

Hans-Jürgen Kuhn

Anspruch, Wirklichkeit und Perspektiven der Gesamtstrategie der KMK zum Bildungsmonitoring

Zusammenfassung

Die 2006 von der KMK beschlossene Gesamtstrategie beschreibt wesentliche Elemente einer auf Evidenz angelegten Bildungspolitik, die auf den 2004/05 beschlossenen bundesweiten Bildungsstandards und der darauf basierenden regelmäßigen Überprüfung des Erreichens dieser Standards aufbaut. Zentrale Bestandteile dieser Strategie werden immer wieder mit unterschiedlichen Argumenten kritisiert. Dabei steht vor allem die Frage nach der Nützlichkeit der erhobenen Befunde von Large-Scale-Assessments wie PISA und TIMSS, aber auch von Ländervergleichen des IQB und von der Durchführung der bundesweiten Vergleichsarbeiten im Mittelpunkt. Der Beitrag analysiert die unterschiedlichen Argumentationsmuster und geht auf von den Kritikern und Kritikerinnen beklagte Akzeptanz-, Transparenz- und Transferdefizite dieser Strategie ein. Für den aktuellen Diskussionsprozess um die Weiterentwicklung der Gesamtstrategie werden zum Abschluss einige Vorschläge entwickelt.

Schlüsselwörter: Bildungsstandards, Gesamtstrategie der KMK, Ländervergleiche, Transferprozesse

Claim, Reality and Perspectives of the Overall Strategy of the KMK

Abstract

The overall strategy, which was enacted by the KMK in 2006, describes essential issues of an evidence-based educational policy, based on the nationwide educational standards enacted in 2004/05 and the hence continuous testing of the achievement of these standards. Central elements of this strategy have been frequently criticized with different arguments. They focus mainly on the question of usefulness of the results of large-scale assessments like PISA and TIMSS, but also of the comparisons across the federal states by the Institute for Educational Quality Improvement (IQB) and the implementation of nationwide comparative testing. The article analyzes the different argumentation patterns and deals with the shortcomings regarding acceptance, transparency and transfer of this strategy, which have been deplored by the critics. Finally, it develops some proposals regarding the topical discussion process about the further development of the overall strategy.

Keywords: educational standards, overall strategy of the KMK, comparisons across the federal states, transfer processes

1. Einleitung

Mit dem sogenannten Konstanzer Beschluss von 1997, der anschließenden Teilnahme an internationalen Schulleistungsstudien wie PISA, TIMSS, PIRLS, der Verabschiedung von nationalen Bildungsstandards für ausgewählte Fächer und der Gründung einer neuen gemeinsamen Ländereinrichtung, dem Institut für Qualitätsentwicklung im Bildungswesen (IQB), hat die KMK in ungewöhnlich schneller Weise auf die seit dem Ende der 1990er-Jahre offenkundigen Leistungsprobleme im Bildungssystem reagiert. Die Klärung von Ursachen für unbefriedigende (Leistungs-)Ergebnisse wurde damit als eine neue Aufgabe der KMK definiert, die vier zentrale Elemente umfasst (vgl. KMK 2006, S. 1):

- Internationale Schulleistungsstudien;
- Zentrale Überprüfung des Erreichens der Bildungsstandards in einem Ländervergleich;
- Vergleichsarbeiten in Anbindung oder Ankopplung an die Bildungsstandards zur landesweiten Überprüfung der Leistungsfähigkeit einzelner Schulen;
- Gemeinsame Bildungsberichterstattung von Bund und Ländern.

Um die davon erwarteten Impulse und positiven Wirkungen auch erreichen zu können, wurde am 2. Juni 2006 die „Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring“ verabschiedet, mit der u.a. auch eine konzeptionelle Verbindung der unterschiedlichen Ebenen geleistet werden sollte. Obwohl festgestellt wurde, dass „Verfahren des Bildungsmonitoring mit Maßnahmen zur Unterrichts- und Qualitätsentwicklung sowie Unterstützung der Schulen verknüpft werden müssen“ (KMK 2006, S. 2), bleibt der gesamte Beschlusstext in diesen Punkten jedoch eher vage und unverbindlich.

Ausgehend von den seit nunmehr zehn Jahren geltenden Bildungsstandards als verbindlichem Referenzrahmen für das Bildungsmonitoring in den Ländern, für die Ländervergleiche des IQB und für die VERA-Vergleichsarbeiten wurden mit der Gesamtstrategie Hoffnungen verknüpft, die aus heutiger Sicht bislang nur unzureichend eingelöst worden sind:

„Schulen können die Ergebnisse ihrer Schülerinnen und Schüler auf mehreren Ebenen vergleichen und erhalten vielfältige Informationen, die sie für ihre pädagogische Arbeit unmittelbar nutzen können. Dadurch wird die angestrebte Verknüpfung zwischen Leistungsüberprüfungen und Maßnahmen der Schulentwicklung entscheidend verbessert sowie die Umsetzung der mit den Bildungsstandards verbundenen Ziele an den Schulen befördert.“ (Ebd., S. 11)

Der folgende Diskussionsbeitrag setzt sich mit Schwierigkeiten und Problemen im Bereich der ersten drei Säulen (Large-Scale-Assessment-Studien (LSA), Ländervergleiche und VERA) auseinander. Dabei werden Vorschläge für die seit 2013 stattfindende Debatte zur Veränderung der Gesamtstrategie entwickelt. Auf die Probleme der nationalen Bildungsberichterstattung als vierter Säule wird an dieser Stelle nicht eingegangen.

Nach fünf PISA-Durchgängen, drei Ländervergleichen durch das IQB und zahlreichen Durchgängen von VERA 3 und 8 sind mindestens drei Problembereiche identifizierbar: Die Gesamtstrategie hat ein Akzeptanzdefizit, ein Transparenzdefizit und ein Transferdefizit.

2. Das Akzeptanzdefizit ...

... bei den Large-Scale Assessments

Sowohl die großen Large-Scale-Assessment-Studien wie TIMSS, PISA, IGLU und andere als auch die bundesweiten Vergleichsarbeiten (VERA) sind von Beginn an kritisch kommentiert worden. Dabei bezog sich die Kritik je nach Akteur und Instrument auf forschungstheoretische, bildungstheoretische und pädagogische Fragen und Implikationen. Die Kritik ist nie abgerissen, sie hat in den letzten Jahren eher zugenommen. Die aktuelle Debatte lässt sich gut durch zwei öffentliche Statements beschreiben:

Am 4. Mai 2014 veröffentlichte Heinz-Dieter Meyer zusammen mit hunderten von Wissenschaftlerinnen und Wissenschaftlern sowie Pädagoginnen und Pädagogen (auch aus Deutschland) einen Offenen Brief an Andreas Schleicher, in dem eine fundamentale Kritik am PISA-Programm der letzten zwölf Jahre geübt wurde. Verbunden mit Hinweisen und Vorschlägen, was man ändern sollte, wurde ein Moratorium gefordert.

Unabhängig davon stellten die GEW, der Grundschulverband, der VBE und der Bayerische LehrerInnenverband am 4. Mai 2014 ein Manifest vor, in dem sie ihre Skepsis bzw. Ablehnung der flächendeckenden Vergleichsarbeiten formulierten und ebenfalls erhebliche Veränderungen an der bisherigen Strategie der KMK bzw. am Vorgehen der Länder forderten.

Damit werden zwei zentrale Säulen der KMK-Gesamtstrategie acht Jahre nach ihrer Einführung schon wieder (LSA-Studien) oder immer noch (VERA) grundsätzlich in Frage gestellt. Das Diskursklima in der Wissenschaft, in der Bildungspolitik und Bildungsadministration und auf Seiten der Schulpraxis hat sich gegenüber posi-

tiven Erwartungen der früheren Jahre deutlich verändert. Die Debatte über den Sinn und Zweck dieser Art von Forschung bzw. wissenschaftlichen Dienstleistung und vor allem über die impliziten und expliziten Versprechen der Gesamtstrategie der KMK wird auf ihre Weiterentwicklung sicherlich Einfluss nehmen.

Insgesamt kann festgestellt werden, dass der Anspruch der KMK,

„[d]ie Konzeption der Testverfahren zur Überprüfung des Erreichens der Bildungsstandards so zu gestalten, dass den Ländern *vor allem qualitatives Wissen* für gezielte Interventionsmöglichkeiten zur Weiterentwicklung des Bildungswesens zur Verfügung gestellt wird“ (KMK 2006, S. 13; Herv. im Original),

nur teilweise eingelöst werden konnte.

Unrealistische Erwartungen an die Bildungsforschung seitens der Politik und Selbstüberschätzungen bei Teilen der empirischen Bildungsforscherinnen und -forscher haben zu Enttäuschungen auf den verschiedenen Ebenen geführt. Zweifel an der Relevanz der empirischen Bildungsforschung zur Verbesserung von Unterricht und Schule mehren sich.

Dieser schleichende Relevanzverlust (und damit auch Akzeptanzverlust) speist sich aus unterschiedlichen Quellen. Zum einen kommt es nach den großen, medial inszenierten Aufregungen durch die Vorstellung der jeweiligen PISA-Ergebnisse, TIMS-Studien, IGLU-Studien und Ländervergleiche inzwischen zu einer Abstumpfung, weil sich die zentralen Ergebnisfiguren wiederholen. Zum anderen hat die Menge der Publikationen, die durch Zweit- und Drittverwertung durch Stiftungen und sonstige Institutionen, die stets mit denselben Primärdaten arbeiten, noch deutlich anwächst, zu einer riesigen Unübersichtlichkeit geführt. Welcher Befund ist neu? Welcher Befund ist wichtig? Welche Aussagen halten einer seriösen Überprüfung stand?

Ein in diesem Zusammenhang immer wieder erhobener Vorwurf lautet, durch PISA werde ein reduzierter und verkürzter Bildungsbegriff einen immer größeren Raum einnehmen. Kompetenzorientierung und für Messungen handhabbare Kompetenzstufenmodelle dominierten die pädagogischen und curricularen Debatten – so die Kritik. Versuche, dieser Verengung durch einen umfassenderen Bildungsbegriff zu entgehen, werden schnell als vermutete Absage an jede Form evidenzbasierter Reformdiskussion kritisiert. Dabei wäre es sehr lohnend, sich wieder stärker auf einen Allgemeinbildungsdiskurs einzulassen. Vor allem, nachdem die Schülerinnen und Schüler in Deutschland in allen drei untersuchten Domänen Leistungen oberhalb des OECD-Mittelwertes erzielen, könnte in dieser Frage eine gewisse Entspannung einsetzen. In enger Anlehnung an Klafki sollte man beispielsweise intensiver über epochaltypische Schlüsselprobleme unserer Gegenwart und der vermutlichen Zukunft sprechen und sich fragen, wie es um das Erreichen des zentralen Ziels der Schulen steht, der nächsten Generation Einsicht in die Mitverantwortlichkeit

für die Herausforderungen der Zukunft zu vermitteln und die Bereitschaft zu fördern, an ihrer Bewältigung mitzuwirken (vgl. Klafki 1996, S. 56).

Dass diese Diskussion keine Alternative zur dominanten Kompetenzorientierung sein muss, zeigt ein Zitat von Jürgen Baumert aus dem ersten PISA-Ergebnisband von 2000:

„Man kann gar nicht ausdrücklich genug betonen, dass PISA keineswegs beabsichtigt, den Horizont moderner Allgemeinbildung zu vermessen, oder auch nur die Umrisse eines internationalen Kerncurriculums nachzuzeichnen. Es ist gerade die Stärke von PISA, sich solchen Allmachtsphantasien zu verweigern.“ (Baumert/Stanat/Demmrich 2001, S. 21)

... bei den Vergleichsarbeiten

Anders als bei der Kritik an den LSAs und Ländervergleichen kommt die Kritik an VERA vor allem von denen, für die die Vergleichsarbeiten eine Hilfe sein sollten. Trotz zahlreicher Aktivitäten der Länder ist die Lehrerschaft in den Schulen bezüglich der Nützlichkeit der flächendeckenden Vergleichsarbeiten gespalten: Während nach Aussagen aktueller Studien (vgl. Richter u.a. 2014) die eine Hälfte dem Vorhaben eher aufgeschlossen und positiv gegenübersteht, ist die andere Hälfte auch zehn Jahre nach der Einführung den Vergleichsarbeiten gegenüber eher skeptisch bis ablehnend eingestellt.

Im Unterschied zu sonstigen Schulleistungstests besteht der besondere Stellenwert von VERA vor allem darin, dass Lehrkräfte die Leistungen ihrer Schülerinnen und Schüler in zentralen Unterrichtsfächern an übergreifend geltenden, objektiven Leistungsmaßstäben – den bundesweit geltenden Bildungsstandards – überprüfen können. VERA richtet seinen Focus auf die Ebene jeder einzelnen Schule und Klasse und trägt auf diese Weise zur Unterstützung der Unterrichts- und Schulentwicklung vor Ort bei – so der Anspruch.

Tatsächlich sind einige der bildungspolitisch Verantwortlichen in den Ländern aus unterschiedlichen Gründen nicht bereit, VERA auch so zu kommunizieren. Je nach Fach werden in sechs bis sieben Ländern den Schulen lediglich die erreichten Lösungshäufigkeiten zu den eingesetzten Testheften mitgeteilt. Die Rückmeldung der Ergebnisse auf der Basis der von der KMK verabschiedeten Kompetenzstufenmodelle (KSM) wird ausdrücklich nicht vollzogen; die ursprünglich gedachte Funktion der VERA-Rückmeldung als Frühwarnsystem auf Schulebene, was das Erreichen der Bildungsstandards angeht, unterbleibt.

Für diesen Verzicht auf eine kriteriale Rückmeldung und die insgesamt schwierige Akzeptanzsituation bei den Vergleichsarbeiten gibt es viele Gründe. Ein Grund dafür ist unter anderem die problematische Ableitung der VERA-Aufgaben aus den Bildungsstandards für die Jahrgangsstufen 4 und 10. Dies führt bei einer Testung in den Jahrgängen 3 und 8 vor allem bei schwachen Schülerinnen und Schülern zu einer grundsätzlichen Überforderung und Demotivierung. Der Informationsgewinn von VERA ist vor allem für Schulen in Brennpunkten seit Jahren eher frustrierend; Lernstandsdiagnostik zur besseren individuellen Förderung auf dieser Basis ist den Lehrkräften so schwer vermittelbar.

Weitere Kritikpunkte aus Sicht der Schulpraxis sind das problematische Verhältnis von Aufwand und Ertrag und die Schwierigkeit, VERA-Ergebnisse als schulinterne Trendmessung nutzen zu können, da es immer wieder zu großen Irritationen durch nicht plausible Varianzen zwischen den Jahren kommt. Aber selbst wenn es gelänge, diese Schwierigkeiten zu lösen, bleibt als Frage derzeit offen, wie die zunehmende Integration bzw. Inklusion und die daraus resultierenden Anforderungen an gemeinsame und zugleich differenzierende Testverfahren auf der Ebene der Testheftkonstruktion bewältigt werden können.

3. Das Transparenzdefizit

3.1 Mangelnde Transparenz bei der Weiterentwicklung der Gesamtstrategie

Nachdem in den ersten Jahren der Neuausrichtung der Bildungspolitik im Zeichen der empirischen Wende die Entwicklung der Bildungsstandards und im Zeichen der Erarbeitung von Kompetenzstufenmodellen (KSM) die Diskussion über Nützlichkeit und Konsequenzen der Ergebnisse noch in einem breiten Dialog zwischen Wissenschaft, Politik, Administration und Schulpraxis geführt wurde, findet die Debatte über inzwischen aufgetretene Herausforderungen, Schwierigkeiten und Perspektiven heute eher in geschlossenen Kreisen statt.

Zu Beginn der 2000er-Jahre fanden beispielsweise in Hubertusstock (Brandenburg) mehrere große PISA-Tagungen statt, wo in großer Runde unter Einbeziehung der Kritikerinnen und Kritiker über alles gesprochen werden konnte. Im September 2004 fand in Berlin in der Akademie der Wissenschaften eine von der KMK organisierte öffentliche ganztägige Veranstaltung mit ca. 250 Teilnehmerinnen und Teilnehmern statt, in der ebenfalls ein offener Dialog geführt wurde.

Die aktuelle Diskussion scheint eher durch politische Ängste geprägt zu sein. Ein Beispiel dafür ist die von der KMK geplante, nicht öffentliche Veranstaltung

im Dezember 2014 in Berlin, in der an einem halben Tag mit handverlesenen Teilnehmerinnen und Teilnehmern in kleiner Runde die aktuellen Probleme der Gesamtstrategie und ihre Lösung diskutiert werden sollen.

Es besteht die große Gefahr, dass sich die Ziele dieses Fachgesprächs – den Entwurf von Eckpunkten der bildungspolitischen Öffentlichkeit vorzustellen, die Transparenz des Überarbeitungsprozesses und die Akzeptanz der vorgenommenen Überarbeitungen zu fördern – nicht erreichen lassen, wenn diejenigen, die aus unterschiedlichen Motiven und auf der Basis unterschiedlicher Erfahrungen zu den Kritikerinnen und Kritikern und zu den Skeptikerinnen und Skeptikern gehören, nicht einbezogen werden. Dies lässt sich auch nicht mit einer pflichtgemäßen Einladung an die Vorsitzenden der großen Lehrerverbände erledigen. Es kommt eben nicht nur darauf an, das Produkt – den neuen Beschluss zur Gesamtstrategie – zu überarbeiten, sondern den Prozess, wie es zu diesem neuen Beschluss kommt, selbst als Chance für einen breiten Dialog mit Kritikerinnen und Kritikern und mit Bedenkenträgerinnen und Bedenkenträgern zu nutzen.

Das bisherige Top-Down-Verfahren bei der Umsetzung der Gesamtstrategie hat sich offensichtlich nicht bewährt. Der Prozess der Erarbeitung einer veränderten Gesamtstrategie sollte deshalb auf breite Akzeptanz aller Akteure angelegt sein.

Zusätzlich müsste die Diskussion zielgruppenspezifischer geführt werden. Bei der Weiterentwicklung der LSA/Ländervergleiche kommt es darauf an, sich im Dialog mit wissenschaftlicher Kritik von außen offen auseinanderzusetzen. Eine Weiterentwicklung der Vergleichsarbeiten braucht vor allem eine konstruktive Öffnung zur Schulpraxis und zu den Lehrerverbänden. Dies könnte beispielsweise mit einer eigenständigen Kommunikationsstrategie (z.B. in mehreren regionalen Foren) umgesetzt werden, bei der Transparenz und Partizipation im weiteren Verfahren eine entscheidende Voraussetzung darstellen. Nur wenn es gelingt, die in- zwischen offenkundigen Probleme und Kontroversen im Dialog mit der Schulpraxis offen und konstruktiv zu bearbeiten, besteht die Chance, den für 2015 geplanten Beschluss der KMK zur überarbeiteten Gesamtstrategie als Einstieg in einen umfassenden Implementationsprozess auch gegenüber Schulen zu nutzen.

3.2 Mangelnde Transparenz bei Sachproblemen

Mangelnde Transparenz gibt es allerdings nicht nur im Diskussionsprozess über die Strategie, sondern auch bei Sachproblemen, die zu entscheiden sind. Zu den offenen Sachproblemen, zu denen es keine fachöffentliche Diskussion gibt, gehört beispielsweise die Entwicklung von Kompetenzstufenmodellen. Das IQB entwickelt und die KMK beschließt. Kaum jemand aus der Bildungspolitik und der Bildungsadministration ist in der Lage, die Rationalität dieser Kompetenzstufenmodelle nachzuvollziehen.

In den Texten der KMK werden zwar Ansprüche und Kriterien formuliert; ob und wie die jeweilige Balance aber umgesetzt wird, ist nur für eine kleine Expertengruppe nachvollziehbar. Kompetenzstufenmodelle sollen u.a.

„[...] motivierende Leistungserwartungen formulieren, die Entwicklungsimpulse an den Schulen auslösen (hierfür dienen insbesondere die Regelstandards plus) und breite bildungspolitische Akzeptanz insbesondere bei den Lehrkräften erreichen“ (IQB 2009).

Je nach Fach und Domäne hat es in der Vergangenheit im Rahmen der vom IQB veranstalteten Anhörungen durchaus kontroverse Diskussionen vor allem mit Fachdidaktikerinnen und Fachdidaktikern sowie Lehrkräften gegeben; trotzdem scheinen die anschließend verkündeten Kompetenzstufenmodelle oft den Charakter einer gewissen Naturgesetzlichkeit zu haben.

Ein weiteres Sachproblem sind die mit diesen Setzungen verbundenen Auswirkungen auf die Berechnung der VERA-Ergebnisse, die ja in Anbindung an diese Modelle berichtet werden. Die dabei immer wieder auftretenden Inkonsistenzen zwischen Leistungsständen bei VERA (Jahrgang 8) und denen im IQB-Ländervergleich (Jahrgang 9) sind nicht wirklich aufgeklärt; hier treten Plausibilitätsvermutungen an die Stelle einer nachvollziehbaren Erklärung.

Zehn Jahre, nachdem die diesen Messungen zugrunde liegenden Bildungsstandards beschlossen wurden, wäre es auch notwendig, auf der Basis der vorliegenden Ergebnisse, die Bildungsstandards selbst einer Evaluation zu unterziehen.

4. Das Transferdefizit

Noch nie wurde in Deutschland so viel und mit so viel Geld im Bereich der empirischen Bildungsforschung gearbeitet. Die Menge der neuen Befunde und Erkenntnisse einerseits und ihre Rezeption und Nutzung durch die Politik, die Bildungsverwaltungen und Landesinstitute der Länder (oder durch Schulen selbst) andererseits stehen in einem krassen Missverhältnis zueinander. Systematische Aufarbeitung, Rezeption und Reflexion der Befunde finden weder auf der Systemebene oberhalb der Einzelschulen noch in den Schulen selbst in ausreichender Weise statt. Im Kern geht es um zwei Fragen: Wie kommt das empirische Wissen in die Bildungspolitik, und wie kommt es in die Schulpraxis?

An dieser Stelle muss man sich darüber verständigen, welches Wissen gemeint ist und auf welche Ebene es sich bezieht. Geht es um beschreibendes Wissen zu den einzelnen Schülerinnen und Schülern, zur Klasse, zum Jahrgang, zur Schule, zum Gesamtsystem? Geht es um Wissen über die Fragestellungen der jeweiligen Studien, zu Rahmenbedingungen, zu Ergebnissen, zu abgeleiteten Schlussfolgerungen, zu

Fragen von Unterricht, zur Leistungserreichung, zur Schulentwicklung oder zu den Folgen von Steuerungshandlungen?

Bezogen auf die Ebene der Bildungspolitik und Bildungsadministration kann festgestellt werden, dass in den letzten 15 Jahren eine neue Dynamik entstanden ist, bei der von Seiten der empirischen Bildungsforschung sehr unterschiedliche Hoffnungen geschürt wurden. Forscherinnen und Forscher haben sich teilweise bewusst oder unbewusst dahingehend missverstehen lassen, dass sie in der Lage wären, der Bildungspolitik durch evidenzbasierte Bildungsforschung eindeutige Empfehlungen geben zu können, was zu tun ist. Es ist Tenorth uneingeschränkt zuzustimmen, der es für nützlicher hält, die wechselseitigen Probleme nach Aufgaben, Adressaten und Organisationsformen zu unterscheiden, aber von der Differenz der Systeme nicht abzusehen (vgl. Tenorth 2011). Bildung als „Konfession“ durch „*evidence-based policy*“ als neuer „Profession“ ersetzen zu können, ist wohl ein zu hoher Anspruch. Zur Profession der Politik sollte es allerdings gehören, „*evidence-informed*“ zu entscheiden, was voraussetzt, dass die der Bildungspolitik zuarbeitende Bildungsadministration einen strukturierten Zugang zu den relevanten Forschungsbefunden besitzt.

In der politischen Diskussion über Transferprozesse im Bildungsbereich wird die Rolle der Bildungsadministration oft unterschätzt. Politische Schlüsse aus zahlreichen Studien können nur gezogen werden, wenn die Teilschritte Rezeption und Reflexion der jeweils aktuellen Befunde in den Bildungsadministrationen vollzogen werden. Dies ist bis heute ein Problem, weil die klassische Rekrutierung der Bildungsadministration grundsätzlich über die Schritte Lehrkraft, Schulleitung, Schulrätin bzw. Schulrat, Oberschulrätin bzw. -schulrat/Ministerialbeamtin bzw. -beamtin läuft. Tatsächlich sind aber in der Bildungsverwaltung neue Kompetenzen erforderlich, die dieser Personenkreis nicht von sich aus mitbringt. Hierzu gehören breite Kenntnisse von Ergebnissen der empirischen Bildungsforschung sowie auch Kenntnisse in der Methodenlehre usw.

Die Fülle an neuen Erkenntnissen und Befunden im Bereich der empirischen Bildungsforschung führt allerdings längst zu einer strukturellen Überforderung der Bildungsadministration (und natürlich auch der Bildungspolitik), die Transfer- und Übersetzungsprozesse, die für eine „*evidence-informed policy*“ erforderlich sind, überhaupt zu gestalten.

Notwendig ist es deshalb für die Bildungsadministration, aber auch für die Schulpraxis, der Frage „*what works*“ systematisch nachzugehen. Hierzu bedarf es Strukturen oder Institutionen, die dauerhaft angelegt sein müssen. Dies wäre *eine* Voraussetzung für eine bessere Nutzung der zahlreichen bereits vorhandenen Forschungsbefunde für Maßnahmen der Schul- und Unterrichtsentwicklung auf allen Ebenen. Der Frage, wie das Wissen in die Praxis kommt, muss künftig höhere Priorität eingeräumt werden.

5. Vorschläge zur Weiterentwicklung der Gesamtstrategie der KMK zum Bildungsmonitoring

Deutschland sollte sich auch weiterhin an den internationalen Large-Scale-Assessments beteiligen. Allerdings sollte der bisher nicht eingelöste Anspruch aufgegeben werden, die Ergebnisse des IQB-Ländervergleichs mit den internationalen Skalen von IGLU, TIMSS und PISA zu verbinden. Eine Umrechnung der Länderergebnisse des IQB auf PISA-Maßstäbe ist nicht nur mit erheblichen methodischen Problemen belastet; sie liefert vermutlich auch kaum Erkenntnisse, die neue Handlungsoptionen nahelegen.

Statt einer weiteren Intensivierung der Forschung mit nationalen PISA-Daten sollten die im Rahmen der Ländervergleiche des IQB gewonnenen Informationen umfassender als bisher aufbereitet und ausgewertet werden. Im Unterschied zu den nationalen PISA-Daten steckt in diesen Daten mehr Potenzial, die innerhalb Deutschlands bestehenden Leistungsunterschiede und unterschiedlichen gesellschaftlichen Kontexte miteinander in Beziehung zu setzen und aufzuklären. Das bisherige Vorgehen bei PISA, aus Querschnittsuntersuchungen im Drei-Jahres-Zyklus Trendaussagen abzuleiten, ist psychometrisch hoch komplex und liefert nur eingeschränkte Informationen. Ein Grund dafür ist die innerhalb Deutschlands sehr unterschiedliche und nicht zeitstabile Verteilung der getesteten 15-Jährigen auf verschiedene Jahrgangsstufen und Schulformen. Die Bezugnahme auf den 9. Jahrgang, wie sie im IQB-Ländervergleich vorgenommen wird, vermeidet die Probleme, die bei PISA mit dem Stichwort „Kohorteneffekte“ umschrieben sind. Um Leistungsdifferenzen nicht nur zu beschreiben, sondern auch erklären zu können, sollten sich die Länder zukünftig zusätzlich auf qualitative Fragestellungen und längsschnittliche Ergänzungen einigen. Die dem IQB bisher zur Verfügung stehenden Mittel reichen für solche Ausweitungen allerdings nicht aus und müssten aufgestockt werden. Zusätzliche Mittel könnte das BMBF über sein Rahmenprogramm „Empirische Bildungsforschung“ über entsprechende Förderbekanntmachungen bereitstellen (vgl. auch BMBF 2014).

Ziel der Bildungspolitik muss es sein, allen Schülerinnen und Schülern den Weg zu einem mittleren Schulabschluss am Ende der Jahrgangsstufe 10 zu ermöglichen. Darauf beziehen sich auch die von der KMK vor zehn Jahren beschlossenen Bildungsstandards. Der methodisch umstrittene und bildungspolitisch fragwürdige Versuch, für sogenannte „Hauptschülerinnen und Hauptschüler“ eigene Hauptschul-Bildungsstandards zu überprüfen, sollte beendet werden. In Schulen mit mehreren Bildungsgängen wird bisher auf der Basis von Lehrkräfteeinschätzungen versucht, die potenziellen „Hauptschülerinnen und Hauptschüler“ zu identifizieren. Tatsächlich gibt es bisher keine empirischen Befunde, die diese Populationsschätzung belastbar überprüfen. Kompetenzstufenverteilungen zu einer auf dieser Basis geschätzten Hauptschulpopulation liefern keine verlässlichen Erkenntnisse, die den jeweils

Verantwortlichen in Schulen und Bildungsadministration bei der Erreichung des oben bezeichneten Ziels helfen könnten. Die im Umbau zu Zwei-Säulen-Modellen befindlichen Schulsysteme der Länder in der Sekundarstufe I machen die Erkenntnisse ohnehin obsolet, und die Förderstrategie der KMK lässt sich auf diesem Wege auch nicht nachvollziehbar evaluieren.

Statt Länderrankings auf der Basis von Mittelwerten zu diskutieren, sollten künftig in der Berichterstattung des IQB die Verteilungen auf die verschiedenen Kompetenzstufen im Mittelpunkt stehen. Besonderes Augenmerk muss dabei auf die Schüleranteile gerichtet werden, die die sogenannten Mindeststandards verfehlen. Dabei sollten die Analysen zusätzlich auch auf Aspekte der ökonomischen und soziokulturellen Kontexte, z.B. in Regionen und Großstädten mit ähnlichen Strukturen, eingehen. Angesichts des in einigen Ländern besonders schwachen Fragebogenrücklaufs und datenschutzrechtlicher Probleme bei der Erhebung der notwendigen Individualauskünfte stellt die z.T. schwache Belastbarkeit der sozioökonomischen und soziokulturellen Angaben ein zusätzliches Problem dar, das künftig besser gelöst werden muss.

Die Gründung des Instituts zur Qualitätsentwicklung im Bildungswesen (IQB) im Jahr 2004 war eine positive und wichtige Entscheidung. Allerdings zeigt sich nach zehn Jahren, dass die von den Ländern gemeinsam aufgebraachten Mittel zur Finanzierung des IQB nicht ausreichend sind, um die inzwischen weiter gestiegenen Anforderungen an das IQB in hoher Qualität zu erfüllen. So ist das IQB mit seinen jetzigen Mitteln nicht in der Lage, den Ländern zeitnah ausführlichere und vertiefende Analysen und Zusatzauswertungen zu den bereits erhobenen Daten zur Verfügung zu stellen. Zudem müsste die fachdidaktische Kompetenz im IQB gestärkt werden. Vor dem Hintergrund restriktiver Ressourcenbewirtschaftung in allen Ländern werden diese kaum bereit sein, eine relevante Aufstockung des IQB-Haushaltes zu beschließen. Eine Alternative könnte darin bestehen, die bestehenden verfassungskonformen Möglichkeiten dazu zu nutzen, dass Bund und Länder das IQB als ein gemeinsames Institut betreiben und finanzieren. Schließlich geht es auch um nationale Bildungsstandards und deren Überprüfung, also um eine nationale Aufgabenstellung.

Die KMK sollte den bei VERA offenkundigen Mangel der problematischen Ankopplung der VERA-Aufgaben an die Kompetenzstufenmodelle der Bildungsstandards der Jahrgangsstufen 4 und 10 beheben, indem die Aufgaben für die Vergleichsarbeiten Jahrgang 3 und 8 durch das IQB künftig auf der Basis der für diese Jahrgangsstufen entwickelten Kompetenzstufenmodelle erarbeitet werden. Nur damit würde eine bessere Passung von SOLL und IST erreicht und könnte der diagnostische Wert der Vergleichsarbeiten erhöht werden.

6. Schlussbemerkung

Die rückblickende Betrachtung der letzten 15 Jahre macht deutlich, dass der formale Anlass der jetzigen Debatte, die Überarbeitung des KMK-Beschlusses von 2006, genutzt werden sollte, um das Verhältnis von Evaluation, Schulpraxis und Bildungspolitik neu auszubalancieren. Die seit 2005 in allen Ländern implementierten Verfahren der externen Evaluation von Schulen auf der Basis umfassender Handlungsrahmen „Schulqualität“ zeigen, wie es gelingen kann, Akzeptanz und Erkenntnisgewinn zu verbinden und den Schulen und der Bildungsadministration nützlich Wissen zur Schulentwicklung zur Verfügung zu stellen. Im Bemühen der Bildungspolitik, Ursachen für unbefriedigende Ergebnisse zu finden, leistet die externe Schulevaluation einen mindestens so großen Beitrag wie ein Ländervergleich, weil sie die Prozesse von Unterricht und Schule evaluiert und mit ihren Rückmeldungen jede Schule mit sehr viel konkreteren Informationen über Stärken und Schwächen versorgt, als dies mit Leistungsergebnissen aus Vergleichsarbeiten möglich ist. In vielen Ländern werden die Schulen inzwischen mit unterschiedlich strukturierten Verfahren auch bei der internen Evaluation unterstützt. Die interne und externe Evaluation von Schul- und Unterrichtsqualität sollte daher als zusätzliche Säule in die Gesamtkonzeption integriert werden.

Damit eine neue Gesamtstrategie umfassend wirksam wird, kommt es nicht nur auf die Verständigung auf der Ebene der KMK und eine Abstimmung mit dem Bund zu einer gemeinsamen Finanzierung des IQB an. Die Hauptaufgabe der Umsetzung liegt auf der Ebene der Länder, die weiterhin in erster Linie für die bildungspolitischen Rahmensetzungen verantwortlich sind. Tatsächlich ist in den letzten Jahren vor allem der Bereich des Messens auf unterschiedlichen Ebenen ausgebaut und institutionalisiert worden; der Bereich der Entwicklung und Unterstützung von Lehrkräften und Schulen kam hingegen konzeptionell und tatsächlich zu kurz. Die Überarbeitung der Gesamtstrategie sollte diesen Prozess nachvollziehen und den Titel des Beschlusstextes ändern. Die überarbeitete Strategie sollte sich nicht auf Monitoring beschränken, sondern eine Gesamtstrategie der Qualitätssicherung beschreiben, die das Monitoring auf der Systemebene bewusst um die Ebene der Einzelschule ergänzt und dabei nicht nur Leistungsergebnisse (VERA), sondern auch Prozessqualitäten (externe Schulevaluation) einbezieht. Messen und Entwickeln müssen dabei konzeptionell zusammengeführt werden.

Das durch Forschung zu generierende Wissen über die Leistungserbringung ist wichtiger als nur das Wissen über die Leistungsergebnisse. Statt einer Weiterentwicklung des Bildungsmonitoring muss die Weiterentwicklung der pädagogischen Praxis in den Mittelpunkt rücken. Gebraucht wird eine neue Balance zwischen Verfahren der internen und externen Evaluation, zwischen Erkenntnisgewinnung, Implementation

und Unterstützung der Schulen. Die KMK hat die Chance, diese Neujustierung im Jahr 2015 politisch zu verabreden – sie sollte sie nutzen.

Literatur und Internetquellen

- Baumert, J./Stanat, A./Demmrich, A. (2001): PISA 2000. Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2000. Opladen: Leske + Budrich, S. 15-68.
- BMBF (2014): Richtlinie zur Förderung von Forschungsvorhaben in Anknüpfung an Large-Scale-Assessments, veröffentlicht im Bundesanzeiger am 04.08.2014.
- IQB (Institut für Qualitätsentwicklung im Bildungswesen) (2009): Kompetenzstufenmodelle. URL: <http://www.iqb.hu-berlin.de/bista/ksm>; Zugriffsdatum: 20.09.2014.
- Klafki, W. (*1996): Neue Studien zur Bildungstheorie und Didaktik. Weinheim: Beltz.
- KMK (2006): Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. Berlin: KMK. URL: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2006/2006_06_02-Bildungsmonitoring.pdf; Zugriffsdatum: 20.9.2014.
- Richter, D./Böhme, K./Becker, M./Pant, H.A./Stanat, P. (2014): Überzeugungen von Lehrkräften zu den Funktionen von Vergleichsarbeiten: Zusammenhänge zu Veränderungen im Unterricht und den Kompetenzen von Schülerinnen und Schülern. In: Zeitschrift für Pädagogik 60, H. 2, S. 225-244.
- Tenorth, H.-E. (2011): Nicht ohne Reputationsverluste. In: Frankfurter Allgemeine Zeitung vom 27.11.2011.

Hans-Jürgen Kuhn, geb. 1953, seit 1991 Referatsleiter für die Schulaufsicht der Sekundarstufe I im Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, von 2000 bis 2013 Referatsleiter für Grundsatzangelegenheiten der Qualitätsentwicklung und -sicherung.

Anschrift: Gleditschstr. 37, 10781 Berlin
Kuhn.Hans-Juergen@web.de