „schweizerischen Bildungsstandards“ Teil des Bildungsmonitorings bilden, das alle vier Jahre durchgeführt werden soll. Das Bildungsmonitoring seinerseits steht im Dienste der „bessere[n] Steuerung des Bildungssystems“ (Maradan & Mangold 2005, S. 7).

Das ist in Deutschland und Österreich – soweit ich sehe – nicht anders. Bildungsstandards dienen der Beobachtung des Bildungssystems, das dank der genaueren Beobachtung besser gesteuert werden soll. Dies ist

zumindest auch die Botschaft der Klieme-Expertise.¹ Dabei liegt in der verbesserten Systemsteuerung kein Selbstzweck; vielmehr steht sie im Dienste der Qualitätssicherung und Qualitätsverbesserung von Schule und Unterricht (vgl. Klieme et al. 2003, p. 9, *passim*; Maradan & Mangold 2005, p. 3f.; Rowe 2000, p. 73f.). Qualitätsverbesserung meint allerdings in erster Linie Steigerung der Lernleistungen der Schülerinnen und Schüler. „Nationale Bildungsstandards“, so die Klieme-Expertise, „setzen den Standard für die Leistungen der *Schule*, und zwar so, dass man ihn an *Individuen* vergleichend messen kann“ (Klieme et al. 2003, p. 55 – Hervorhebung W.H., vgl. auch ebd., p. 139).

Interessant an diesem Zitat ist die Akzentverschiebung von den Schülerinnen und Schülern zur Schule: Es sind *deren* Leistungen, die zur Diskussion stehen. Die Schülerinnen und Schüler sind nur die Datenträger, an denen sich entscheidet, ob eine Schule die geforderte Qualität erbringt oder nicht. Obwohl man sich das Lernen nicht anders denn als (individuellen) Vorgang vorstellen kann, der durch die Lernenden selbst hervorgebracht wird, stehen Bildungsstandards in einem Denkraum, der in der Schule eine Kausalursache für die Schülerleistungen sieht. Dafür steht das paradoxe Konzept der *Output-Steuerung*: Die Schulen sollen über ihren Output gesteuert werden, und dieser besteht aus den Lernleistungen der Schülerinnen und Schüler.

Instanz der Output-Steuerung ist die Politik. Sie definiert die „Maßstäbe, nach denen der ‚Output‘ erfasst und bewertet werden kann“ (Klieme et al. 2003, p. 99). Doch der Clou – das „kennzeichnende Merkmal“, wie ich einleitend sagte – der standardbasierten Schulreform liegt darin, dass sich

¹ Bildungsstandards stellen „einen Fortschritt für die Steuerung unsers Schulsystems dar...“ (Klieme et al. 2003, p. 29). Sie „werden ... international als normative Vorgaben für die Steuerung von Bildungssystemen verstanden“ (ebd., p. 32).

die Steuerung nicht auf die Systemebene beschränkt. Vielmehr wird erwartet, dass sie sich auf die unteren Ebenen – Einzelschule, Unterricht und Individuen – überträgt.

(2) Doch wie soll dies möglich sein? Ich komme zum zweiten Teil meiner Ausführungen. Wie kann sich ein Steuerungsimpuls, den die Politik auf der Systemebene gibt, auf die unteren Ebenen des Bildungssystems übertragen? Meine Antwort ist, dass dies nur im Rahmen eines technokratischen Reformansatzes gelingen kann. Den Schulen werden Leistungsziele vorgegeben, die sie in einem definierten Zeitraum bzw. zu einem bestimmten Zeitpunkt in der schulischen Karriere eines Kindes erreichen müssen (*content* und *performance standards*), die Zielerreichung wird mittels Tests überprüft (*assessment* bzw. *monitoring*), die Schule muss über ihre Leistungen Rechenschaft ablegen (*accountability*), und je nach Diskrepanz zwischen Ist- und Soll-Wert werden Maßnahmen getroffen (*high-stakes decisions, sanctioning*).

James Popham, der die standardbasierte Schulreform in den USA mit kritischer Aufmerksamkeit begleitet, führt die Logik des Reformansatzes auf das Zweck-Mittel-Schema zurück. Im ersten Schritt werden Ziele gesetzt, die man auf einem operationalen Niveau messbar macht. Im zweiten Schritt werden die Ziele umgesetzt. Im dritten Schritt wird mittels standardisierter Verfahren überprüft, ob die Ziele erreicht wurden oder nicht. „SBE [Standards-based education, W.H.]“, so Popham, „is nothing more than a posh *ends-means* model wherein content standards represent intended ends, teaching constitutes the means for achieving those ends, and test results supply the evidence regarding whether the means did, in fact, achieve the intended ends. ... SBE is, at bottom fundamentally nothing

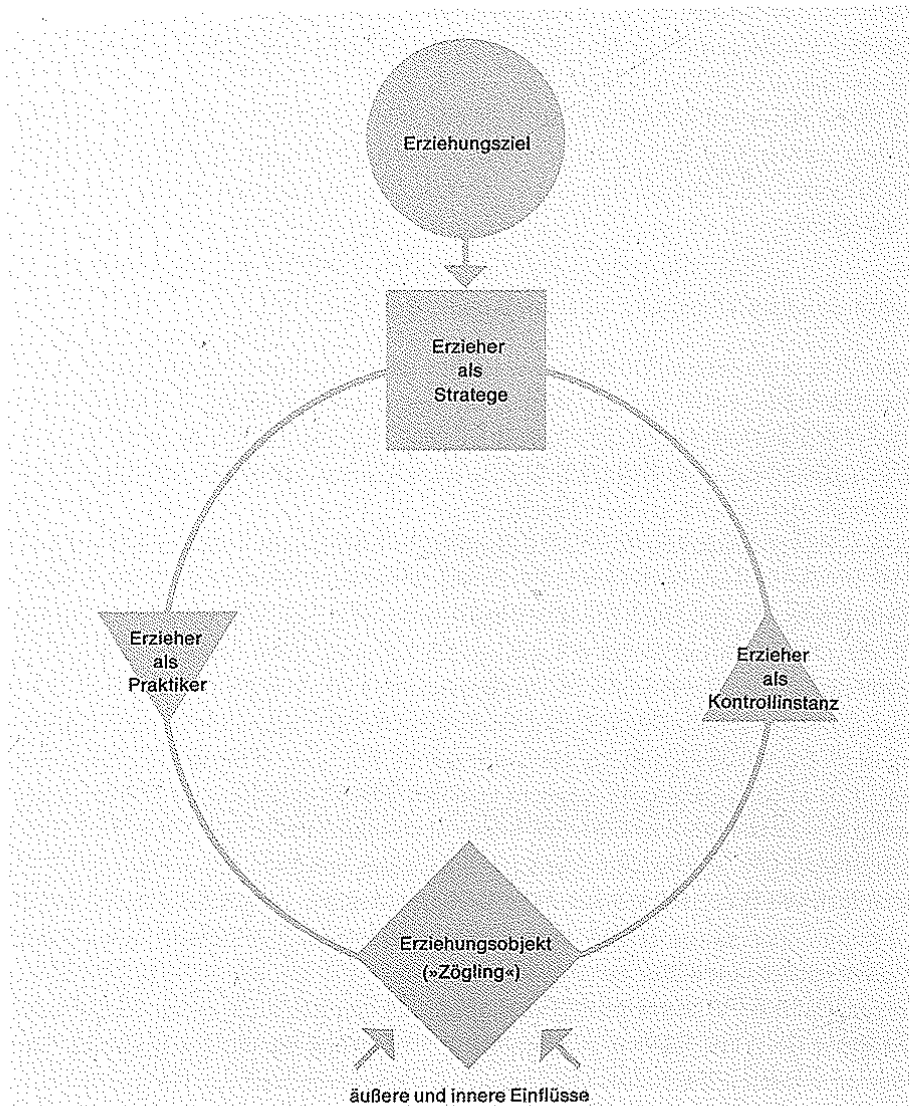
more than the educational application of the traditional ends-means paradigm” (Popham 2004, p. 17f.).

Gegen dieses Modell kann man zunächst so viel vorbringen, wie sich gegen eine rein zweckrationale Begründung von Schule und Unterricht einwenden lässt. Ethisch beleuchtet, ist das Schema utilitaristisch² und unterbietet eine an (ethischen) Prinzipien orientierte Begründung von Bildung und Erziehung.

Die Kritik braucht sich aber nicht auf ethische Argumente zu beschränken. Historisch kann man die Geschichte der Schulreformen durchleuchten und Fälle aufsuchen, wo auf gleiche Weise versucht wurde, den Unterricht zu verbessern. Dabei wird man feststellen, dass es bisher noch selten gelungen ist, mit einem technokratischen Ansatz eine nachhaltige Verbesserung der Schule zu bewirken. Man denke an die Reformen der 1970er und 80er Jahre: an die Curriculumtheorie, die Lernzieltaxonomien, den programmierten Unterricht oder die kybernetische Didaktik.

Was Letztere anbelangt, zeigt sich der technokratische Ansatz schon rein in der Anschauung:

² „It is essentially a utilitarian ethic that underlies test-driven curricular reform, one based on means-ends arguments“ (Mathison & Freeman 2004, p. 90).

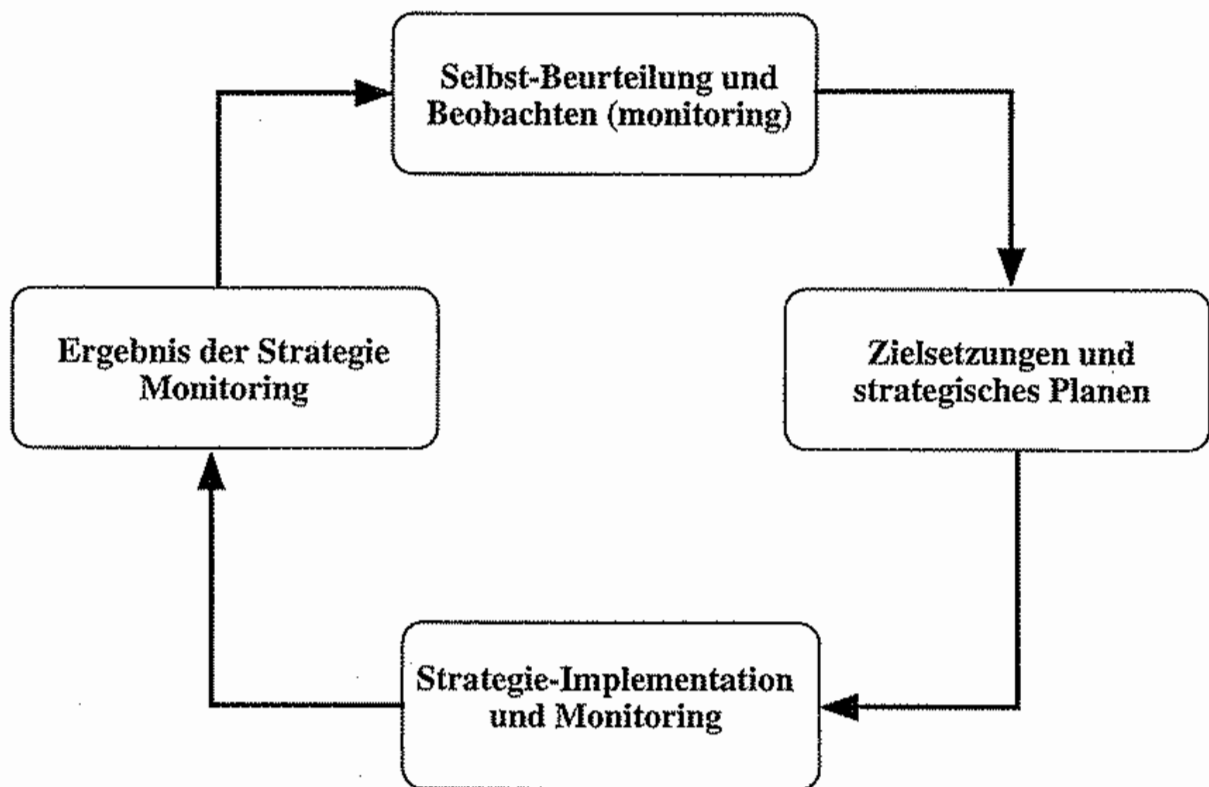


Um den »Zögling« mit Sicherheit zum Ziel zu steuern, muß der Erzieher immer wieder den erreichten Zustand feststellen, Erziehungsstrategien entwickeln und entsprechende Erziehungsmaßnahmen einleiten.

von Cube 1970, p. 82

Was hier dargestellt ist, unterscheidet sich nicht wesentlich vom Grundmodell der standardbasierten Schulreform. Offensichtlich lässt sich das Zweck-Mittel-Schema – erweitert um eine Feedback-Schlaufe – nicht nur

auf der Ebene der politischen Steuerung des Bildungssystems installieren, sondern auch auf der tiefer liegenden Ebene der pädagogischen und didaktischen „Steuerung“ des Unterrichts, ja selbst auf der Ebene der „(Selbst-)Steuerung“ der Lernprozesse. Die folgende Darstellung stammt zwar nicht aus dem Kontext der standardbasierten Schulreform, sondern aus einem psychologischen Lehrbuch, bringt aber recht anschaulich zur Darstellung, wie man sich den Zusammenhang vorstellen muss:



Seel 2000, p. 230

Doch nicht nur auf der Unterrichts- und der Individualebene, auch auf der Schulebene lässt sich das Schema anwenden, wie die Evaluationsforschung zeigt. Ich gebe Ihnen ein klassisches Beispiel: den Initialtext der schulischen Evaluationsforschung, Henry M. Levins „A Conceptual Frame-

work for Accountability in Education“ (1974). Hier finden Sie folgende Abbildung:

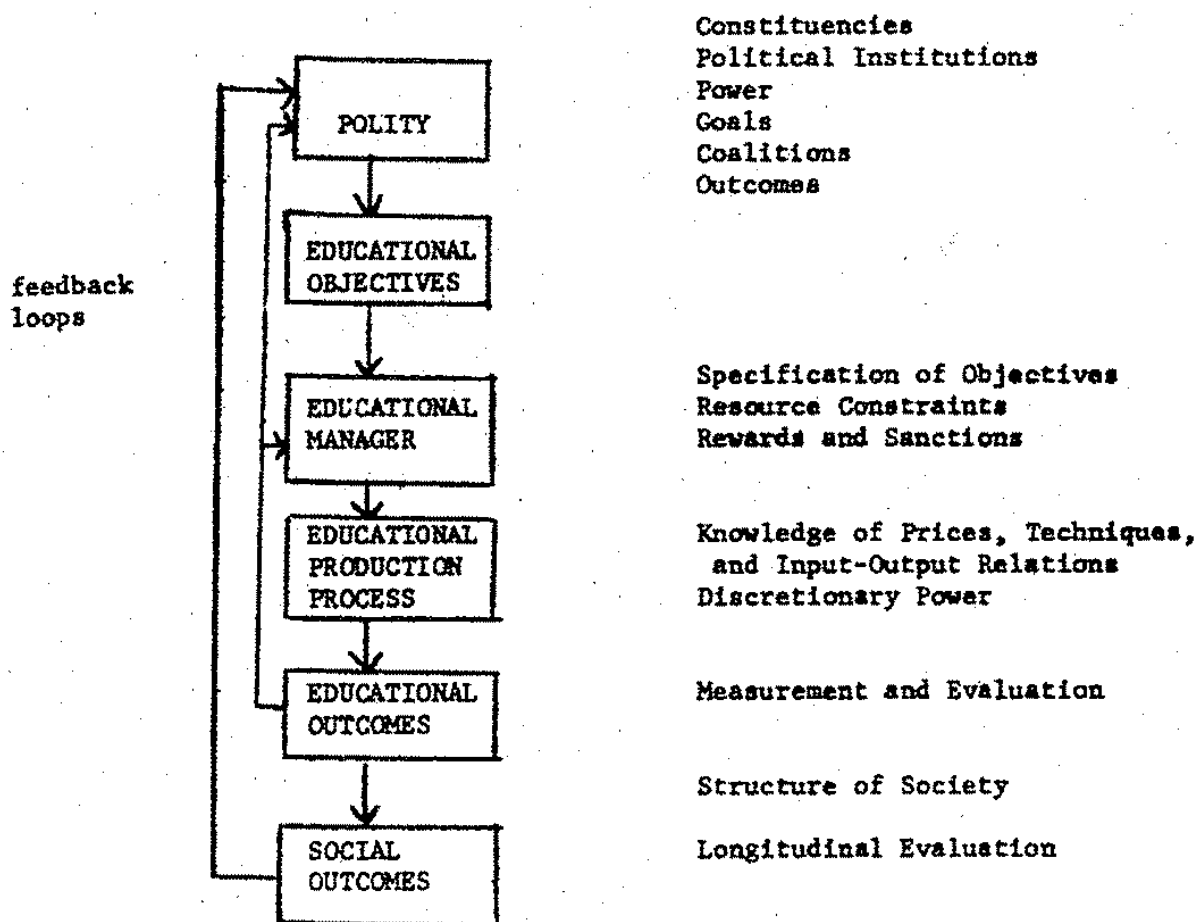


FIG. 1.—Accountability system for education

Levin 1974, p. 385

Das Schema ist etwas anders gezeichnet als bei von Cube oder Seel, bringt aber dieselbe Logik zum Ausdruck. Zudem erkennen Sie in diesem Schema, worum es m.E. bei der standardbasierten Reform im Kern geht, nämlich um die Verbindung der verschiedenen Ebenen des Schulsystems.

Wir haben es nicht einfach mit *Analogien* zu tun, sondern mit der Überzeugung, die vier Ebenen würden nach *derselben* Logik funktionieren und könnten demnach von der höchsten Ebene aus gesteuert werden. Es wird erwartet, dass sich die politische Steuerung auf den tieferen Ebenen –

Einzelnschule, Unterricht und Individuen – fortsetzt. Hier liegt m.E. der kennzeichnende Unterschied, der die standardbasierte Schulreform von früheren technokratischen Reformansätzen unterscheidet.

Wenn Sie diese Analyse akzeptieren, dann stellt sich die Frage, wie denn die unteren Ebenen des Schulsystems an die oberste Ebene angebunden werden. Womit werden die politischen Steuerungsimpulse auf die Schul-, Unterrichts- und Individualebene übertragen?

Die Antwort scheint auf den ersten Blick einfach zu sein. Es sind nämlich die Bildungsstandards, die die Verknüpfung garantieren sollen. Bildungsstandards sind so definiert, dass sie nicht nur für die Systemebene, sondern auch für die unteren Ebenen des Schulsystems gelten. In der Klieme-Expertise heißt es ausdrücklich, Standards seien „ein Kriterium für die Lernergebnisse jeder Schule, jeder Klasse und konsequenterweise jeder einzelnen Schülerin und jedes einzelnen Schülers“ (Klieme et al. 2003, p. 48).³ Demnach sind es die Standards, die als Transformationsriemen zwischen den Ebenen des Schulsystems wirken.

So richtig die Antwort ist, so wenig vermag sie zu überzeugen. Denn was genau ist der Mechanismus, der die Übertragung der Steuerungsimpulse garantiert? Aus meiner Sicht gibt es zwei Antworten auf diese Frage: eine angloamerikanische gewissermaßen und eine zentraleuropäische. Nehmen Sie diese Etiketten nicht zu eng; ich will damit lediglich andeuten, dass wir es mit zwei unterschiedlichen Ansätzen der standardbasierten Bildungsreform zu tun haben, die sich einerseits geographisch lokalisieren

³ Ähnlich heißt es in einem Aufsatz von Drüke-Noe, Keller und Blum (2008, p. 374), Bildungsstandards falle eine „Steuerungsfunktion auf systemischer Ebene, auf schulischer Ebene und auf der individuellen Ebene der Schülerinnen und Schüler zu“.

lassen, andererseits aber nicht zwingend gegenseitig auszuschließen brauchen. – Doch was sind die beiden Ansätze?

(3) Ich beginne mit dem *angloamerikanischen*. Ein Kernmoment des technologischen Modells, das der standardbasierten Schulreform zugrunde liegt, bildet der Ist-Soll-Vergleich. An diesem Vergleich entscheidet sich, ob eine Schule die gesetzten Ziele erreicht und ihre Verantwortung (*accountability*) wahrgenommen hat oder nicht. Die Soll-Komponente liegt in den Bildungsstandards, die Ist-Komponente wird mittels standardisierter Tests erfasst. Insofern liegt das Scharnier zwischen politischer und pädagogischer Steuerung der Schule in den Tests, mittels derer die Bildungsstandards überprüft werden.

Die zentrale Bedeutung der Tests für die standardbasierte Schulreform lässt sich an einer Bemerkung in der Klieme-Expertise ablesen, wonach ein „erster Höhepunkt der Arbeit mit Standards ... dann erreicht sein (wird), wenn die Schulen die Ergebnisse aus Bildungsmonitoring und Evaluation erhalten. Sie erfahren dann, inwieweit es ihnen gelungen ist, die angestrebten Kompetenzen bei ihren Schülerinnen und Schülern zu entwickeln“ (Klieme et al. 2003, p. 117).⁴

Wie sehr den Tests die gesuchte Scharnierfunktion zukommt, zeigt auch die Tatsache, dass ein Grossteil der vorliegenden Forschung zur standardbasierten Schulreform an genau diesem Punkt ansetzt, d.h. an den Tests und deren Konsequenzen für Schule und Unterricht. Analysiert werden u.a. die Qualität der Tests, deren Einfluss auf den Unterricht und deren Auswir-

⁴ Auch bei Halbheer und Reusser (2008, p. 260) heißt es, die Tests zur Überprüfung von Bildungsstandards bildeten „das Herzstück dieser Reform“ der Schule.

kungen auf den Lehrerberuf. Ich möchte auf diese drei Punkte kurz eingehen.

Qualität der Tests

Die Qualität der Tests ist vor allem dann von kritischer Bedeutung, wenn sie Auswirkungen auf der Individualebene haben, d.h. mit Konsequenzen für die Individuen – seien es die Lehrkräfte oder die Schülerinnen und Schüler – verbunden sind. In den USA nennt man dies *High-Stakes Testing* bzw. *High-Stakes Decisions*. Diese stellen ein entscheidendes Merkmal der Standard-basierten Schulreform in den angloamerikanischen Ländern dar.

Es ist notorisch strittig, ob die Qualität der verfügbaren Tests ausreichend ist, um verlässliche Entscheidungen auf der Individualebene zu treffen, die zum Beispiel die Versetzung in die nächste Klasse oder die Ausstellung eines Abschlusszeugnisses betreffen.⁵ Wo die Probleme liegen, kann eine Analyse von Popham zeigen.

Popham konstatiert zwei fundamentale Fehler („monumental mistakes“), die bei der Implementierung der standardbasierten Schulreform in den USA gemacht werden: *erstens* eine Inflation von Lernzielen (*content standards*) und *zweitens* Tests, die unterrichtsinsensitiv (*instructionally insensitive*) sind. Die beiden Fehler werden dazu führen, dass eine Verbesserung des Unterrichts nicht eintreten und die Reform als Ganze scheitern wird (vgl. Popham 2004, p. 15f.).

⁵ In der Klieme-Expertise heißt es (mit Bezug auf die USA) wohl zu Recht, „zu häufiges Testen und zu oberflächliches Messen“ könnten „negative Folgen für die Qualität des Unterrichts haben“ (Klieme et al. 2003, p. 103). Zudem wird von der Verwendung der Tests für individualdiagnostische Urteile explizit abgeraten (vgl. ebd., p. 10, 49, 107ff.). Selbst hinsichtlich der Bildungsstandards wird davor gewarnt, diese „als Kriterien für Notengebung, Zertifizierung und Selektionsentscheidungen zu missbrauchen“ (ebd., p. 48).

Was den ersten Fehler anbelangt, so kritisiert Popham, dass man den Fachlehrkräften überlassen habe, die curricularen Standards zu entwickeln. Da jeder an sein eigenes Fach gedacht und niemand eine übergeordnete Perspektive eingenommen habe, seien Listen mit Hunderten von Zielen entstanden (vgl. Popham 2004, p. 18ff.) – viel zu viele, als dass sie in der verfügbaren Zeit unterrichtet, geschweige denn getestet werden könnten.

Was den zweiten Fehler anbelangt, so würden zwei Arten von Tests verwendet: auf der einen Seite normorientierte Leistungstests, die zwar geeignet sind, Individuen nach ihrer Leistungsfähigkeit zu unterscheiden, aber nicht dafür konstruiert sind, um Lernleistungen nach Zielkriterien zu beurteilen, und auf der anderen Seite standardbasierte Tests, die zwar speziell für die Aufgabe der Prüfung von Schulleistungen entwickelt wurden, aber aufgrund der Inflation von curricularen Standards bestenfalls *einen Teil* des Unterrichts abdecken und schlimmstenfalls Stoff prüfen, *der gar nicht unterrichtet wurde*. Das ist *dann* kein Problem, wenn die Messung von Standards lediglich eine Monitoringfunktion hat. Wenn die Tests aber Konsequenzen auf der Individualebene haben oder der Lehrkraft ein Feedback zu ihrem Unterricht geben sollen, dann sind sie völlig ungeeignet. „To use students’ performances on instructionally insensitive tests to identify instructionally effective (or ineffective) schools is every bit as silly as it sounds. Yet that is precisely what is going on in most parts of our nation“ (Popham 2004, p. 15).⁶

⁶ Aus meiner Sicht besteht hier auch ein ethisches Problem. Die Tests, mit denen Standards überprüft werden, müssen valide sein, d.h. sie müssen das messen, was im Unterricht vermittelt wurde. Wenn sie diese curriculare Validität nicht haben (wie auch die PISA-Tests), ist unklar, ob sie den intendierten Nutzen überhaupt erbringen können. Die Situation ist umso gravierender, wenn anstelle von Tests, die spezifisch für die Überprüfung von Bildungsstandards entwickelt wurden, handelsübliche kommerzielle Tests verwendet werden, die normorientiert sind. Mittels dieser Tests lässt sich schlicht nicht sagen, ob

Es bestehen weitere Probleme mit den *High-Stakes Tests*, auf die ich hier aber nicht näher eingehen kann.⁷ Die Probleme haben zur Folge, dass die Lehrkräfte die Tests und Testergebnisse oft als wenig hilfreich empfinden für Entscheidungen, die ihren Unterricht betreffen (Smith 1991; Ingram, Louis & Schroeder 2004). Zudem fühlen sie sich durch die Tests unter Druck gesetzt, was sich negativ auf den Unterricht auswirkt.

Einfluss auf den Unterricht

Es braucht wenig Fantasie, um sich auszumalen, dass Lehrkräfte ihren Unterricht unter Bedingungen eines *High-Stakes Testing* an den Tests ausrichten. Und dies in mehrerer Hinsicht. Erstens *curricular*: Sie treffen die Stoffauswahl so, dass der Unterricht mit dem übereinstimmt, was durch die Tests abgeprüft wird. Zweitens *didaktisch*: Sie orientieren sich am Testformat und richten den Unterricht formal an den *Paper-and-Pencil-* und *Multiple-Choice-Tests* aus. Drittens *methodisch*: Sie verengen ihre Methodik auf traditionelle Lehrformen („*direct instruction*“) und beschränken sich auf das Einpauken von Wissen. Alle drei Reaktionen lassen sich aufgrund verschiedener Studien belegen (vgl. Darling-Hammond 2004; Diamond & Spillane 2004; Dorgan 2004; Ellis 2008; Kantor & Lowe 2006; Herman 2004; Linn 2000, 2003; Mathison & Freeman 2004; McCloskey & McMunn

der Unterricht gut oder schlecht war bzw. ob die erzielten Ergebnisse der Schülerinnen und Schüler auf den Unterricht oder auf etwas anderes zurückzuführen sind.

⁷ Die teilweise gravierenden Mängel der Tests, mit denen Standards überprüft werden, haben verschiedene professionelle Organisationen und Berufsverbände in den USA veranlasst, vor einem einseitigen Gebrauch der Tests bzw. Testergebnisse zu warnen (vgl. AERA 2000; Mathison 2004, p. 3). Gewarnt wird vor allem vor der Verwendung einzelner Tests für folgenreiche Entscheidungen (High-Stakes Decisions). Vgl. die oft zitierte Äusserung von Heubert und Hauser (1999, p.15): „An educational decision that will have a major impact on a test taker should not be made solely or automatically on the basis on a single test score.“

2000; McNeil 2000; Rex & Nelson 2004, p. 1292ff.; Rowe 2000, p. 75f.; Sandholtz et al. 2004, p. 1188 ff.; Smith 1991; Stecher 2002).

Wirklich bedeutsam als Determinanten des Unterrichts – und das ist eine eher beunruhigende Feststellung – sind demnach nicht die Bildungsstandards, sondern die *Tests*, mittels derer die Standards überprüft werden. Dafür steht der etwas schillernde Begriff des *Teaching-to-the-Test*. Der Unterricht wird verengt auf jene Fächer und Themen, die getestet werden. Aufgegeben werden Vertiefungsthemen, Projekte, Museumsbesuche, Gruppenunterricht, Exkursionen etc. Kenneth Rowe (2000, p. 76) spricht (in Bezug auf Großbritannien) von einem „test-dominated curriculum“. Andernorts ist von „measurement-driven instruction“ (Popham 1987), „data-driven teaching“ (Ingram et al. 2004, p. 1281) oder „test-driven instruction“ die Rede.

Auch hier kann ich aus zeitlichen Gründen auf weitere Details nicht eingehen. Erwähnt sei lediglich, dass in verschiedenen Arbeiten vom Ausschluss leistungsschwacher Schülerinnen und Schüler vom Unterricht, ihrer Überweisung in Sonderklassen, von der selektiven Förderung von Schülerinnen und Schülern, der Anpassung des Leistungsniveaus des Unterrichts an die Standards, der Vorbereitung der Schülerinnen und Schüler auf die Testsituation, vom Fernhalten schwacher Schülerinnen und Schüler von den Leistungstests und von Betrugereien bei der Testdurchführung berichtet wird (vgl. Allington & McGill-Franzen 1992; Booker-Jennings 2005; Brown 2001, p. 376f.; Cala 2004, p. 152; Cizek 2001; Darling-Hammond 2004, p. 157ff.; Diamond & Spillane 2004, p. 1155, 1166f.; Dorgan 2004, p. 1218, 1223f.; Ellis 2008, p. 1144; Mathison & Freeman 2004, p. 87; Nichols & Berliner 2005; Rennert-Ariev 2008, p. 129ff.; Rowe 2000, p. 76).

Dass all dies auch Auswirkungen auf den Lehrerberuf hat, kann nicht erstaunen.

Auswirkungen auf den Lehrerberuf

In einer Studie im Staate Texas hat Linda McNeill (2000) festgestellt, dass sich die Lehrkräfte von den Standards und den Tests zur Überprüfung der Standards bedrängt fühlen. Sie empfinden eine Art schizophrene Spaltung zwischen dem, was sie für die Tests tun, und dem, was sie für die Schülerinnen und Schüler tun. In ihrem professionellen Selbstverständnis herausgefordert, sehen sie sich genötigt, „[to] choose between their personal survival in the system or their students' education“ (McNeil 2000, p. 192).

In einem vergleichbaren Dilemma befinden sich die von Sandra Mathison und Melissa Freeman (2004) untersuchten Lehrkräfte im Staat New York. Sie schwanken zwischen dem Anspruch auf Individualisierung des Unterrichts, an dem viele Lehrkräfte festhalten möchten, und der Notwendigkeit, den normativen Vorgaben auf Standardisierung der Schülerleistungen zu entsprechen (vgl. Mathison & Freeman 2004, p. 89).⁸

Viele Lehrkräfte fühlen sich in ihrer Professionalität bedrängt (vgl. Kanna-pel, Coe, Aagard, Moore & Reeves 2000; Rex & Nelson 2004, Smith 1991). Die Administration externer Tests empfinden sie als Misstrauensbekundung gegenüber ihrer beruflichen Kompetenz (vgl. Hill 2004, p. 1120; Meier 2004, p. 71ff.). Nicht nur Linda McNeill, Sandra Mathison und Melissa Freeman, auch eine Reihe weiterer Autorinnen und Autoren sprechen von einer *Deprofessionalisierung* des Lehrerberufs durch die standard-

⁸ Es ist auffällig, wie in vielen Texten aus den USA, aber auch aus anderen Ländern von Dilemmata und Widersprüchen der Standard-basierten Schulreform die Rede ist, Dilemmata und Widersprüche, von denen vor allem die Lehrkräfte betroffen sind (vgl. z.B. Gamoran 2007; McNeil 2000, Rowe 2000, p. 77; Sandholtz et al. 2004, p. 1200).

basierte Schulreform. Die Arbeit der Lehrkräfte wird auf Technologie reduziert. „Far from the reflective practitioner or the empowered teacher, those optimistic images of the 1980s, the image we project of teachers in the world after testing reform is that of interchangeable technicians receiving the standard curriculum from above, transmitting it as given (the presentation manual never leaving the crook of their arms), and correcting multiple-choice responses of their pupils“ (Smith 1991, p. 11).

(4) In *High-Stakes Tests* liegt *eine* Antwort auf die Frage nach dem Scharnier, das die politische mit der pädagogischen Ebene des Schulsystems im Rahmen der standardbasierten Schulreform verbindet. Es ist dies die angloamerikanische Antwort. Ich habe jedoch von zwei Antworten gesprochen. Dieser zweiten, *zentraleuropäischen* Antwort möchte ich nun nachgehen.

Tests müssen nicht zwingend mit *High-Stakes* verbunden sein; sie müssen nicht zwingend dazu dienen, Entscheidungen zu treffen, die für die Betroffenen weit gehende Konsequenzen haben. Insofern kann man sich eine standardbasierte Reform vorstellen, die ohne *High-Stakes Testing* auskommt.⁹ Und es scheint, dass wir uns in Deutschland, Österreich und der Schweiz – zumindest vorläufig – in dieser Situation befinden.

Was die Schweiz anbelangt, so lässt man sich von Seiten der EDK dahingehend vernehmen, dass die Bildungsstandards weder direkte Konsequenzen für die Schülerinnen und Schüler noch für die Lehrkräfte haben sollen. Es gehe nicht darum, „die Ergebnisse des Projekts HarmoS [wie

⁹ In diesem Sinne unterscheidet Scott Thompson (2001) eine gute von einer schlechten Standardbewegung (*standards movement*). Die gute orientiert sich an „authentischen“ Tests, während die schlechte Tests benutzt, um Druck auf die Schulen auszuüben.

das Reformprojekt in der Schweiz heißt, W.H.] für die Evaluation der Lehrpersonen oder die Beurteilung und Selektion der Schülerinnen und Schüler zu nutzen, sondern [nur] um die Steuerung des Schulsystems“ (EDK 2004, p. 13). Ich bin offen gestanden skeptisch, ob wir uns auf diese Verlautbarungen verlassen können. Zu Vieles weist in eine andere Richtung (vgl. Herzog 2008a, 2008b, 2008c).

Trotzdem, die standardbasierte Schulreform setzt in unseren Ländern einen anderen Akzent als in den USA, Großbritannien oder Australien. Statt auf *High-Stakes Testing* setzt man auf Kompetenzen. Genauer gesagt, sind es nicht Kompetenzen, sondern *Kompetenzmodelle*, die in Deutschland, Österreich und der Schweiz als Scharnier fungieren, um die Regelkreise der Politik und der Pädagogik miteinander zu verknüpfen.

Die Klieme-Expertise weist den Kompetenzmodellen die Aufgabe zu, „... die Ziele, die Struktur und die Ergebnisse fachlicher Lernprozesse zu beschreiben. Sie bilden die Komponenten und *Stufen der Kompetenzentwicklung* von Schülerinnen und Schülern ab und bieten somit eine Orientierung für schulisches Lehren und Lernen“ (Klieme et al. 2003, p. 135 – Hervorhebung W.H.). Erst die Kompetenzmodelle „... geben den Standards eine Orientierungskraft für den Unterricht, indem sie *unmittelbar einsichtig* und nachvollziehbar, illustriert an *konkreten Anforderungen*, demonstrieren, welche *Entwicklungs- und Niveaustufen* fachliche Kompetenzen haben“ (ebd. – Hervorhebungen W.H.).

Die beiden Zitate, die kurz aufeinander folgen, sind in doppelter Weise interessant. Erstens steht nicht wie bei den *High-Stakes Tests* die Outputmessung im Vordergrund, sondern die Darstellung der (fachlichen) Lernprozesse. Die Kompetenzmodelle sollen „kumulatives Lernen über mehrere Jahrgänge und Niveaustufen hinweg darstellen“ (Klieme et al. 2003, p. 50) und damit den nach Stufen unterscheidbaren *Prozess des Erwerbs von*

Kompetenzen abbilden. Explizit heißt es in der Klieme-Expertise, die den Bildungsstandards zu Grunde liegenden Kompetenzmodelle würden die „*Stufen der Kompetenzentwicklung* von Schülerinnen und Schülern [in einem Fach, W.H.] ab(bilden)“ (ebd., p. 135 – Hervorhebung W.H.). Und: Indem sie der „Logik des Wissenserwerbs“ (ebd., p. 75) folgen, würden sie die „Schritte beim Erwerb von Kompetenzen“ (ebd., p. 77), ja deren „Entwicklungslogik“ (ebd., p. 136) entfalten.¹⁰

Damit sind wir beim zweiten Moment, das die Kompetenzmodelle so interessant macht. Von den Kompetenzmodellen wird nämlich nicht nur erwartet, dass sie den Prozess des fachlichen Lernens abbilden und damit die Aussicht auf ein Curriculum im strengen Sinn eröffnen (vgl. Klieme et al. 2003, p. 50). Vielmehr sollen sie die Stufen des fachlichen Lernens konkret *beschreiben* und *anschaulich* darstellen.

Es wäre eine reizvolle Aufgabe, die Klieme-Expertise nach der Verwendung von Wörtern wie „klar“ und „konkret“ zu durchforsten. Wie zentral diese Wörter sind, möchte ich anhand einiger Beispielen zeigen. Bildungsstandards bestimmen „konkrete Kompetenzanforderungen“, heißt es an einer Stelle (Klieme et al. 2003, p. 99). „Sie arbeiten in klarer und konzentrierter Form heraus, worauf es in unserem Schulsystem ankommt“ (ebd., p. 47 – im Original ganzer Satz hervorgehoben). Indem „klare und verbindliche Erwartungen“ (ebd., p. 48) in die Bildungsstandards eingehen, erhalte der Unterricht „klare Zielstellungen“ (ebd., p. 51) und lasse sich „klar fokussieren“ (ebd., p. 50). Dadurch würden die Ziele und Anforderungen der Schule „transparent“ (ebd., p. 48). Indem die Kompetenzen und deren Stu-

¹⁰ Ähnlich geht man im schweizerischen HarmoS-Projekt davon aus, die den Bildungsstandards zu Grunde liegenden Kompetenzmodelle würden die „Entwicklungsverläufe von Kompetenzen *sichtbar* machen“ (Maradan & Mangold 2005, p. 4 – Hervorhebung W.H.).

fen „konkret beschrieben“ (ebd., p. 19) würden, werde „*sichtbar*, in welcher Richtung und auf welchem Weg sie [die Schülerinnen und Schüler, W.H.] gefördert werden können“ (ebd., p. 135 – Hervorhebung W.H.). Etc. Die Beispiele ließen sich leicht vermehren.¹¹

Was hier aufscheint, ist ein „Positivismus höherer Ordnung“ (um einen Ausdruck von Helmuth Plessner zu gebrauchen). Als ob Kompetenzen in irgendeiner Hinsicht anschaulich sein könnten.¹² Kompetenzen sind theoretische Begriffe, „hypothetische Konstrukte“, die sich weder einfach *beschreiben*, noch irgendwie *sichtbar* machen lassen.¹³ Dass uns die Kompetenzstufen jemals anschaulich vor Augen liegen und die Lehrkräfte ihren Unterricht „kompetenzbasiert“ an konkreten Entwicklungsstufen ausrichten könnten, halte ich für pure Ideologie.

¹¹ „Klarheit“, „Konkretheit“ und „Transparenz“ sind Ausdrücke, die im Kontext der Standard-basierten Schulreform, jedenfalls im deutschsprachigen Raum, häufig verwendet werden. Um noch ein Beispiel zu geben: „Bildungsstandards sollten ... das tun, wofür sie entwickelt worden sind, nämlich Klarheit geben, worauf es in der Schule ankommt“ (Specht & Lucyshyn 2008, p. 324).

¹² Man denke an die Definition von Weinert, an der sich die Klieme-Expertise genauso wie das HarmoS-Projekt orientiert. Danach sind Kompetenzen „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert, zit. nach Klieme et al. 2003, p. 72).

¹³ Bei der EDK heißt es, Bildungsstandards würden sich „auf die genaue [!] Beschreibung [!] der aufeinander folgenden Kompetenzniveaus [in einem Fachbereich, W.H.] stützen“ (EDK 2006, p. 23) bzw. auf „einer genauen [!] Beschreibung[!] von skalierten Kompetenzniveaus (beruhen)“ (ebd., p. 25). Sie würden „genau [!] beschreiben [!], welche Basiskompetenzen alle Schülerinnen und Schüler ... erworben haben sollen“ (ebd.). Dank der „Genauigkeit [!] bei der Beschreibung [!] der Kompetenzniveaus und der erwarteten Lernfortschritte (ebd., p. 26) werde es zudem möglich sein, „die Schülerbeurteilung zu verbessern“ (ebd.).

Dass trotzdem unbeirrt und fast hektisch an Kompetenzmodellen gearbeitet wird, erklärt sich m.E. aufgrund der Funktion, die ihnen zukommt. Die Kompetenzmodelle sollen jene Leistung erbringen, die in den USA den *High-Stakes Tests* aufgetragen ist: Sie sollen die Verknüpfung der politischen mit der pädagogischen Regulierungsschleife garantieren, da die standardbasierte Schulreform nur so Aussicht hat, auf der Unterrichtsebene überhaupt anzukommen.

Indem Kompetenzmodelle die „Stufen der Kompetenzentwicklung“ (Klieme et al. 2003, p. 135) abbilden, haben sie eine *curriculare* Bedeutung (und nicht eine Assessment-Funktion wie die *High-Stakes Tests*). Sie schreiben vor, und zwar ziemlich streng, *wie unterrichtet werden soll*. Dadurch vermögen sie einen ähnlichen Zwang auf den Unterricht auszuüben wie die *High-Stakes Tests*. Wo alles klar und konkret ist und anschaulich vor Augen liegt, was zu tun ist, da gibt es keinen Spielraum für Interpretationen und Artikulationen des Unterrichts.¹⁴ Wie die Lehrkraft zu unterrichten hat, schreiben ihr die Kompetenzmodelle klar und unmissverständlich vor! Dadurch wird Machbarkeit suggeriert, die eine vergleichbare Wirkkraft verspricht wie die *High-Stakes Tests* in den USA.

(5) Damit komme ich zum Schluss meiner Ausführungen, die ich nicht beenden möchte, ohne meine Argumentation noch etwas zuzuspitzen. Den Kern und zugleich das Neue der standardbasierten Schulreform gegenüber früheren Ansätzen einer zweckrationalen Erneuerung der Schule sehe ich im Anspruch, alle Ebenen des Bildungssystems nach derselben techni-

¹⁴ Insofern bezweifle ich, dass die in der Klieme-Expertise mehrfach anzutreffende Behauptung, wonach Bildungsstandards den Schulen und Lehrkräften mehr Freiräume bei der Unterrichtsgestaltung bringen würden (vgl. Klieme et al. 2003, p. 9, 123, 139), zutreffend ist.

schen Logik zu steuern, indem ihre Regulierungsschlaufen miteinander verknüpft werden. Dabei fungieren Bildungsstandards als Kommunikationsmedium, um die Ebenen sprachlich aufeinander abzustimmen. Erzwungen wird die Verknüpfung aber nicht mit den Bildungsstandards als solchen, sondern auf der einen, angloamerikanischen Seite mittels *High-Stakes Tests*, die für die Einzelschulen und die Individuen (Lehrkräfte sowie Schülerinnen und Schüler) gravierende Konsequenzen haben können, und auf der anderen, zentraleuropäischen Seite mittels *Kompetenzmodellen*, denen eine curricular normierende Funktion zugewiesen wird.

Die angloamerikanische Variante eines Transformationsmechanismus zwischen Politik und Pädagogik ist – wie die Erfahrungen aus den USA zeigen – nicht sonderlich empfehlenswert, da sie zu drastischen Einengungen und Verhärtungen des Unterrichts führt. Die zentraleuropäische Variante ist vielleicht empfehlenswerter. Doch ist sie mit dermaßen hohen Ansprüchen an die Entwicklung von Kompetenzmodellen verbunden, dass sich die Frage aufdrängt, ob wir solche Kompetenzmodelle jemals haben werden.

Falls wir sie *nicht* haben sollten, stellt sich die Frage, ob uns nicht – und vielleicht schneller als uns lieb sein könnte – amerikanische Verhältnisse drohen.

Anmerkung 2015: Für den Literaturnachweis verweise ich auf andere Referate sowie Publikationen zum Thema der standardbasierten Schulreform. Das vorliegende Referat ist in gekürzter Form erschienen: Walter Herzog (2010). Besserer Unterricht dank Bildungsstandards und Kompetenzmodellen? In Gehrmann, Axel; Hericks, Uwe; Lüders, Manfred (Eds.), *Bildungsstandards und Kompetenzmodelle. Eine Verbesserung der Qualität von Schule, Unterricht und Lehrerbildung* (S. 37-46). Bad Heilbrunn: Klinkhardt.