

---

Norbert Maritzen

## **Glanz und Elend der KMK-Strategie zum Bildungsmonitoring**

Versuch einer Bilanz und eines Ausblicks

---

### **Zusammenfassung**

*Mit ihrer Gesamtstrategie zum Bildungsmonitoring hat die Kultusministerkonferenz im Jahre 2006 vor allem ihre Beteiligung an bzw. Beauftragung von Schulleistungstests (PISA, TIMSS, PIRLS, IQB-Ländervergleich, VERA) und die Entwicklung von Bildungsstandards in einen systematischen Zusammenhang gestellt. Der Beitrag bilanziert die historisch zu nennende Leistung, die die Implementierung eines Systems der Dauerbeobachtung des Bildungsbereichs für die Ländergemeinschaft bedeutet. Es werden zentrale Probleme benannt, die im Zuge der Entfaltung der Strategie nunmehr sichtbar werden. Schließlich werden einige Linien einer möglichen Weiterentwicklung der Strategie skizziert.*

*Schlüsselwörter: Bildungsmonitoring, Kultusministerkonferenz, PISA, VERA, IQB-Ländervergleich, Qualitätssicherung*

### **Splendor and Misery of the KMK-Strategy of Educational Monitoring**

An Attempt of Taking Stock and Looking Forward

#### **Abstract**

*With its overall strategy of educational monitoring, the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (KMK) has put its participation in resp. assignment of student achievement tests (PISA, TIMSS, PIRLS, IQB-comparison of the German states, VERA) and the development of educational standards into a systematic framework. This article balances the historic effort of the German states to implement a system of permanent surveillance in the field of education. It specifies key problems, which have become visible in the pro-*

*cess of the strategy. Finally, some lines of a possible further development of the strategy are sketched.*

*Keywords: educational monitoring, Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany, PISA, VERA, IQB-Comparison of the German states, quality management*

Nachdem sich Deutschland 1995 zum ersten Mal an der TIMS-Studie beteiligt hatte, stellte der Konstanzer Beschluss der Kultusministerkonferenz (KMK) vom 24.10.1997<sup>1</sup> eine bildungspolitische Wende dar, da die Ländergemeinschaft sich unter öffentlichem Druck entschloss, ihre Abstinenz hinsichtlich einer systematischen und vergleichenden empirischen Überprüfung von Schülerleistungen aufzugeben. Es folgten in den Jahren danach die Beschlüsse zur regelmäßigen Teilnahme an PISA, TIMSS und IGLU, zur Entwicklung von Bildungsstandards für die Grundschule und die Sekundarstufe I (später auch für das Abitur<sup>2</sup>) und deren Überprüfung im Ländervergleich, zur kooperativen Implementierung testbasierter Vergleichsarbeiten (VERA) auf Länderebene und zu einer Bildungsberichterstattung gemeinsam mit dem Bund. Die KMK hat ihre sukzessiv entstandenen Aktivitäten 2006 in einer „Gesamtstrategie zum Bildungsmonitoring“<sup>3</sup> systematisiert. Zugleich wurde sowohl auf Länderebene als auch länderübergreifend und mit dem Bund mit der Gründung wissenschaftlicher Einrichtungen die Infrastruktur gelegt, um die mit diesen Vorhaben verbundenen Aufgaben zu bewältigen.

Nach mehr als 15 Jahren, in denen sich dieses umfassende standardbasierte Assessmentprogramm sukzessive entfaltet hat, mehren sich die Hinweise darauf, dass die Entwicklung einer kritischen Bilanzierung bedarf, vielleicht auch einer strategischen Neuausrichtung. Im Folgenden sollen deshalb zunächst die historisch zu nennenden Leistungen (Abschnitt 1), aber auch die Brüche, Ungleichzeitigkeiten und Inkonsequenzen (Abschnitt 2) skizziert werden, die mit der Implementierung der Gesamtstrategie zum Bildungsmonitoring verbunden waren. Anschließend werden die Spezifik des mit den Maßnahmen generierten Wissens und die Schwierigkeiten der Nutzung betrachtet (Abschnitt 3). Schließlich werden einige Linien herausgearbeitet, entlang derer eine Weiterentwicklung der Strategie diskutiert werden könnte (Abschnitt 4).

1 Vgl. URL: [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/1997/1997\\_10\\_24-Konstanzer-Beschluss.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1997/1997_10_24-Konstanzer-Beschluss.pdf); Zugriffsdatum: 20.08.2014.

2 Bei den Abiturstandards verzichtete die KMK auf die empirische Überprüfung im Ländervergleich auf der Basis von Kompetenzstufenmodellen.

3 Vgl. URL: [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2006/2006\\_08\\_01-Gesamtstrategie-Bildungsmonitoring.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2006/2006_08_01-Gesamtstrategie-Bildungsmonitoring.pdf); Zugriffsdatum: 20.08.2014.

## 1. Die Monitoringstrategie der KMK als historische Leistung

Die nach den Plöner Beschlüssen der KMK vom 02.06.2006 veröffentlichte Gesamtstrategie zum Bildungsmonitoring<sup>4</sup> besteht aus vier miteinander verbundenen Elementen: den internationalen Schulleistungsuntersuchungen (PIRLS/IGLU, TIMSS, PISA), der zentralen Überprüfung der Bildungsstandards im Ländervergleich (IQB-Ländervergleich), den Vergleichsarbeiten zur landesweiten Überprüfung der Leistungsfähigkeit einzelner Schulen (VERA 3 und VERA 8) und der gemeinsamen Bildungsberichterstattung von Bund und Ländern. Vergleicht man den inzwischen erreichten Entwicklungsstand des Bildungsmonitorings mit der Situation noch in den 1990er-Jahren, sind die erreichten Erfolge offenkundig:

- Der Ländergemeinschaft ist es gelungen, einen bis dato nicht für möglich gehaltenen Konsens über ein länderübergreifendes Programm der Leistungsmessung zu erzielen und damit Transparenz hinsichtlich zentraler Leistungsindikatoren zu schaffen. Damit verbunden war ein breit akzeptierter Einstieg aller Länder in die Schulleistungsevaluation auf Schul- bzw. Unterrichtsebene auf der Basis standardisierter Tests (VERA), die nachhaltige Rückwirkungen auf Referenzmaßstäbe von Lehrkräften bei der Leistungsdiagnostik (Einübung des Umgangs mit kriterialen Bezugsnormen) hat. In diesem Zusammenhang wurden eine länderübergreifende Kooperation auf fachlicher Ebene etabliert, die Testadministration institutionalisiert, die Professionalisierung des Personals vorangetrieben und auch die Rückmeldeformate deutlich instruktiver gestaltet.
- Das Assessmentprogramm wurde systematisch auf einen normativen, empirischer Überprüfung zugänglichen Referenzrahmen bezogen, der in diskursiv entwickelten Standards und in Kompetenzmodellen operationalisiert ist. Insbesondere in der bildungstheoretischen Debatte im Kontext von PISA 2000 und mit der so genannten Klieme-Expertise wurden Grundlagen für einen an empirische Forschung anschlussfähigen Bildungsbegriff gelegt und kritisch diskutiert (vgl. Benner 2007; Hericks 2004; Klieme u.a. 2003; Koch 2004; Tenorth 2004a, 2004b).
- Die Erklärungskraft der internationalen Assessments wurde durch strategische Erweiterung/Vertiefung (z.B. Stichprobenerweiterung, längsschnittliche Erweiterung, Ergänzung der Erhebungsinstrumente, Erweiterung der internationalen Frameworks, Entwicklung innovativer Testformate u.a.m.) systematisch gestärkt.
- Die KMK gründete 2004 das Institut zur Qualitätsentwicklung im Bildungswesen (IQB)<sup>5</sup> und 2010 mit dem Bund das Zentrum für internationale Bildungsvergleichsstudien (ZIB)<sup>6</sup>. Die Mehrheit der Länder bündelten ihre Aufgaben des Bildungsmonitorings unterschiedlich weitgehend in bereits existierenden oder in

4 Vgl. URL: <http://www.kmk.org/presse-und-aktuelles/pm2006/ergebnisse-der-314plenarsitzung.html>; Zugriffsdatum: 20.08.2014.

5 Vgl. URL: <http://www.iqb.hu-berlin.de/>; Zugriffsdatum: 20.08.2014.

6 Vgl. URL: <http://zib.education/startseite.html>; Zugriffsdatum: 20.08.2014.

eigens dafür neu gegründeten Einrichtungen, die in der Regel nun als Dienstleister Monitoringaufgaben für die zuständigen Bildungsministerien übernehmen (vgl. Fickermann 2014; Rürup 2014).

- Die von zahlreichen Mythen gekennzeichnete Wahrnehmung und Selbstbeschreibung des Bildungssystems und der Bildungspolitik wurden in Folge der ersten Ergebnisveröffentlichungen von PISA gleichsam entzaubert und mit Evidenzansprüchen konfrontiert. Die KMK reagierte darauf in einem ersten Schritt mit der Definition von konsensfähigen, zukunfts-offenen Handlungsfeldern,<sup>7</sup> die vielleicht gerade aufgrund ihrer Vagheit, die den Erklärungsanspruch von PISA deutlich überzog, die politische „Verarbeitung“ der Ergebnisse ermöglichten (siehe hierzu auch Tillmann u.a. 2008).
- Kennzeichnend für die diskursive Lage war eine deutliche Veränderung der Aufmerksamkeitsstrukturen von Öffentlichkeit und Medien: Schulleistungsuntersuchungen wurden schlagzeilentauglich. Wichtiger aber war, dass das Programm nachhaltig das Agenda-Setting im politischen und öffentlichen Raum beeinflusst hat. Themen, die vorher so nicht virulent waren, sind seither aus der bildungspolitischen Diskussion nicht mehr wegzudenken, wie z.B. der Zusammenhang von Bildung, Humankapital und internationalem Wettbewerb; die Bedeutung von Bildung im Lebenslauf; das Thema Bildungsgerechtigkeit; das Schicksal von „Risikogruppen“ (ein neuer Begriff für Bildungsverlierer und -verliererinnen) und deren Kosten für den Sozialstaat und eine demokratische Gesellschaft; der Stellenwert frühkindlicher Bildung für die gesamte Bildungskarriere; die Bedeutung der Integration von Zuwanderern und ihre sozialen Kosten; regionale Disparitäten in Deutschland; schließlich auch das ordnungspolitische Thema des Bildungsföderalismus, der sowohl als Innovationsbarriere problematisiert als auch als wettbewerbsfördernd neu begründet wurde.

Mit der Gesamtstrategie hat die KMK den Versuch unternommen, zum einen die bis 2006 zum Teil unverbunden entstandenen Aktivitäten quasi nachholend zu einem in sich schlüssigen Konzept zu bündeln, zum anderen die Selbstverpflichtung auf eine gesamtstaatliche, langfristige Strategie zur Transparenz- und Qualitätssicherung festzuschreiben. Dafür hatte sie bereits mit der Einrichtung der Amtschefscommission „Qualitätssicherung in Schulen“ im Jahr 1997 eine länderübergreifende Koordination der Dauerbeobachtung und Qualitätssicherung auf Staatssekretärs- und Arbeitsebene geschaffen. Dass es der KMK gelungen ist, das Programm zu entwerfen und umzusetzen, war angesichts der chronisch zu nennenden Entscheidungsschwäche ihrer Gremien alles andere als erwartbar. Der Erfolg verdankt sich einer spezifischen historischen Konstellation.

---

7 Vgl. URL: <http://www.kmk.org/fileadmin/pdf/PresseUndAktuelles/2002/massnahmen.pdf>; Zugriffsdatum: 20.08.2014.

- Nach Veröffentlichung der TIMSS-Ergebnisse 1996 entstand erstmals eine öffentliche Aufmerksamkeit für Leistungs- und Gerechtigkeitsdefizite des deutschen Bildungswesens, die die Diskussion aus der rein wissenschaftlichen in die öffentliche und mediale Sphäre holte und mit Problemlöseerwartungen auch auf den bildungspolitischen Diskurs einwirkte.
- Der Strategie des Bundes, in internationalen Gremien der OECD die Beteiligung Deutschlands an Large-Scale-Assessments voranzutreiben und die Länder mit Angeboten der Mischfinanzierung oder mit Aktivitäten im Rahmen seiner Ressortforschungskompetenz – z.B. den Expertisen zur Entwicklung von Bildungsstandards (vgl. Klieme u.a. 2003), zur Implementation der Bildungsstandards (vgl. Oelkers/Reusser 2008), zu konzeptionellen Grundlagen der Bildungsberichterstattung (vgl. Rauschenbach u.a. 2004; Baethge/Buss/Lanfer 2003, 2004), zu Kompetenzen von Kindern im Elementarbereich (vgl. Fthenakis u.a. 2005; Roßbach/Weinert 2008), zur Förderung von Lesekompetenz und zur Sprachförderung (vgl. Artelt u.a. 2005; Ehlich u.a. 2005, 2008a, 2008b; Redder u.a. 2011) u.a.m. – vor sich herzutreiben, hat die Entscheidungsprozesse auf KMK-Ebene erheblich beschleunigt, da man sich auf Länderseite schwerlich den Vorwurf gefallen lassen wollte, international den Anschluss zu verpassen. So wurde bereits vor dem Konstanzer Beschluss die Beteiligung an TIMSS und auch an PISA entschieden.
- Es gab eine günstige personelle Konstellation auf der Entscheiderenebene insofern, als wenige Kultusminister und vor allem ihre auch in der Sache kundigen, wissenschaftsaffinen Amtschefs mit Einfluss in der Ländergemeinschaft einerseits und wenige Wissenschaftler und Wissenschaftlerinnen mit politischer Sensibilität und Einfluss im Wissenschaftsbetrieb andererseits den Wechsel der KMK zu mehr Transparenz und einer längerfristig angelegten Bearbeitungsstrategie für skandalträchtige Defizite durchsetzen konnten.
- Obwohl bis dahin kaum erwartbar, konnte es – maßgeblich vorangetrieben durch das KMK-Sekretariat und die Amtschefskommission „Qualitätssicherung in Schulen“ – gelingen, die Strategie durch Gründung des IQB institutionell abzusichern.
- Im Zuge der Föderalismusreform wurde 2006 trotz der Abschaffung einer gemeinsamen Bildungsplanung (Auflösung der Bund-Länder-Kommission (BLK)) an der Möglichkeit eines „gemeinsamen Zusammenwirke[n]s von Bund und Ländern zur Feststellung der Leistungsfähigkeit des Bildungswesens im internationalen Vergleich und bei diesbezüglichen Berichten und Empfehlungen“ (Art. 91b Abs. 2 GG) festgehalten, deren nähere Ausgestaltung am 21.05.2007 in einem Verwaltungsabkommen geregelt wurde. Es wurde eine entsprechende Steuerungsgruppe eingerichtet, die von einem wissenschaftlichen Beirat beraten wird.

Vergegenwärtigt man sich die bildungspolitische Diskussion noch der 1990er-Jahre, so sind die beschriebenen Leistungen, auf die sich die Ländergemeinschaft mit der Entwicklung und Implementierung der Gesamtstrategie zum Bildungsmonitoring berufen kann, nicht hoch genug einzuschätzen.

## 2. Inkonsistenzen der Strategieumsetzung und Grenzen der politischen Folgenbearbeitung

Nicht unerwartet ist die sukzessive Entfaltung des komplexen Monitoringprogramms mit konzeptionellen und strategischen Webfehlern behaftet, die im fortgeschrittenen Zustand der Umsetzung ebenso offenkundig werden wie das historisch Unerhörte. Dies erstaunt nicht bei einem Programm, das wissenschaftlich höchst anspruchsvoll ist, seine konzeptionellen und infrastrukturellen Grundlagen im Zuge der Entfaltung zum Teil selbst mit schaffen muss und bei seiner Initiierung im deutschsprachigen Raum ohne Beispiel war. Auch die wesentlich weiter fortgeschrittenen Entwicklungen z.B. in Großbritannien oder den USA taugten angesichts grundsätzlicher Unterschiede in den Schulsystemen kaum als Blaupause.

Brüche zeigen sich zunächst in konzeptionellen oder technischen Details, z.B.:

- Theoretische Modelle insbesondere zur Modellierung sozialer Herkunftsmerkmale von Schülerinnen und Schülern sind in den Wellen der internationalen Assessments (TIMSS, PIRLS, PISA) variiert worden mit der Folge eingeschränkter Vergleichbarkeit der Befunde. Das hat auch in der öffentlichen Diskussion von Befunden phasenweise für erhebliche Irritationen gesorgt und empirische Grundlagen der Problembeschreibungen diskreditiert.<sup>8</sup>
- Im Rahmen von PISA kamen Trendanalysen der OECD und des nationalen Konsortiums aufgrund unterschiedlicher methodischer Konzepte und unkontrollierter Varianz in den Item-Schwierigkeiten zu widersprüchlichen Befunden, was zum Teil zu öffentlich ausgetragenen Verwerfungen mit der OECD geführt hat (vgl. grundlegend zu Trendanalysen in Large-Scale-Assessments Mazzeo/von Davier 2013).
- Noch offen sind die Konsequenzen für die Durchführung und für die Entwicklung der Befundmuster, die mit einer Umstellung von PISA auf IT-basiertes Testen (*technology-based assessment*) einhergehen werden.
- Die Verknüpfung der internationalen Assessments (TIMSS, PIRLS, PISA) mit dem IQB-Ländervergleich durch ein systematisches *Equating* (Einordnung der Ergebnisse auf einer gemeinsamen Skala) ist bisher nicht systematisch durchgeführt worden mit der Folge, dass die Befunde der beiden Untersuchungsebenen (internationales Benchmarking vs. Länderbenchmarking) nicht hinreichend aufeinander bezogen werden können und in der öffentlichen Kommunikation beziehungslos nebeneinander stehen.
- Das Programm der Entwicklung von Kompetenzstufenmodellen für die IQB-Ländervergleiche ist – eher aus Kostengründen – reduziert worden. Die Bildungsstandards für den Hauptschulabschluss sind praktisch aus dem Ländervergleichs-

<sup>8</sup> Verwendet werden Modelle zum sozialen Status mit unterschiedlichen Implikationen für die Aufklärung von Leistungsvarianzen wie die EGP-Klassifikation (nach Erikson/Goldthorpe/Portocarero), der ESCS (*Economic, Social and Cultural Status*) und der HISEI (*Highest International Socio-Economic Index of Occupational Status*); siehe dazu Ehmke/Siegle 2005.

programm herausgenommen worden. Was man nach einem empirischen Erhebungsdurchgang auch über den Realismus der seinerzeit normativ gesetzten Bildungsstandards lernen könnte, ist bisher nicht ansatzweise in den Blick genommen worden.

- Noch nicht hinreichend geklärt ist der Umgang mit der zunehmenden Anzahl von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf bei den Tests in allgemeinen Schulen (Beteiligungsmodalitäten, Testkonstruktion, Berichterstattung).
- Die Durchführungsbedingungen von VERA weisen trotz eines inzwischen erreichten Minimalkonsenses<sup>9</sup> eine erhebliche Varianz zwischen den Ländern auf. Dies birgt das Risiko einer Diskreditierung der Verfahren und einer Korruption der intendierten Funktion.
- Mit dem Anspruch, VERA auch auf der Basis von Kompetenzstufenmodellen zu implementieren, wird das Risiko einer unzulässigen Verknüpfung der VERA-Befunde mit denen des IQB-Ländervergleichs eingegangen.
- Die Kalibrierung und Normierung der VERA-Tests erfolgt unter Bedingungen, die im Anschluss an den flächendeckenden Einsatz der Tests schwer händelbare Varianzen in den Befunden produzieren (unplausible Varianz über die Zeit, zwischen den Domänen bzw. Subdomänen etc.).
- Der gemeinsame Bezug des IQB-Ländervergleichs, von VERA und der länder-eigenen zentralen Prüfungsaufgaben auf ein und denselben normativen Referenzrahmen (Bildungsstandards) führt bei disparaten Befunden zu Erklärungsnotständen mit dem Risiko des Legitimationsverlustes: Wie ist bei gleichem Referenzrahmen z.B. zu erklären, dass der Anteil der Schülerinnen und Schüler, die keinen Hauptschulabschluss erhalten, weit unter dem in Assessments festgestellten Anteil der Schülerinnen und Schüler liegt, die die Mindeststandards nicht erreichen?

Die angedeuteten Probleme mögen vielleicht eher konzeptioneller oder technischer Natur sein. Kennzeichnend ist aber, dass bisher ihre Implikationen für die Integrität und Akzeptanz des Programms deutlich unterschätzt wurden.

Hinzu tritt als bildungspolitische Folge, dass der Strategiewechsel der KMK in Gestalt öffentlicher Thematisierung persistenter Problemlagen sich immer deutlicher als geeignet erwiesen hat, auch die Koordinationsleistung und -bereitschaft der Länder mit Blick auf eine zügige und effiziente Bearbeitung der Folgeprobleme an ihre Grenze zu führen. Das hängt auch damit zusammen, dass das Monitoring im internationalen oder im Ländervergleich die Verantwortungszuschreibung eindeutiger gemacht hat: Die Entstehungskontexte der Probleme sind identifizierbar geworden und liegen in der Bearbeitungshoheit der Länder, negativ ausgedrückt: Sie sind Folge ihrer politischen und administrativen Versäumnisse, für die es nun – wenn

---

9 Vgl. URL: [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2012/2012\\_03\\_08\\_Weiterentwicklung-VERA.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_03_08_Weiterentwicklung-VERA.pdf); Zugriffsdatum: 20.08.2014.

auch umstrittene – Evidenz gibt. Das bedeutet angesichts einer Ungleichverteilung der Problemlagen auf die Länder in gesamtstaatlicher Perspektive eine permanente Zerreiprobe fr die Verantwortungsgemeinschaft der KMK. Im politischen Diskurs eines Wettbewerbsfderalismus verfngt da wenig der Hinweis darauf, dass die diagnostizierten Problemlagen die Lebenschancen der nachwachsenden Generation insgesamt betreffen, also eine intergenerationelle Dimension haben und mit Langzeitkosten nachhaltig das Wirtschafts- und Sozialsystem Deutschlands insgesamt beeinflussen werden. Simpler formuliert: Dass fr Bayern oder Sachsen ein unmittelbares solidarisches Handlungsinteresse gegeben ist, wenn in Stadtstaaten im Zuge von Large-Scale-Assessments bildungsbezogene und soziale Folgeprobleme der grostdtischen Metropolenbildung sichtbar werden, ist bisher kein konsensuales Argument in der bildungspolitischen Diskussion der Lndergemeinschaft.

Obwohl das Monitoring geeignet ist, durch Sichtbarmachung der innerstaatlichen Heterogenitt, auch der Heterogenitt innerhalb der Lnder, den grundgesetzlichen Auftrag der Sicherstellung gleichwertiger Lebensverhltnisse auf die Agenda zu setzen (vgl. kritisch Block/Klemm 2005), ist es der Lndergemeinschaft in der KMK bisher gelungen, die eigentlichen „Gretchenfragen“ aus dem politischen Diskurs erfolgreich auszuklammern. Man denke z.B. an die potenziellen Folgen der Heterogenittsdiagnose fr die Anerkennung von Abschlssen, fr die Ressourcenallokation auf die Bildungsphasen oder fr einschneidende, konzertierte Interventionen zugunsten von Bildungsverliererinnen und -verlierern. Die strukturellen Grenzen der Verstndigungsbereitschaft und Koordinationsfhigkeit der Lnder, die als Schwierigkeit im Umgang mit sich allmhlich entfaltenden Implikationen des oben beschriebenen Strategiewechsels beschrieben werden knnen, zeigen sich exemplarisch in folgenden Punkten:

- Wenn naheliegende Schlussfolgerungen aus Ergebnisberichten des Monitorings finanzielle Leistungen erfordern, die nicht zu lnderspezifischen Priorittensetzungen passen, sieht sich die KMK als Lndergemeinschaft auerstande, gesamtstaatliche Verantwortung zu bernehmen. Sie verffentlicht in solchen Fllen regelmig disparate bersichtskataloge von Manahmen, die die Lnder ohnehin durchfhren (siehe z.B. die sogenannte „Frderstrategie fr Leistungsschwache“<sup>10</sup>). Eine konsistente Gesamtstrategie der Lndergemeinschaft zur Problembearbeitung wird so nicht erkennbar. Wenn es einmal zu einem lnderbergreifend abgestimmten Handlungsprogramm kommt (z.B. Bildung durch Sprache und Schrift (BISS)<sup>11</sup>), ist eine Zustimmung aller nur mglich, sofern die Lnder den Status quo ihrer ohnehin finanzierten Manahmen fortschreiben knnen und der Bund die zustzlichen Kosten im Rahmen seiner Zustndigkeit fr Ressortforschung komplett bernimmt (zur Kritik an BISS vgl. z.B. Fickermann 2013).

10 Vgl. URL: [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2010/2010\\_03\\_04-Foerderstrategie-Leistungsschwaechere.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_03_04-Foerderstrategie-Leistungsschwaechere.pdf); Zugriffsdatum: 20.08.2014.

11 Vgl. URL: <http://www.biss-sprachbildung.de/>; Zugriffsdatum: 20.08.2014.



- Ergebnisberichte des laufenden Monitorings (PISA, IQB-Ländervergleich, Bildungsberichte, demografische Analysen u.a.m.) legen auch eine sehr grundsätzliche Auseinandersetzung mit schulstrukturellen Fragen nahe, ohne diese Auseinandersetzung auf überkommene, zum Teil ritualisierte Muster der Schulformdebatte zu verengen. Auf Ebene der Ländergemeinschaft ist diese Auseinandersetzung kaum möglich, da entsprechende länderspezifische politische Festlegungen kaum folgenreich problematisierbar sind.
- In der Ländergemeinschaft gar nicht thematisiert werden Maßnahmenpakete als Schlussfolgerung aus Befunden des Bildungsmonitorings, die mit hohem Risiko politischer „Abstrafung“ verbunden sind. So sind beispielsweise Fragen der durchgängigen Akademisierung des Personals in der frühkindlichen Erziehung, der Struktur der Lehrerbesoldung und der Lehrerarbeitszeit, der Aufwertung des Grundschullehreramt, der Veränderung insgesamt der Personalstruktur an Schulen alles Themen, die – obwohl hochvirulent – nicht ernsthaft im Diskurs der Ländergemeinschaft vorangetrieben werden können. Zu zahlreich sind die finanziellen, rechtlichen, ideologischen oder standespolitischen Implikationen, als dass man sich als Ländergemeinschaft an die Bearbeitung solcher „Baustellen“ traute. Am Schicksal des Themas Lehrerbildung in der KMK lässt sich das trefflich studieren (siehe hierzu auch den Beitrag von Terhart in diesem Heft).
- Die Entwicklung vieler Indikatoren des internationalen Assessments zeigen im Zeitraum zwischen 2000 und 2009 für Deutschland positive Entwicklungstrends (mittlere Leistungskennwerte, Koppelung von Herkunft und Leistung, Anteil der Risikogruppe) (vgl. Klieme u.a. 2010). Gleichwohl hat die Ländergemeinschaft bisher keinen Weg gefunden, mit weiter bestehenden zentralen Herausforderungen in einer Langzeitperspektive wirkungsvoll umzugehen (Leseförderung als Querschnittsaufgabe, Rückstand der Jugendlichen mit Migrationshintergrund, geschlechtsspezifische Leistungsunterschiede, geringe Besetzung der Leistungsspitze). Angesichts demografischer Veränderungen werden, wenn diese Herausforderungen nicht in einem Zeitfenster der nächsten fünf bis sieben Jahre durchgreifend angegangen werden, auch die positiven Entwicklungstrends im internationalen Vergleich wieder verspielt werden.

Die skizzierte Gemengelage zeigt: Bildung als intergenerationelle, gesamtgesellschaftliche Aufgabe stößt im Kontext bildungsföderaler staatlicher Verfasstheit auf deutliche Grenzen der Bearbeitbarkeit, jedenfalls so lange, wie – anders als in anderen Staaten – nicht-staatliches, öffentliches, zivilgesellschaftliches Engagement in Bildungsfragen so unterentwickelt und unorganisiert bleibt wie hierzulande. Zugleich erhöht der einmal eingeschlagene Weg eines Bildungsmonitorings gleichsam das Risiko permanenter Sichtbarkeit und Kritisierbarkeit dieses Skandalons.

### 3. Wahrnehmung und Nutzung des Monitoring-Wissens

Die wissenschaftlich-technologische Entfaltung des Bildungsmonitorings hat trotz der angedeuteten konzeptionellen Probleme auf (inter-)nationaler Ebene, zum Teil auch auf Länderebene ein beachtliches Niveau erreicht. Die Verwendungspraktiken für die so generierte Evidenz bleiben – das muss man feststellen – dahinter zurück. Es führt kein direkter Weg vom Messen zum Gestalten, weder in der Politik noch im Unterricht. Zu unterschiedlich sind die jeweils erforderlichen Handlungslogiken. Zugespitzt formuliert: Die Problemlösekapazität steht auf allen Ebenen – oben wie unten – in deutlichem Missverhältnis zum Problemwissen.

Kennzeichnend für die öffentliche Wahrnehmung ist

- zum einen – positiv – die Tatsache, dass – unterstützt durch einen deutlich weiterentwickelten Bildungsjournalismus – der Bildungsbereich als solcher und insbesondere sichtbar werdende Problemlagen überhaupt Gegenstand öffentlicher Rezeption und Debatten, zum Teil auch der medialen Skandalisierung geworden sind;
- zum anderen – eher kritisch – die Tatsache, dass eine Fokussierung auf Rankings im Staaten- oder Ländervergleich die Diskursstrukturen und Reaktionsmuster der Akteure zu deutlich bestimmt.

Zu beiden genannten Aspekten haben politische Akteure nicht selten ein gebrochenes Verhältnis. Die Befunde des Monitorings sind so zahlreich und uneinheitlich, dass sie ohne Schwierigkeiten selektiv rezipiert und zu Legitimationszwecken in die je spezifische politische Agenda eingeschrieben werden können. Rankings, deren Problematik und Grenzen hinreichend bekannt sind, übernehmen die Funktion, im politischen Wettbewerb jenseits sachlicher Triftigkeit eine anschauliche, leicht kommunizierbare Ordnung von Gewinnern und Verlierern zu erzeugen. Nicht erstaunlich ist, dass Dauerverlierer Rankings problematisieren, Dauergewinner sie zelebrieren und beide auf diese Weise die Gemengelage im politischen Diskurs stabilisieren. Das wird so bleiben trotz der Vorschläge, zu anderen Formen von Vergleichen zu kommen (Stadt vs. Land, Metropolenvergleich, Ländergruppen, Regionen u.a.m.).

Inwieweit Verfahren und Ergebnisse des Monitorings auch Bildungsadministrationen tendenziell überfordern, kann man an immer noch vorherrschenden Rezeptionsstrukturen und Reaktionsmustern anlässlich der Veröffentlichung von Vergleichsuntersuchungen studieren. Diese Strukturen und Muster sind geprägt durch die Dominanz einer festgelegten Form des „Tagesgeschäfts“, zu der wissenschaftlich-empirisch generierte Erkenntnisse in aller Regel einen nur schwer überbrückbaren Erklärungsabstand haben. Nicht selten führt dies in Bildungsverwaltungen zu dem Kuriosum, dass von den Monitoring-Verantwortlichen auch noch die Formulierung

von bildungsplanerischen Konsequenzen erwartet wird. Wenn sie es dann aber tun, wird dies als unzulässiger Übergriff abgewehrt.

Die Nutzung von Ergebnissen der Assessments auf Schul- und Unterrichtsebene ist ebenso voraussetzungsreich. Es gibt dazu mit der „Konzeption der Kultusministerkonferenz zur Nutzung der Bildungsstandards für die Unterrichtsentwicklung“<sup>12</sup> Überlegungen der Ländergemeinschaft und auch vier gemeinsame Vorhaben nach dem Ende der BLK<sup>13</sup> (vgl. zur Implementation von Bildungsstandards auch die umfangreiche Expertise von Oelkers und Reusser 2008). Im Zuge der Implementierung der Gesamtstrategie hat sich aber erst allmählich eine Vorstellung davon entwickelt, dass die Funktionen der unterschiedlichen Assessments mit Folgen für die Anlage und Durchführung der Tests, für die Aufbereitung und Rückkopplung der Ergebnisse, für die Zugänglichkeit von Daten usw. sauber unterschieden werden müssen. Die Grundfunktion von VERA 3 und VERA 8 als Instrument der Schul- und Unterrichtsentwicklung erfordert eine systematische Fokussierung der Instrumente und Verfahren auf diesen Zweck, d.h. auch eine Zurückhaltung bei der konzeptionellen Anbindung an den IQB-Ländervergleich oder bei der Überfrachtung mit Ansprüchen des Systemmonitorings. Bis dato ist die Positionierung der Länder in dieser Hinsicht weder klar noch einheitlich, was in einzelnen Ländern zu erheblichen Akzeptanzproblemen geführt hat.

Von unmittelbarer Relevanz für den Aspekt der Nutzung von Large-Scale-Assessments ist die Frage, welche Art von Wissen mit ihnen generiert wird. In der politischen Diskussion ist die Frage ebenfalls virulent; gelegentlich wird bereits so getan, als sei sie beantwortet: PISA und Co. produzierten allenfalls Beschreibungswissen. Exemplarisch kann der Antrag der Hamburgischen SPD-Bürgerschaftsfraktion „Ursachenermittlung statt Länder-Bashing“ (Drs. 20/9119) zitiert werden:

„Unbestreitbar liefern nationale Bildungsstudien wichtige Ergebnisse. Sie zeigen auf, wo Schülerinnen und Schüler im nationalen Vergleich stehen. Offen lassen sie leider, wo genau die Ursachen für das gute oder schlechte Abschneiden der Schülerinnen und Schüler eines Bundeslandes zu verorten und welche schulpolitischen Verbesserungen vorzunehmen sind. Man weiß, dass Stadtstaaten in der

---

12 Vgl. URL: [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2010/2010\\_00\\_00-Konzeption-Bildungsstandards.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2010/2010_00_00-Konzeption-Bildungsstandards.pdf); Zugriffsdatum: 19.09.2014.

13 – ProLesen – Auf dem Weg zur Leseschule; vgl. URL: <http://www.leseforum.bayern.de/index.asp?MNav=6>; Zugriffsdatum: 20.08.2014.

– Udikom – Aus- und Fortbildung der Lehrkräfte in Hinblick auf Verbesserung der Diagnosefähigkeit, Umgang mit Heterogenität, individuelle Förderung; vgl. URL: <http://www.kmk-udikom.de/>; Zugriffsdatum: 20.08.2014.

– for.mat – Bereitstellung von Fortbildungskonzeptionen und -materialien zur kompetenz- bzw. standardbasierten Unterrichtsentwicklung; vgl. URL: <http://www.kmk-format.de/>; Zugriffsdatum: 20.08.2014.

– EMU – Evidenzbasierte Methoden der Unterrichtsdiagnostik und -entwicklung. URL: <http://www.unterrichtsdiagnostik.info>; Zugriffsdatum: 20.08.2014.

Regel schlechter abschneiden als Flächenländer. Es gibt jedoch keine Erkenntnisse darüber, welche Faktoren dafür ursächlich sind und wie genau sie sich auswirken.“

Das ist richtig. Kennzeichnend für die politische Diskussion ist jedoch, dass Large-Scale-Assessments (gemeint sind eigentlich Rankings) grundsätzlich problematisiert werden oder Monitoring und „Ursachenermittlung“ gegeneinander ausgespielt werden. Insofern man sich in der Politik unter letzterem in Zeiten der Haushaltsnot kostengünstigere Forschung vorstellt, könnte sich das als Trugschluss erweisen. Auch die Vermeidung von Legitimationsproblemen durch Wegfall von Rankings kann sich durch einen strategischen Schwenk zur „Ursachenermittlung“ als illusionär erweisen, steigt doch mit genauerer Kenntnis von Ursachen auch der Druck, problemadäquat und -spezifisch zu handeln.

Gängig geworden ist die Unterscheidung von Diagnosewissen, Erklärungswissen und Handlungswissen. Bildungsmonitoring produziert in aller Regel Diagnosewissen für unterschiedliche Ebenen des Bildungssystems. Damit bleibt aber tendenziell das Bedürfnis von Bildungsakteuren, seien es Politiker und Politikerinnen oder Lehrkräfte, auf der Strecke, für diagnostizierte Zustände auch Wissen zu erhalten, das diese Zustände erklären kann, und ferner auch Wissen, das professionelle praktische Bearbeitungsstrategien für die festgestellten Zustände anleiten kann (Professionswissen). Wissen des Bildungsmonitorings bewährt sich in diesem Sinne in Entscheidungspraktiken nur dann, wenn es seiner Funktionalität als Diagnosewissen gerecht wird, wenn es die immer begrenzte Verarbeitungskapazität für Wissenskomplexität von Wissensnutzern und -nutzerinnen berücksichtigt, aber auch, indem es sich als anschlussfähig an andere Wissensformen mit praktischer Relevanz für Wissensnutzer und -nutzerinnen erweist. Alle Ebenen des Bildungssystems brauchen Diagnosewissen, um Stärken und Schwächen zu identifizieren, brauchen Erklärungswissen, um das eigene professionelle Handeln zu reflektieren und weiterzuentwickeln, brauchen Handlungswissen und ‚Tools‘, um passende, bewährte Maßnahmen einzuleiten und durchzuführen. Bei einer Weiterentwicklung der Strategie des Bildungsmonitorings wäre diesem komplexen Verständnis von Wissen Rechnung zu tragen.

#### **4. Weiterentwicklung der Monitoringstrategie**

Wie kann es weitergehen angesichts der skizzierten Voraussetzungen und Restriktionen? Für eine Reformulierung der Gesamtstrategie gibt es, sofern das Assessmentprogramm nicht grundsätzlich in Frage gestellt wird, unterschiedlich weit gehende Optionen. Für den Diskussionsprozess wird grundsätzlich wichtig sein, unterschiedliche Argumentationsebenen, die immer gleichzeitig virulent sind, deutlich zu unterscheiden:

- die Ebene der konzeptionellen Schlüssigkeit des Monitoringprogramms,
- die Ebene der Finanzierbarkeit des Monitoringprogramms,
- die Ebene der politischen Opportunitäten im Umgang mit Befunden des Monitorings,
- die Ebene der administrativen Verarbeitbarkeit der Durchführung und Auswertung des Monitorings.

Ob es gelingen wird, den Diskussions- und Entscheidungsfindungsprozess in der Ländergemeinschaft so anzulegen, dass die auf den genannten Ebenen liegenden Motive für eine Weiterentwicklung der Gesamtstrategie immer klar sind, wird man sehen.

Betrachtet man die Fragen pragmatisch, scheint zunächst eine grundsätzliche Änderung der Beteiligung an internationalen Assessments wenig wahrscheinlich, weil sie aufgrund der Kooperation mit dem Bund, der internationalen Einbindung sowie der Institutionalisierung im ZIB fest verankert sind. Ob Deutschland in den internationalen Steuerungsgremien nachhaltig Einfluss auf die grundsätzliche Anlage der Erhebungsprogramme nehmen kann, ist wenig wahrscheinlich. Für den IQB-Ländervergleich gilt im Grundsatz das gleiche, wenngleich die Länder hier souveräner in der Gestaltung sind. Politisch dürfte eine grundsätzliche Abkehr von einer rankingorientierten Berichterstattung auch auf nationaler Ebene kaum durchsetzbar sein. Ihr instruktiver Symbolgehalt ist im kompetitiven „Polit-Spiel“ um „richtige Lösungen“ wahrscheinlich unverzichtbar. Viel gewonnen wäre schon, wenn die Berichterstattung künftig stärker auf das Erreichen der Bildungsstandards, also auf kriterial referenzierte Verteilungsmuster fokussierte. Veränderungsspielräume dürften u.U. auch gegeben sein bei der Gestaltung der Erhebungsintervalle (z.B. zeitliche Streckung in der Sekundarstufe I auf 5 Jahre) oder bei der Fokussierung der Testdomänen.

Gründlich reflektiert werden muss die konzeptionelle und strategische Anlage von VERA. Hier kann man grundsätzlich die Frage stellen, ob es nötig ist, dass die Länder dieses Instrument der Schul- und Unterrichtsentwicklung durch länderübergreifende aufwändige Abstimmungsprozeduren mit dem Anspruch überfrachten, Baustein einer nationalen Monitoringstrategie zu sein, zumal das IQB sich immer weniger in der Lage sieht, übergreifende Qualitätssicherungsaufgaben wahrzunehmen und zu finanzieren. Eine solche Entwicklungsrichtung liefe auf eine stärkere konzeptionelle Abkoppelung vom Paradigma hinaus, das mit dem IQB-Ländervergleich gegeben ist, und ermöglichte länderspezifische Ausdifferenzierungen, verlangte aber auch von den Ländern Investitionen in den Ausbau eigener, stärker formativer Testsysteme. Das Szenario birgt deshalb auch das Risiko, dass einzelne Länder sich ganz von schul- und unterrichtsnahen, entwicklungsorientierten Assessments verabschieden. Will man nicht so weit gehen, erscheint gleichwohl eine Weiterentwicklung in Richtung längs-

schnittlich angelegter und für Lehrkräfte instruktiverer Test- und Rückmeldeformate nötig.

Will man gerade die Entwicklungsfunktion datengestützter Rückmeldungen stärken, sind die schulbezogenen Ergebnisdaten von VERA zum einen durch die Rückmeldung von Daten zu innerschulischen Prozessen und durch systematische Bereitstellung von einzelschulischen Kontextdaten (Zusammensetzung der Schülerschaft, Spezifika des schulischen Umfeldes usw.) zu ergänzen. Zu überlegen ist deshalb, den Verfahren, die diese Informationen liefern (z.B. Schulinspektionen und amtliche Statistiken), einen besonderen Platz im Rahmen der Gesamtstrategie zu geben. Zum anderen sind erhebliche Anstrengungen erforderlich, um das Nutzenpotenzial eines so erweiterten Monitoringkonzepts zu identifizieren und mit Professionalisierungs- und Unterstützungsmaßnahmen für Schulen auf allen Ebenen zum Tragen zu bringen. Auf diese Art und Weise würde die Entwicklungsfunktion des Monitorings gestärkt.

Zentrales Desiderat ist also eine Weiterentwicklung der Gesamtstrategie, die geeignet ist, künftig verstärkt Handlungswissen bereitzustellen. Bildungspolitik, Bildungsadministration und natürlich auch Lehrkräfte benötigen neben ihrer professions- und erfahrungsbezogenen zunehmend mehr wissenschaftliche „Evidenz“, um rational entscheiden zu können. Die Umsetzung des Steuerungsparadigmas „evidenz- bzw. evaluationsbasiert“ muss auch für die Anlage von Forschungsvorhaben Konsequenzen haben. Es gilt stärker als bisher „praktische“ Probleme in den Blick zu nehmen, diese auch in der Praxis mit hinreichend großen Stichproben zu untersuchen und so Wissen zu generieren, das tatsächlich dazu beiträgt, die Praxis zu verbessern bzw. Qualität zu entwickeln und zu sichern. Es müssen zudem verstärkt auch solche Forschungsvorhaben gefördert werden, die Entwicklungsarbeiten in einen Forschungskontext mit den üblichen Standards empirischer Bildungsforschung (experimentelles Design mit einer Zufallsauswahl der Beteiligten und einer Kontrollgruppe etc.) stellen. Das Rahmenprogramm des BMBF zur Förderung der empirischen Bildungsforschung ist eine gute Grundlage für eine strategisch ausgerichtete Forschungsförderung, die – in Abstimmung mit dem BMBF – in der Lage ist, den Interessen der Länder an Handlungswissen nachzukommen. Es würde sich also in Ergänzung zu einer eventuellen Reformulierung der Gesamtstrategie zum Bildungsmonitoring anbieten, als KMK für eine Wiederbelebung des zuletzt reduzierten Rahmenprogramms zu werben.

## Literatur und Internetquellen

- Artelt, C./McElvany, N./Christmann, U./Richter, T./Groeben, N./Köster, J./Schneider, W./Stanat, P./Ostermeier, C./Schiefele, U./Valtin, R./Ring, K. (2005): Förderung von Lesekompetenz (Bildungsforschung, Band 17). Bonn/Berlin: BMBF. URL: [http://www.bmbf.de/pub/bildungsreform\\_band\\_siebzehn.pdf](http://www.bmbf.de/pub/bildungsreform_band_siebzehn.pdf); Zugriffsdatum: 20.08.2014.
- Baethge, M./Buss, K.-P./Lanfer, C. (2003): Konzeptionelle Grundlagen für einen Nationalen Bildungsbericht: Berufliche Bildung und Weiterbildung/Lebenslanges Lernen (Bildungsforschung, Band 7). Bonn/Berlin: BMBF.
- Baethge, M./Buss, K.-P./Lanfer, C. (Hrsg.) (2004): Expertisen zu den konzeptionellen Grundlagen für einen Nationalen Bildungsbericht – Berufliche Bildung und Weiterbildung/Lebenslanges Lernen (Bildungsforschung, Band 8). Bonn/Berlin: BMBF. URL: [http://www.bmbf.de/pub/expertisen\\_zd\\_konzept\\_grundlagen\\_fn\\_bildungsbericht\\_bb\\_wb\\_III.pdf](http://www.bmbf.de/pub/expertisen_zd_konzept_grundlagen_fn_bildungsbericht_bb_wb_III.pdf); Zugriffsdatum: 20.08.2014.
- Benner, D. (Hrsg.) (2007): Bildungsstandards. Chancen und Grenzen, Beispiele und Perspektiven. Paderborn: Schöningh.
- Block, R./Klemm, K. (2005): Gleichwertige Lebensverhältnisse im Bundesgebiet? Demografische, ökonomische, institutionelle und familiäre Bedingungen des Lernens im Bundesländervergleich. Essen. URL: <https://www.gew-bw.de/Binaries/Binary4061/Klemm-Studie.pdf>; Zugriffsdatum: 20.08.2014.
- Ehlich, K., in Zusammenarbeit mit Bredel, U., Garne, B., Komor, A., Krumm, H.-J., McNamara, T., Reich, H.H., Schnieders, G., ten Thile, J.D. und Van den Bergh, H. (2005): Anforderungen an Verfahren der regelmäßigen Sprachstandserhebung als Grundlage für die frühe und individuelle Förderung von Kindern mit und ohne Migrationshintergrund (Bildungsforschung, Band 11). Bonn/Berlin: BMBF. URL: [http://www.bmbf.de/pub/bildungsreform\\_band\\_elf.pdf](http://www.bmbf.de/pub/bildungsreform_band_elf.pdf); Zugriffsdatum: 20.08.2014.
- Ehlich, K./Bredel, U./Reich, H.H. (Hrsg.), in Zusammenarbeit mit Falk, S., Guckelsberger, S., Kemp, R.F., Knopp, M., Komor, A., Landua, S., Maier-Lohmann, C., Mihaylov, V., Sirim, E., Soutanian, N. und Trautmann, C. (2008a): Referenzrahmen zur altersspezifischen Sprachaneignung (Bildungsforschung, Band 29/I). Bonn/Berlin: BMBF. URL: [http://www.bmbf.de/pub/bildungsforschung\\_band\\_neunundzwanzig.pdf](http://www.bmbf.de/pub/bildungsforschung_band_neunundzwanzig.pdf); Zugriffsdatum: 20.08.2014.
- Ehlich, K./Bredel, U./Reich, H.H. (Hrsg.), in Zusammenarbeit mit Falk, S., Guckelsberger, S., Kemp, R.F., Knopp, M., Komor, A., Landua, S., Maier-Lohmann, C., Mihaylov, V., Sirim, E., Soutanian, N. und Trautmann, C. (2008b): Referenzrahmen zur altersspezifischen Sprachaneignung – Forschungsgrundlagen (Bildungsforschung, Band 29/II). Bonn/Berlin: BMBF. URL: [http://www.bmbf.de/pub/bildungsforschung\\_bd\\_neunundzwanzig\\_zwei.pdf](http://www.bmbf.de/pub/bildungsforschung_bd_neunundzwanzig_zwei.pdf); Zugriffsdatum: 20.08.2014.
- Ehmke, T./Siegle, T. (2005): ISEI, ISCED, HOMEPOS, ESCS: Indikatoren der sozialen Herkunft bei der Quantifizierung von sozialen Disparitäten. In: Zeitschrift für Erziehungswissenschaft 8, H. 4, S. 521-539.
- Fickermann, D. (2013): Maßnahmen zur Verbesserung der Effizienz und der Effektivität der (vor-)schulischen Sprachförderung. In: Die Deutsche Schule 105, H. 2, S. 169-184.
- Fickermann, D. (2014): Einrichtungen zur Qualitätssicherung und -entwicklung als „nachgeordnete Dienststellen besonderer Art“. In: Die Deutsche Schule 106, H. 3, S. 231-239.
- Fthenakis, W.E./Gisbert, K./Griebel, W./Kunze, H.-R./Niesel, R./Wustmann, C. (2005): Auf den Anfang kommt es an. Perspektiven für eine Neuorientierung frühkindlicher Bildung (Bildungsforschung, Band 16). Bonn/Berlin: BMBF. URL: [http://www.bmbf.de/pub/bildungsreform\\_band\\_16.pdf](http://www.bmbf.de/pub/bildungsreform_band_16.pdf); Zugriffsdatum: 20.08.2014.
- Hericks, U. (2004): Grundbildung, Allgemeinbildung und Fachunterricht. In: Zeitschrift für Erziehungswissenschaft 7, H. 2, S. 192-206.

- Klieme, E./Avenarius, H./Blum, W./Döbrich, P./Gruber, H./Prenzel, M./Reiss, K./Riquarts, K./Rost, J./Tenorth, H.-E./Vollmer H.-J. (2003): Zur Entwicklung nationaler Bildungsstandards. Expertise (Bildungsforschung, Band 1). Bonn/Berlin: BMBF. URL: [http://www.bmbf.de/pub/zur\\_entwicklung\\_nationaler\\_bildungsstandards.pdf](http://www.bmbf.de/pub/zur_entwicklung_nationaler_bildungsstandards.pdf); Zugriffsdatum: 20.08.2014.
- Klieme, E./Jude, N./Baumert, J./Prenzel, M. (2010): PISA 2000-2009. Bilanz der Veränderungen im Schulsystem. In: Klieme, E./Artelt, C./Hartig, J./Jude, N./Köller, O./Prenzel, M./Schneider, W./Stanat, P. (Hrsg.): PISA 2009. Bilanz nach einem Jahrzehnt. Münster u.a.: Waxmann, S. 277-300. URL: [http://www.pedocs.de/volltexte/2011/3539/pdf/Klieme\\_et.Al\\_Bilanz\\_D\\_A.pdf](http://www.pedocs.de/volltexte/2011/3539/pdf/Klieme_et.Al_Bilanz_D_A.pdf); Zugriffsdatum: 20.08.2014.
- Koch, L. (2004): Allgemeinbildung und Grundbildung, Identität oder Alternative? In: Zeitschrift für Erziehungswissenschaft 7, H. 2, S. 183-191.
- Mazzeo, J./von Davier, M. (2013): Linking Scales in International Large-Scale Assessments. In: Rutkowski, D. (Hrsg.): Handbook of International Large-Scale Assessment. Background, Technical Issues, and Methods of Data Analysis. London: Chapman Hall/CRC Press, S. 229-258.
- Oelkers, J./Reusser, K., unter Mitarbeit von Berner, E., Halbheer, U. und Stolz, S. (2008): Qualität entwickeln – Standards sichern – mit Differenz umgehen (Bildungsforschung, Band 27). Bonn/Berlin: BMBF. URL: [http://www.bmbf.de/pub/bildungsforschung\\_band\\_siebenundzwanzig.pdf](http://www.bmbf.de/pub/bildungsforschung_band_siebenundzwanzig.pdf); Zugriffsdatum 20.08.2014.
- Rauschenbach, T./Leu, H.R./Lingenauber, S./Mack, W./Schilling, M./Schneider, K./Züchner, I. (2004): Non-formale und informelle Bildung im Kindes- und Jugendalter. Konzeptionelle Grundlagen für einen nationalen Bildungsbericht (Bildungsforschung, Band 6). Bonn/Berlin: BMBF.
- Redder, A./Schwippert, K./Hasselhorn, M./Forschner, S./Fickermann, D./Ehlich, K. sowie Becker-Mrotzek, M./Krüger-Potratz, M./Rossbach, H.-G./Stanat, P./Weinert, S., unter Mitarbeit von Kulik, M., Worgt, M. und Zech, C. (2011): Bilanz und Konzeptualisierung von strukturierter Forschung zu „Sprachdiagnostik und Sprachförderung“ (ZUSE-Berichte, Bd. 2). Hamburg: ZUSE. URL: [http://www.dabil.de/zuse/501publikation/zuse\\_berichte\\_02.pdf](http://www.dabil.de/zuse/501publikation/zuse_berichte_02.pdf); Zugriffsdatum: 20.08.2014.
- Roßbach, H.G./Weinert, S. (Hrsg.) (2008): Kindliche Kompetenzen im Elementarbereich: Förderbarkeit, Bedeutung und Messung (Bildungsforschung, Band 24). Bonn/Berlin: BMBF. URL: [http://www.bmbf.de/pub/bildungsforschung\\_band\\_vierundzwanzig.pdf](http://www.bmbf.de/pub/bildungsforschung_band_vierundzwanzig.pdf); Zugriffsdatum: 20.08.2014.
- Rürup, M. (2014): Landesinstitute für Qualität als neue Akteure im deutschen Schulwesen? Governancetheoretische Annäherungen an eine Alltagstheorie. In: Die Deutsche Schule 106, H. 3, S. 201-218.
- Tenorth, H.-E. (2004a): Stichwort: „Grundbildung“ und „Basiskompetenzen“. Herkunft, Bedeutung und Probleme im Kontext allgemeiner Bildung. In: Zeitschrift für Erziehungswissenschaft 7, H. 2, S. 169-182.
- Tenorth, H.-E. (2004b): Bildungsstandards und Kerncurriculum. Systematischer Kontext, bildungstheoretische Probleme. In: Zeitschrift für Pädagogik 50, H. 5, S. 650-661.
- Tillmann, K.-J./Dederling, K./Kneuper, D./Kuhlmann, C./Nessel, I. (2008): PISA als bildungspolitisches Ereignis. Fallstudien in vier Bundesländern. Wiesbaden: VS.

*Norbert Maritzen*, geb. 1952, Direktor des Instituts für Bildungsmonitoring und Qualitätsentwicklung (IfBQ), Hamburg.

Anschrift: IfBQ, Beltgens Garten 25, 20537 Hamburg  
E-Mail: [Norbert.Maritzen@ifbq.hamburg.de](mailto:Norbert.Maritzen@ifbq.hamburg.de)