



27 Qualität entwickeln – Standards sichern – mit Differenz umgehen

Bildungsforschung Band 27

Qualität entwickeln - Standards sichern – mit Differenzen umgehen

Impressum

Herausgeber

Bundesministerium
für Bildung und Forschung (BMBF)
Referat Bildungsforschung
11055 Berlin

Bestellungen

Schriftlich an den Herausgeber
Postfach 30 02 35
53182 Bonn

oder per

Tel.: 01805-262 302

Fax: 01805-262 303

(0,14 Euro/Min. aus dem deutschen
Festnetz)

E-Mail: books@bmbf.bund.de

Internet: <http://www.bmbf.de>

Bonn, Berlin 2008

Autoren:

Prof. Dr. Jürgen Oelkers,
Pädagogisches Institut der Universität Zürich
Prof. Dr. Kurt Reusser,
Pädagogisches Institut der Universität Zürich

Beiträge

Esther Berner, Pädagogisches Institut der Universität Zürich
Uli Halbheer, Pädagogisches Institut der Universität Zürich
Stefanie Stolz, Pädagogisches Institut der Universität Zürich

Jürgen Oelkers/Kurt Reusser

Unter Mitarbeit von Esther Berner,

Ueli Halbheer & Stefanie Stolz

Expertise:

Qualität entwickeln –

Standards sichern –

mit Differenz umgehen

Januar 2008

Inhalt

Vorwort

1	Auftrag, Vorgehen und Ziele der Expertise	11
2	Entwicklung und Diskussion der zentralen Konzepte	17
2.1	Historische Erfahrungen mit Qualitätssicherung in Schulen	18
2.2	Kompetenz als Schlüsselkategorie	23
2.3	Evaluationen und Accountability	29
2.4	Konzepte von Bildungsstandards	36
2.5	Implementation und Schulentwicklung	45
2.6	Kritik an Standards und Tests	52
3	Stand der Forschung und Entwicklung in vier ausgewählten Ländern	61
3.1	Vereinigte Staaten	66
3.1.1	Der Wandel der Bildungspolitik	67
3.1.2	Verständnis und Bedeutung von Standards	76
3.1.3	Überprüfung von Standards: Tests und Accountability	80
3.1.4	Erfahrungen mit der Implementation von Standards	91
3.1.5	Theoretische Ansätze der Implementation	103
3.1.6	Veränderungen und Problemfelder der Lehrerbildung	114
3.2	Niederlande	130
3.2.1	Der Wandel der Bildungspolitik	132
3.2.2	Verständnis und Bedeutung von Standards	135
3.2.3	Überprüfung von Standards: Tests und Prüfungen	139
3.2.4	Qualitätssicherung und Evaluation	141
3.3	Schweden	148
3.3.1	Der Wandel der Bildungspolitik	150
3.3.2	Verständnis und Bedeutung von Standards	159
3.3.3	Überprüfung von Standards: Tests und Prüfungen	162
3.3.4	Qualitätssicherung und Evaluation	165
3.4	England	175
3.4.1	Der Wandel der Bildungspolitik	176
3.4.2	Verständnis und Bedeutung von Standards	196
3.4.3	Überprüfung von Standards: Tests und Prüfungen	205
3.4.4	Qualitätssicherung und Evaluation	210
3.5	Auswertung des Ländervergleichs	222

4	Verständnis, Strategien, Akteure, Ebenen und Kontexte der Implementation von Bildungsstandards	231
4.1	Begriff und theoretisches Verständnis einer Implementation von Bildungsstandards	234
4.2	Leitideen und Strategien einer neuen Qualitätspolitik	240
4.2.1	Anforderungen und Leitideen	240
4.2.2	Strategien der Einführung und Sicherung von Bildungsstandards	250
4.3	Ebenen und Akteure	270
4.3.1	Nationale Ebene: Bund, Länder und Kantone	272
4.3.2	Regionale Vernetzung der Schulentwicklung	277
4.3.3	Einzelne Schulen	279
4.3.4	Akteure: Schulleitungen, Lehrkräfte, Schüler und Eltern	281
4.3.5	Werkzeuge und Verfahren	284
4.4	Kontext- und Unterstützungssysteme	285
4.4.1	Aus- und Weiterbildung der Lehrkräfte	287
5	Werkzeuge und Verfahren der Implementation	303
5.1	Systemebene	305
5.1.1	Ausgearbeitete Bildungsstandards und darauf bezogene Tests	306
5.1.2	Bildungsmonitoring	319
5.1.3	Rückmeldung von Ergebnissen	326
5.1.4	Rahmenlehrpläne, Kerncurricula, Lehrmittel	335
5.1.5	ICT-Werkzeuge, Schul- und Bildungsserver	346
5.1.6	Schulaufsicht und Schulinspektion	352
5.1.7	Reformprojekte und Modellversuche	358
5.2	Schulebene	366
5.2.1	Schulleitung	369
5.2.2	Schulprogrammarbeit	375
5.2.3	Integrierte Selbst- und Fremdevaluation	378
5.2.4	Kooperation der Lehrkräfte in Fachgruppen und Netzwerken	390
5.3	Unterrichts- und Klassenebene	399
5.3.1	Unterrichtsentwicklung als Kernaufgabe	402
5.3.2	Lehrmittel und Aufgabenkulturen	408
5.3.3	Abrufbare Fach- und Leistungstests	412
5.3.4	Fachspezifisch-pädagogisches Coaching	417

5.3.5	Videobasierte Analyse und Reflexion von Unterricht	421
5.3.6	Lehren und Lernen – Perspektive und Verantwortung der Schülerinnen und Schüler	426
6	Stand der Diskussion und Entwicklung in den Auftragsländern	431
6.1	Deutschland	433
6.1.1	Diskussionen und Problemwahrnehmungen	433
6.1.2	Konzeption und Funktion von Bildungsstandards	435
6.1.3	Überprüfung der Leistungen: Tests und Bildungsmonitoring	447
6.1.4	Aktueller Stand und Strategien der Umsetzung von Bildungsstandards	453
6.2	Österreich	463
6.2.1	Diskussionen und Problemwahrnehmungen	463
6.2.2	Konzeption und Funktion von Bildungsstandards	466
6.2.3	Überprüfung der Leistungen: Tests und Bildungsmonitoring	472
6.2.4	Aktueller Stand und Strategien der Umsetzung von Bildungsstandards	474
6.3	Schweiz	476
6.3.1	Diskussionen und Problemwahrnehmungen	476
6.3.2	Konzeption und Funktion von Bildungsstandards	479
6.3.3	Überprüfung der Leistungen: Tests und Bildungsmonitoring	481
6.3.4	Aktueller Stand und Strategien der Umsetzung von Bildungsstandards	482
6.4	Luxemburg	485
6.4.1	Diskussionen und Problemwahrnehmungen	485
6.4.2	Konzeption und Funktion von Bildungsstandards	487
6.4.3	Überprüfung der Leistungen: Tests und Bildungsmonitoring	489
6.4.4	Aktueller Stand und Strategien der Umsetzung von Bildungsstandards	490
7	Zusammenfassung und Schlussfolgerungen	495
7.1	Zusammenfassung zentraler Aussagen der Expertise	495
7.2	Schlussfolgerungen	509
	Literaturverzeichnis	523

Vorwort

Bildungsstandards sind seit etwa fünf Jahren Thema einer lang anhaltenden öffentlichen Diskussion in Deutschland, Österreich, der Schweiz und Luxemburg. Die Diskussion ist kontrovers, was auch damit zu tun hat, dass über Bildungsstandards konkret wenig bekannt ist und der Begriff Assoziationen weckt, die eine „Standardisierung“ nach Art einer Industrienorm nahe legen. Daher bestehen vor allem in der Lehrerschaft Befürchtungen, mit der Einführung von Standards würden bestehende Freiheiten beschnitten oder eine unnötige Regulierung werde angestrebt. Auf der anderen Seite gelten Bildungsstandards leicht als Allheilmittel, dem die Lösung aller Probleme zugetraut wird. Wenigstens ist auffällig, dass auf Fragen, wie die weitere Entwicklung des Bildungswesens gestaltet werden soll, oft unisono mit „Standards“ geantwortet wird.

In dieser Situation hilft zunächst ein Blick ins Ausland. Es gibt im europäischen Vergleich ganz unterschiedliche Versuche, im Bereich der öffentlichen Schulen Bildungsstandards zu entwickeln, eher zurückhaltende wie in Schweden oder eher forcierte wie in England, aber der Trend ist unverkennbar, die Erwartungen an das Bildungssystem zu präzisieren und die Steuerung auch auf die Ergebnisse einzustellen. Die klassische Form der Bildungssteuerung ist weitgehend nur auf den Input eingestellt, sie betrifft Vorschriften und Ressourcen, vernachlässigt aber die Resultate, soweit darunter Mehr und Anderes verstanden wird als Noten und Zeugnisse. Was „Outputsteuerung“ genannt wird, ist eine Konsequenz der Einführung von Bildungsstandards und inzwischen in vielen europäischen Ländern bildungspolitische Realität oder ernsthafte Entwicklungsaufgabe.

Vor allem amerikanische Forschungen verweisen auch auf Gefahren und Risiken dieser Entwicklung, die oft unter dem Stichwort „high stakes testing“ diskutiert werden. Bildungsstandards sind mit Leistungstests verbunden, die zu verschiedenen Zwecken eingesetzt werden können. Mit solchen Tests gibt es in Zentraleuropa wenige Erfahrungen, aber ohne Tests kann es eine Outputsteuerung kaum geben. Darunter ist mehr zu verstehen als nur die regelmäßige Erhebung von Testdaten, aber nur *mit* solchen Daten lässt sich die Steuerung des Bildungssystems umstellen. Neu sind also nicht nur Bildungsstandards im Sinne präzisierter und transparenter Leistungserwartungen, sondern auch bestimmte Formen der Leistungsüberprüfung.

Die nachstehende Expertise geht davon aus, dass die Implementation von Bildungsstandards als Prozess der *Schulentwicklung* verstanden werden muss. Bildungsstandards sind keine isolierte Größe; wer auf Standards und neue Formen der Leistungsüberprüfung setzt, hat ein Praxisfeld vor sich, das für die Entwicklung gewonnen werden will. Daher vertritt die Expertise, basierend auf den Ergebnissen der Implementationsforschung, einen auf die Akteure bezogenen Ansatz. Wandel setzt *sense-making* voraus, verlangt aber auch dafür geeignete *Instrumente*. Hier liegt die zweite Botschaft der Expertise: Ohne neue Tools wird es nicht möglich sein, das Praxisfeld wirkungsvoll in Richtung Bildungsstandards zu verändern.

Eine Steuerung allein über den Output ist jedoch nicht ausreichend. Die Daten müssen in das Praxisfeld rückübersetzt werden, aus „Output“, anders

gesagt, müssen wieder „Input“ und qualitätsvolle Prozesse werden. Hier liegt ein dritter Punkt. Ohne gezielten Aufbau von neuartigen Rückmeldesystemen ist es nicht möglich, mit den Akteuren wirksam zu kommunizieren. Schulentwicklung, viertens, findet heute verstärkt in regionalen Netzen statt, ist also nicht mehr allein auf die einzelne Schule konzentriert. Zudem bedarf sie innovativer Formate der Fortbildung, die sich auf den Unterricht als Kernaufgabe richten. Auf diesen Punkt sollte auch die Weiterbildung der Lehrkräfte konzentriert werden. Wer sie für die Arbeit mit Bildungsstandards gewinnen will, muss sie so ausbilden, dass sie vom Nutzen neuer Instrumente und Verfahren überzeugt sind und diese im Unterricht auch einsetzen können.

Dafür gibt es inzwischen gute Beispiele, die zeigen, dass Entwicklung vor Ort möglich ist. Die Unterscheidung von Ebenen ist ein fünftes Anliegen der Expertise. Die Implementation von Bildungsstandards setzt komplexe Prozesse auf verschiedenen Ebenen voraus, die einer einfachen Top-Down-Sicht von „Educational Governance“ widersprechen. Die Akteure reagieren auf Probleme ihres Umfelds, nicht auf Erlasse, während andererseits die höheren Ebenen der Aufsicht und Verwaltung auf genau diese Probleme eingestellt sein müssen. Das setzt, sechstens, neue Formen der Evaluation, einschließlich Selbstevaluation voraus, worunter es ebenfalls gute Ansätze gibt, die sich in erweiterter Form umsetzen lassen.

Insgesamt sehen wir einen langwierigen und aufwändigen Prozess, der bildungspolitische Kontinuität verlangt. Das ist unsere siebte und vielleicht wichtigste Botschaft. Die Implementation von Bildungsstandards wird Jahre dauern und soll die Schule nachhaltig verändern. Das ist nur möglich, wenn die Richtung nicht ständig durch Wechsel in der Politik verändert wird. Ein zentrales Resultat unseres internationalen Vergleichs untermauert diese Botschaft: Wo kein politischer Konsens über die grundsätzliche Richtung herrscht, ist die Entwicklung mühsam, weil immer wieder neu angesetzt wird. Wenn der Wandel stetig sein soll, ist Rückhalt über die Parteigrenzen hinweg notwendig. Dafür sprechen Beispiele aus England und Schweden.

Das Verständnis von Prozessen der Implementation von Innovationen in einem komplexen System, wie es der Bildungsbereich darstellt, erfordert, das macht der Ansatz unserer Expertise deutlich, Forschungszugänge und -ergebnisse aus den unterschiedlichsten erziehungswissenschaftlichen Bereichen, aber auch aus daran anschließenden Gebieten von der Policy-Forschung bis zur Psychologie. Entsprechend waren wir bei unserem Unternehmen an verschiedenen Stellen auf das Spezialwissen von Kolleginnen und Kollegen angewiesen. Unser Dank gebührt besonders PD Dr. Urs Moser und Dr. Dominik Petko, deren Beiträge zu Tests und Bildungsmonitoring bzw. ICT-Werkzeugen direkt in die entsprechenden Kapitel eingeflossen sind. In verschiedenen Phasen der Kon-

zeption und Ausarbeitung konnten wir auf Anregungen und Rückmeldungen des „Konsortiums“ in der Funktion eines wissenschaftlichen Beirates zurückgreifen. Wir danken sämtlichen Mitgliedern: Prof. Dr. Herbert Altrichter, Dr. Brigitte Bents-Rippel, Prof. Dr. Lucien Criblez, Prof. Dr. Kurt Czerwenka, Prof. Dr. h.c. Rolf Dubs, Prof. Dr. Ferdinand Eder, Prof. Erich Ettlin, Prof. Dr. Eckhard Klieme, Prof. Dr. Rudolf Messner, Prof. Dr. Manfred Prenzel, Dr. Ute Pres, Dr. Erich Ramseier, Prof. Dr. Hermann Saterdag; besonders wertvoll waren Ergänzungen, die uns Herr Czerwenka, Herr Eder, Herr Ramseier und Herr Saterdag auch außerhalb der regelmäßigen Zusammenkünfte beim Verfassen einzelner Kapitel haben zukommen lassen. Zu nennen sind schließlich verschiedene Experten, die sich uns an ganz unterschiedlichen Punkten der Arbeit als Gesprächs- und Interviewpartner zur Verfügung gestellt und mit wichtigen Einschätzungen und Informationen unterstützt haben, nämlich Prof. Dr. Mats Ekholm, Prof. Dr. Lauren Resnick, Prof. Dr. Jaap Scheerens, Prof. Dr. Richard Shavelson, Prof. Dr. Cees Glas, Prof. Dr. Jos Letschert, Prof. Dr. Olaf Köller, Prof. Dr. Werner Blum, Dr. Götz Bieber, Mag. Josef Lucyshyn, Dr. Werner Specht, Dr. Helmut Heugl, Dr. Anton Strittmatter, Ingo Fischer, Dr. Thomas Riecke-Baulecke, Prof. Dr. Fritz C. Staub, Claudia Lena Schnetzler und Prof. Dr. Jürgen Rost.

Die vorliegende Expertise zur Frage der Implementation von Bildungsstandards ist am Pädagogischen Institut der Universität Zürich unter der Leitung von Prof. Dr. Jürgen Oelkers und Prof. Dr. Kurt Reusser und unter weiterer Mitarbeit von Esther Berner (Mitautorin bzw. Autorin der Kapitel 3 und 6), Ueli Halbheer (Mitautor der Kapitel 4 und 5) und Stefanie Stolz (Mitautorin des Kapitels 3) entstanden. Im Bereich des Sekretariats und Lektorats leisteten uns Frau Christina Hartmann, Frau Sonja Geiser und Frau Jonna Truniger unentbehrliche Dienste. Wir alle danken den Vertreterinnen und Vertretern der Auftragsländer, Dr. Werner Specht (BMBWK), Olivier Maradan (EDK), Michel Lanners (MENFP) und ganz besonders Dr. Dorothee Buchhaas-Birkholz (BMBF), für ihr Engagement, erwiesenes Vertrauen und ihre Geduld bis zur definitiven Fertigstellung des Textes.

Zürich, den 22. Januar 2008

Jürgen Oelkers/Kurt Reusser

1 Auftrag, Vorgehen und Ziele der Expertise

Im Auftrag¹ der Expertise *Qualität entwickeln – Standards sichern – mit Differenz umgehen* sind zwei allgemeine Zielsetzungen formuliert:

- (1) Auswertung internationaler Forschungsliteratur und Erfahrungen zu Prozessen der Qualitätsentwicklung und -sicherung unter besonderer Berücksichtigung der Steuerung durch Standards;
- (2) Herausarbeitung von Modellen zur Steuerung von Prozessen der Qualitätsentwicklung und -sicherung unter Berücksichtigung der Steuerung durch Standards und unter Bedingungen kultureller Differenz.

Zur Bearbeitung dieser Ziele befasst sich die Expertise im Kern mit der Implementation von Bildungsstandards im Rahmen eines umfassenden Konzepts der Qualitätsentwicklung von Schule und Unterricht. Standards müssen im Praxisfeld verankert sein und Nachhaltigkeit zeigen. Die Expertise geht der Frage nach, unter welchen Bedingungen und mit welchen Maßnahmen dies auf den verschiedenen Ebenen des Bildungssystems erreicht werden kann.

Unter „Implementation“ wird kein singulärer und kein allein technischer Vorgang verstanden. Vielmehr beschreibt der Begriff ein Bündel von Maßnahmen, die für die Bildungssysteme im deutschsprachigen Raum zum Teil neue Aufgaben darstellen, die dauerhaft realisiert werden müssen. Im Kern geht es um innovative Formen der Qualitätssicherung und Qualitätsentwicklung, die von einem systemweiten Bildungsmonitoring durch Testverfahren bis zur lokalen, kooperativen Schul- und Unterrichtsentwicklung reichen. Die Etablierung von Testkulturen ist lediglich *ein* Mittel aus einer breit gefächerten Palette von Werkzeugen im Dienste einer wirkungsvolleren Steuerung und Qualitätsentwicklung des Bildungswesens. Grundlegend lässt sich festhalten: Die Entwicklung von *Qualität* im Bildungswesen hat primär mit der *Sicherung von Standards* zu tun, die erreicht werden müssen und nicht unterschritten werden dürfen.

Der Fokus der Expertise bezieht sich auf verschiedene Problemfelder, die im Laufe der Untersuchung immer wieder angeschnitten und aus unterschiedlichen Perspektiven behandelt werden:

- Politische Erwartungen an eine Qualitätsentwicklung mit Bildungsstandards,
- Zielsetzungen und Erfahrungen im Umgang mit Bildungsstandards,
- Evaluation durch Tests und andere Verfahren inklusive Rückmeldung von Ergebnissen an Schulen und Lehrkräfte,
- Zeiträume der Entwicklung und Formen der Überprüfung des Erreichten,

¹ Auftraggeber der Expertise sind: das deutsche Bundesministerium für Bildung und Forschung (BMBF), das österreichische Bundesministerium für Bildung, Wissenschaft und Kultur (BMBWK) bzw. das Bundesministerium für Unterricht, Kunst und Kultur (BMUKK), die Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) und das luxemburgische Ministère de l'Éducation nationale et de la Formation professionnelle (MENFP).

- Akzeptanz bei den Akteuren,
- Instrumente zur Nutzung von Bildungsstandards,
- Kommunikationsprofile des Projekts „Implementation von Bildungsstandards“ und Unterstützungssysteme.

Im Einzelnen geht die Expertise der Frage nach, wie die zentralen Konzepte des Projekts „Implementation von Bildungsstandards“ begrifflich gefasst werden können, was ihre Entwicklung ausmacht und vor allem, wie sie als Steuerungsmittel im System so verankert werden können, dass sie die Ebenen des Unterrichts und der Akteure erreichen. Hier liegt die zentrale Botschaft der Expertise. An ihrem Ende stehen Schlussfolgerungen, die sich an die Entscheidungsträger in den vier Auftragsländern richten.

Das Problem der *Differenz* wird an verschiedenen Stellen der Expertise angesprochen, ohne einen eigenen Gliederungsabschnitt zu erhalten. Der Ausgangspunkt für die gesamte Standard-Diskussion war die Frage, wie viel Differenz in staatlichen Bildungssystemen zulässig ist und welche Formen von Einheitlichkeit es geben soll. Ein hoher Grad von Individualität des Unterrichts ist oft gebunden an eine weitgehende Unvergleichbarkeit der Resultate, weil Lehrplanziele selten als verbindlich verstanden werden, die Methoden- und Mittelwahl freigestellt ist und geeignete Formen der Überprüfung fehlen. Auf der anderen Seite werden Standards von den Lehrkräften oft als „Uniformierung“ und Beschneidung ihrer Freiheiten wahrgenommen. Die Ausübung des Lehrberufes wird individuell verstanden, aber das widerspricht in gewisser Hinsicht dem Auftrag öffentlicher Bildung, der egalitär ausgerichtet ist. In diesem Spannungsfeld bewegt sich auch die Diskussion um eine Einführung von Bildungsstandards.

Das Thema Differenz hat noch einen anderen Aspekt, nämlich den der Chancengleichheit. Formen und Grade individueller, sozialer, aber auch kultureller Differenz gehören zu den Eingangsbedingungen jeder heutigen Schulerfahrung. Schulen müssen sich ständig auf ungleiche Voraussetzungen einstellen und angemessen darauf reagieren. Standards sind notwendig, nicht nur um eine angenäherte Gleichheit des curricularen Angebots zu sichern, sondern auch um den Prozess, der in negativer *und* in positiver Weise den Umgang mit Heterogenität und damit Differenz einschließt, kontrollieren zu können. Schließlich verlangt, unter der Voraussetzung einer zunehmend heterogenen Schülerschaft, gezielte Förderung adaptive Maßstäbe, die nicht zufällig erscheinen dürfen. In diesem Sinne sind Standards eine Antwort auf die Frage, welche Chancen Schülerinnen und Schüler mit unterschiedlichen Voraussetzungsprofilen haben, die Ziele schulischer Lehrgänge auch wirklich erreichen zu können.

Schließlich gibt es einen dritten Bedeutungsaspekt von Differenz, der in der vorliegenden Expertise ebenfalls an verschiedenen Stellen thematisiert wird. Gemeint ist, dass nationale Bildungssysteme politisch und historisch gewordene soziokulturelle Konfigurationen darstellen, in denen sich die auf Schule und Lernen bezogenen Werte, Deutungsmuster und Praktiken einer Kultur widerspiegeln. Auch wenn Sprache und Rhetorik der gegenwärtigen Bildungsreform in verschiedenen nationalen Systemen angenähert gleich sind, ist und bleibt die Praxis doch in vielerlei Hinsichten divergent. Schulen in Deutschland, der Schweiz, in Österreich und Luxemburg werden notwendigerweise einen guten Teil ihrer historischen Besonderheiten wahren. Die Entwicklung und Implementation von Bildungsstandards muss darauf auch in einem sich *globalisierenden* Bildungsmarkt in gebührendem Maße Rücksicht nehmen.

In den deutschsprachigen Ländern ist eine Kultur der Sicherung von Qualitätsansprüchen mit Hilfe von Tests, Evaluationen und Rückmeldungen bisher wenig entwickelt. Das erklärt bestimmte Reaktionen der Abwehr. Auf der anderen Seite geht es um den Auftrag, die Legitimität und die Finanzierung öffentlicher Bildung. Die Ziele, Inhalte und Aufgaben der Bildung sind so formuliert, dass sie – unterschieden nach Stufen und Lehrgängen – für *alle* Lernenden gelten sollen. Ob und wie dieser Anspruch erfüllt wird, kann nicht länger offen bleiben. Die grundsätzliche Frage, warum es überhaupt Bildungsstandards geben soll und was sie relevant macht, beantwortet sich hier. Öffentliche Bildung kann und muss *rechenschaftspflichtig* verstanden werden. Damit ist das entscheidende Relevanzkriterium formuliert. Dieses Kriterium verlangt, dass transparent gemacht werden muss, ob und wie die elementaren Ziele der Bildung, über die sich die Gesellschaft verständigt hat, erreicht oder verfehlt wurden. Zudem muss dafür Verantwortung übernommen werden. Die Standarddiskussion hat hier ihren rationalen Kern.

Dies heißt aber nicht, dass es *ein* „bestes“ Standardkonzept gibt, das sich überall gleich anwenden oder einfach von einem Land auf das andere übertragen ließe. Nach Lage der Dinge kann man nicht einfach fremde Systeme kopieren, und seien sie noch so erfolgreich. Auch das eigene System lässt sich nicht ein zweites Mal komplett neu erfinden, wie manche Kommentare zu unterstellen scheinen, sondern nur – analog einem Schiff auf hoher See – langsam umbauen und weiterentwickeln. Gegen radikale Phantasien eines Neubaus des Bildungssystems steht die historische Macht des nach dem Gesundheitswesen zweitgrößten gesellschaftlichen Systems, das sich aufgrund seiner historischen Prämissen weiterentwickeln wird. Auch die Systemsteuerung kann nicht einfach neu erfunden, sondern nur in mehr oder weniger großen Schritten entwickelt werden.

Das zeigt sich nicht zuletzt beim Thema dieser Expertise. Bildungsstandards können im internationalen Vergleich „auf völlig unterschiedlichen theoretischen Annahmen, Wertvorstellungen und gesellschaftlichen Zielen“ basieren, während die Akzeptanz dieser Standards von genau diesen Annahmen, Werten und Zielen bestimmt wird (Klieme/Avenarius/Blum et al. 2003, S. 35). Sie lassen sich nicht theoretisch aus der Welt schaffen, und sie sind je nach Land unterschiedlich. Ein solcher Befund gilt auch für die vier Auftragsländer dieser Studie. Trotz ähnlicher Reformsemantik, Ausgangslagen und Referenzen sind die tatsächlichen Strategien der Systementwicklung durchaus verschieden.

Das bedeutet aber nicht, Unvergleichbarkeit oder die Einzigartigkeit aller Verhältnisse annehmen zu müssen. Im Blick auf Lösungen für spezifische Problemstellungen lässt sich aus den verschiedenen Erfahrungen der nationalen Systeme lernen. Auch lassen sich Schwierigkeiten vergleichen und Verschiebungen in den Problemen nachvollziehen. Die Effekte bestimmter Strategien können nützliche Informationen darstellen, mit denen sich abschätzen lässt, was vermieden oder was angestrebt werden sollte. Schließlich können auch Modelle, Verfahren und soziale Werkzeuge, die sich anderswo bewährt haben, übernommen und angepasst werden. Die nachstehende Expertise verfolgt diesen Weg des punktuellen Vergleichs, der die Randbedingungen des eigenen Systems beachtet.

Für die Analyse und ihre Schlussfolgerungen werden Daten und Befunde aus der internationalen Forschungsliteratur herangezogen sowie Aussagen aus Interviews mit Expertinnen und Experten aus mehreren Ländern. Hinzu kommen Dokumente aus der laufenden Entwicklungsarbeit in den einzelnen Ländern und weitere Quellen, die Aufschluss geben über erfolgversprechende, aber auch anforderungsreiche und riskante Strategien der Implementation. Die damit verbundenen Erfahrungen lassen sich im Blick auf bestimmte Probleme generalisieren, sofern die nationalen, regionalen und lokalen Gegebenheiten gebührend Beachtung finden. Adressat sind die vier Auftragsländer, also die Bundesrepublik Deutschland, die Republik Österreich, die Schweizerische Konföderation sowie das Großherzogtum Luxemburg. Der Aufbau der Expertise ist wie folgt:

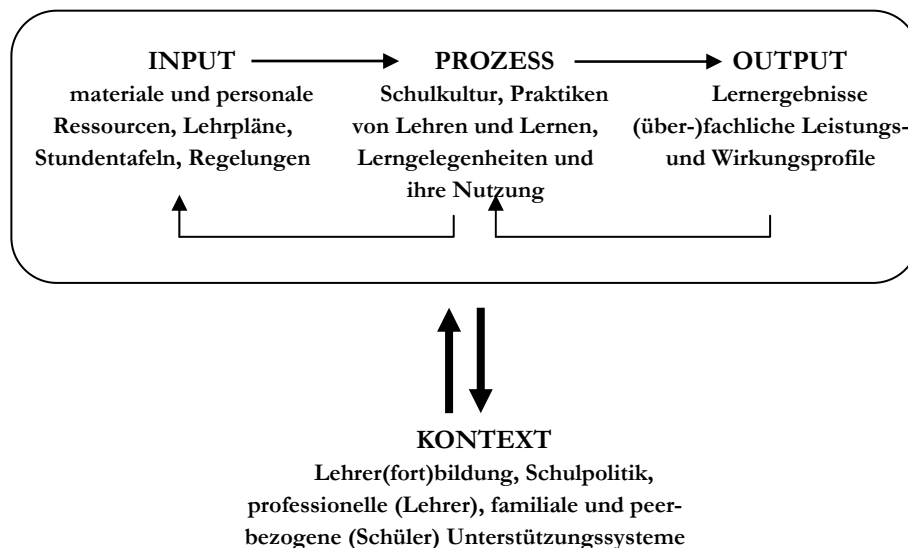
- In einem ersten Teil wird die Entwicklung und Diskussion der zentralen Konzepte dargelegt. Begriffe wie „Kompetenz“, „Standards“ und „Implementation“ sind nicht einfach politisch gesetzt worden, sondern haben eine bestimmte theoretische und empirische Klärung hinter sich (Kapitel 2).

- In einem zweiten Schritt werden einschlägige Ergebnisse der Forschung und Entwicklung von Bildungsstandards aus vier ausgewählten Ländern dargelegt (Kapitel 3).
- Darauf aufbauend wird ein Modell der Implementation entwickelt, in dem Leitideen, Strategien, Ebenen, Akteure, Verfahren und Kontexte unterschieden werden (Kapitel 4).
- Viertens werden konkrete Verfahren und Instrumente beschrieben, die auf verschiedenen Ebenen Anwendung finden können (Kapitel 5).
- Daran anschließend wird auf den Stand der Qualitätsentwicklung und speziell der Entwicklung von Bildungsstandards in den vier Auftragsländern dieser Studie eingegangen (Kapitel 6).
- Am Ende stehen Schlussfolgerungen und Empfehlungen (Kapitel 7).

2 Entwicklung und Diskussion der zentralen Konzepte

Der heutige Diskurs über Bildungsstandards geht zurück auf einen inzwischen international weitgehend akzeptierten Wechsel der Perspektive in der Bildungspolitik. Genauer müsste von einer *Erweiterung* der Perspektive die Rede sein. Die Erweiterung betrifft grundlegend die politische Aufmerksamkeit, die sich bislang auf die *Systemeingänge* bzw. die Input-Faktoren eines *Input-Prozess-Output-Modells* der Funktionsweise von Bildungssystemen (Abb. 1) konzentrierte und weder die Qualität der *Prozesse* noch der *Kontexte* und der *Resultate* sonderlich beachtete. Die Verschiebung des Augenmerks *von der Strukturqualität und der Prozessqualität* auf die Output-Seite des Bildungswesens und damit das Bestreben, dessen *Ergebnisqualität fassbar* zu machen, ist eine neue Politikstrategie. Sie hat zu tun mit der Frage der Effektivität, also des Einsatzes der Ressourcen, die nur mit Ergebnissen beantwortet werden kann. Neue Strategien der Systemsteuerung und der Systementwicklung im Bildungswesen setzen vor allem hier an.

Abbildung 1: Einfaches Modell der Funktionsweise von Bildungssystemen



Im Rahmen dieser Perspektivenerweiterung haben Bildungsstandards eine internationale Karriere erlebt, in deren Verlauf sich auch im deutschen Sprachraum bildungspolitische Überzeugungen verändert haben. Dabei ist nicht einfach nur ein Wandel der Expertensprache zu beobachten, sondern auch eine Veränderung der Politik. Die Aufmerksamkeit richtet sich nunmehr auch auf die Resultate des Unterrichts und damit zusammenhängend auf die Leistungen des Schulsystems insgesamt. Der Wandel der Politik, der mittlerweile in den meisten OECD-Ländern beobachtet werden kann, steht in einem engen Zusammenhang mit den Ergebnissen der TIMS-Studie und der beiden PISA-Stu-

dien, die einen internationalen Wettbewerb evoziert haben. Der Wandel betrifft mehr als nur die Leistungen der Schülerinnen und Schüler, die ja auch bisher schon beurteilt und bewertet wurden. Die Diskussion der Bildungsstandards steht im Zusammenhang mit neuen Konzepten der *Qualitätssicherung*, die über das hinausgehen, was in den deutschsprachigen Ländern von ihrer Tradition her üblich war und ist. Damit ist aber nicht das gesamte Thema der Qualitätssicherung neu, wie in einem ersten Schritt gezeigt werden soll.

2.1 Historische Erfahrungen mit Qualitätssicherung in Schulen

Die Geschichte der staatlichen Schule kann seit ihren Anfängen als Entwicklung hin auf eine zunehmende Standardisierung im allgemeinen Wortsinne beschrieben werden. Versteht man Standards als Festlegungen von Inhalten, Normen, Formaten, Prozeduren und Zielen, die im gesamten Geltungsbereich von Bildungsinstitutionen oder in Teilbereichen Anwendung finden sollen und die Erwartungen darstellen, deren Erfüllung mehr oder weniger verbindlich festgelegt ist, so liegt darin noch nichts grundsätzlich Neues. In dieser Hinsicht bestehen Schulen geradezu aus Standards, die mehr sind als lediglich staatliche Vorschriften in Papierform.

Offizielle Regelungen, formelle wie informelle Normen oder Maßstäbe hinsichtlich einer sich allmählich herausbildenden *grammar of schooling* (Tyack/Tobin 1994) finden sich bereits im 19. Jahrhundert bezogen auf Schulzeit und Schulbesuch, Durchführung des Unterrichts, Aufteilung der Schülerinnen und Schüler (z.B. in Jahrgangsklassen), des Lernraumes und der Lernzeit, Studententafeln, Lehrpläne und Lehrmittel, Notengebung, Zeugnisse und Berechtigungen, Lehrerbildung und Lehrberechtigung, Aufgaben und Verfahren der Schulaufsicht etc. Damit sind Felder genannt, die bezüglich Inhalt, Leistung und Ressourcen traditionell in irgendeiner Weise zunehmend standardisiert worden sind. Dazu kommen viele meist informelle methodische Regeln in der Gestaltung des Unterrichts, die noch erheblich älter sind.

Mit der Verstaatlichung des Bildungswesens im 19. Jahrhundert ist zudem noch eine weitergehende Standardisierung festzustellen, die die lokale und regionale Organisation von Schule und Unterricht erweiterte und allmählich zum Aufbau nationaler Systeme führte. Anlass dafür waren meist diagnostizierte Mängel und Defizite des vorhandenen Systems, das heißt, Standards waren stets zur Qualitätsverbesserung gedacht. Dafür bildete sich eine bestimmte Form der politischen Argumentation in Bezug auf den überlokalen und internationalen Vergleich heraus, wie sie auch heute noch üblich ist (Gonon 1998).

Der Vergleich zielte auf die Benennung und die Behebung der eigenen Schwächen, was voraussetzte, die Entwicklungen außerhalb des lokalen Raumes, insbesondere im Ausland, stark zu machen und als Vorbild hinzustellen. Es war dann immer schwierig, Ideal und Wirklichkeit der anderen Systeme klar zu unterscheiden, weil nur deren Idealisierung das Argument der eigenen Rückständigkeit nach innen hin stark machte.

Diese Probleme der vergleichenden Systembeobachtung finden sich bis heute. Unabhängig von der Neigung der Idealisierung des Fremden entwickelten sich Standards vor allem seit der Mitte des 19. Jahrhunderts in allen Bildungssystemen. Für bestimmte Probleme wurden vergleichbare Lösungen gefunden, die oft auf punktuellen Übertragungen aus anderen Systemen beruhten. Historische Beispiele auf der Stufe des Unterrichts waren etwa neue Lehrmittel wie große Landkarten für den Anschauungsunterricht, Lernhilfen wie Lesekarten für die Selbstinstruktion, Lernmedien wie Rechenschieber für den Fachunterricht, Formen von Klassenarbeiten und Rückmeldepraktiken der Lehrkräfte nach vollzogenen Leistungen oder neue Formen der Gruppierung und Anforderungsdifferenzierung unter den Schülerinnen und Schülern im Lernraum.

Auf dieser Ebene von didaktischen Unterrichtsmitteln fanden und finden Innovationen schnell Verbreitung, ohne Landesgrenzen zu wahren. Nicht zufällig ist Brauchbarkeit im Unterricht eine zentrale Voraussetzung für wirksame Implementation. Werkzeuge, Medien und Technologien, in diesem Sinne *Know-How*, das sich in „Tools“ kristallisiert, können viel leichter übertragen werden als relativ abstrakte, sprach- und kulturgebundene Theorien, die komplexe Deutungsarbeit verlangen.

Aus diesen historischen Erfahrungen lassen sich allgemeine Schlussfolgerungen ziehen: Standards legen für einen bestimmten Geltungsbereich Maßstäbe oder Verfahren fest. Sie bestimmen so den Erwartungshorizont und reglementieren die Praxis. Akzeptanz bei den Akteuren finden sie vor allem aufgrund ihrer Brauchbarkeit. Es ist falsch, davon auszugehen, dass Standards allein staatliche Vorschriften darstellen. Auch die sich entwickelnde Profession selbst hat Standards entwickelt. Historisch wirksam gewesen sind Standards immer nur als effektive Problemlösungen und Werkzeuge, die sich in irgendeiner Form als nützlich erweisen mussten. Demgegenüber müssen sowohl Theorien als auch Reglemente angepasst und in die Lehrerbildung sowie die Schulorganisation übersetzt werden, was nie vollumfänglich möglich ist.

Vorschriften ohne Platz im System werden zudem leicht abgewiesen oder nur widerwillig implementiert, was ihren Nutzen beeinträchtigt. (Bildungs-) Standards müssen deshalb *im* System, gemeinsam mit dessen Akteuren entwi-

ckelt werden. Dabei müssen Ebenen und Anschlüsse beachtet werden. Lehrplanrevisionen wirken sich nur begrenzt auf Lehrmittel aus, und Innovationen des Primar- oder Grundschulunterrichts beeinflussen anschließende Stufen wiederum nur sehr bedingt. Auch die verschiedenen Kulturen *im* System lassen sich kaum auflösen oder angleichen. Hinzu kommt, dass Standards historisch nie als Systemplanung und Versuch der Gesamtsteuerung eingeführt oder realisiert worden sind. Die Entwicklung erfolgte immer punktuell und unter Wahrung der intern gegebenen Differenzierungen. Daher ist in der Vergangenheit auch kein Masterplan implementiert, sondern versucht worden, gegebene Problemlösungen zu verbessern.

Historisch gewachsene Standards lassen sich nicht nur für die Praxis des Unterrichtens nachweisen. Standards sind auf ganz unterschiedlichen Stufen und in allen Bereichen des Bildungssystems entwickelt worden, denkt man etwa an die Benutzung gleicher oder ähnlicher Lehrbücher auf gleichen Stufen, die Versetzung oder Promotion der Schülerinnen und Schüler von niederen in höhere Stufen, die Verteilung der Lernzeit nach Unterrichtsfächern, die damit verbundenen curricularen Vorschriften oder die Zulassungsbedingungen zu einzelnen Schulzweigen. Dazu gehören auch die Grundnormen der Raumausstattung, der Zeittakt des Unterrichts und die Einstellung des Lernens auf die Raum-Zeit-Bedingungen. Das Verhältnis von Aufgaben und Leistungen ist maßgebend für jede Form von Unterricht. In jedem System gibt es Bewertungssysteme, die sich weniger unterscheiden, als es den Anschein hat. Psychometrische Tests haben seit Beginn des 20. Jahrhunderts eine eigene Geschichte. Und schließlich sind in allen Bildungssystemen Standards für die Professionalisierung des Lehrpersonals entwickelt worden (Ravitch 1995, S. 34f.; siehe auch Resnick 1982; Resnick/Resnick 1985; Aldrich 2000; Ballantyne 2002; Oelkers 2004).

Neu ist also weder das grundsätzliche Bestreben, in einem stark individualisierten Berufsfeld Standards als eine wichtige Voraussetzung für die Gleichheit der Bildungschancen zu setzen, noch ist es der Tatbestand ihres realen Vorkommens im Schulgefüge, wie in der heutigen Diskussion gelegentlich unterstellt wird. Die staatliche Schule war nie einfach zentriert auf individuelle Anliegen oder Bedürfnisse, vielmehr musste immer ein gleicher Auftrag an alle Lernenden einer bestimmten Kategorie erfüllt werden, wobei schon früh nach Typen oder Stufen unterschieden wurde. Die Standardisierung war somit auch nie gleich, sondern reagierte auf verschiedene Bedingungen, zu denen das Alter der Lernenden ebenso wie die Unterschiede in der Leistungsfähigkeit zählten. Auch Standards für Fördermaßnahmen wurden entwickelt, ebenso solche für die

Kommunikation der Schule mit dem Umfeld oder die Anforderungen an die Kompetenz der Lehrkräfte.

Und seit jeher besteht zwischen Standards – geteilten Normen und Erwartungen – auf der einen Seite und der tatsächlichen Gestaltung des Unterrichts auf der anderen immer auch eine charakteristische Spannung, die sich in der seit dem 19. Jahrhundert entwickelten staatlichen Schule verschärft hat. Vom „Kult der Effizienz“ (Callahan 1962) war längst vor der heutigen Diskussion die Rede, und der Vorwurf einer „bürokratischen“ Systemsteuerung findet sich bereits in der deutschen Volksschulliteratur des 19. Jahrhunderts an vielen Stellen. Die Kritik, obwohl anhaltend und stereotyp, hat die tatsächliche Systementwicklung kaum beeinflusst. Ausschlaggebend für die Entwicklung des Systems waren und sind bis heute neben der Qualifikation der Lehrkräfte die Standards des Curriculums, der Schulorganisation und der Input-Seite des Bildungssystems insgesamt. Sie sollen Schulen vergleichbar halten und sind an sich nichts Neues oder gar Ungewöhnliches.

Auch die Ausbildung von Lehrerinnen und Lehrern ist nicht beliebig, sondern setzt ebenfalls bestimmte Standards von Zeit, Organisation und Ergebniskontrolle voraus. Jede Schulstunde stellt eine Standardisierung dar, da sie nämlich die Zeit mit einer bestimmten Maßeinheit normiert, die in allen Bereichen der Organisation gilt. Diese Normierung ist das Ergebnis eines langen historischen Prozesses, der bei allen Anstrengungen, die Qualität von Schule und Unterricht zu verbessern, vorausgesetzt werden muss. Zudem muss das Gewicht der bestehenden Organisation beachtet werden. Es sind große und komplexe Systeme, die mit den gegebenen Standards ein hohes Maß an Sicherheit in der Gestaltung des Arbeitsalltags verbinden.

Was ist dann aber *neu* in der heutigen Diskussion um die Einführung von *Bildungsstandards*? Grundsätzlich neu sind die Funktion und Ausgestaltung von Standards sowie deren Platzierung im gesamten Bildungssystem. Auf der Systemebene betrafen „Standards“ bislang die Steuerung von Schule und Unterricht vor allem durch Struktur- und Programmvorgaben (Schulorganisation, Stundentafeln, Lehrpläne, Mittelallokation), das heißt durch den Bereich des „*Inputs*“ von Schule und nur sehr schwach durch die Überprüfung von Ergebnissen (zum Beispiel via Schulaufsicht, Noten, Übertrittsquoten, Abschlussprüfungen). Ausgegangen wurde von der selbstverständlichen (jedoch kaum überprüften) Übereinstimmung von offiziellen Vorgaben, verinnerlichten Absichten und realisierten Ergebnissen. Neu besteht die politische Absicht darin, sich bei der Steuerung von Bildungssystemen vermehrt auch am „*Output*“, das heißt an objektivierbaren, regelmäßig überprüften Leistungen einer Schule

bzw. den empirisch gemessenen Lernergebnissen der Schülerinnen und Schüler zu orientieren.

Auf der Ebene des Lehrerhandelns wurden die Ergebnisse des Unterrichts zwar schon bisher numerisch, mit Hilfe des Notenschemas beschrieben, jedoch gruppennormorientiert und ohne Bezug auf ein fachinhaltlich-qualitatives, den Referenzrahmen eines zufällig zusammengesetzten Klassenzimmers übergreifendes Lernzielkriterium. Konkret wurden und werden Schulnoten aus der Sicht der Lehrkraft erteilt, sie geben einen Klassendurchschnitt wieder, erfassen aber nicht oder nur sehr oberflächlich eine tatsächlich erreichte, durch materielle Fachstandards umschriebene Kompetenz, weil qualitativ-inhaltliche Bezugsnormen fehlen. Die heutige Notenorientierung der Schule ist daher nur sehr bedingt auch eine Ergebnisorientierung im Sinne einer Orientierung an vergleichbaren inhaltlichen Bildungsstandards oder Zielvorgaben.

Lehrpläne, die das Angebot bestimmen sollen, sind überfrachtet und lassen zugleich zu viel offen. Die Leistungen der Schülerinnen und Schüler beziehen sich auf Aufgaben, die in aller Regel den Lehrmitteln entnommen sind. Die Zielorientierung der Lehrmittel ist jedoch häufig eher schwach, gleichzeitig fehlen verbindliche und handhabbare Beschreibungen von kompetenzbezogenen Niveaus der Zielerreichung. Das heißt: Noten beziehen sich auf den Durchschnitt einer Klasse, aber geben keinen Aufschluss über die in einem Fach oder einem Lerngebiet erreichten tatsächlichen Kompetenzen. Was die Schülerinnen und Schüler am Ende eines Prozesses tatsächlich wissen und können, muss anders bestimmt werden.

Die politische Neuorientierung am „Output“ oder den Ergebnissen der Lernerfahrungen in Schule und Unterricht setzt eine Neuordnung oder Neugewichtung des Diskurses voraus und so bestimmte Schlüsselkategorien, die zuvor weniger Gewicht hatten, noch gar nicht profiliert waren oder in anderen theoretischen Kontexten verwendet wurden. Vier allgemeine Konzepte sind dabei grundlegend, nämlich erstens *Kompetenz*, zweitens *Feedback* und *Evaluation*, drittens *(Bildungs-)Standards* und viertens *Implementation* von neuen Formen der Qualitätssicherung. Diese Konzepte haben unterschiedliche Ausgangsbedingungen und sind erst allmählich zu einem politisch-pädagogischen Diskurs verknüpft worden. Die Entwicklung des Diskurses soll im Folgenden kurz skizziert werden; sie ist im Blick auf die Begriffe inzwischen international maßgebend.

2.2 Kompetenz als Schlüsselkategorie

Seit den siebziger Jahren des 20. Jahrhunderts, insbesondere seit der „kognitiven Wende“ und so der Ablösung des Behaviorismus in der Sozialwissenschaft, steht die Frage der Kompetenz im Mittelpunkt fachlicher und auch politischer Auseinandersetzungen (vgl. auch Klieme/Hartig im Druck). Oft wurde diese verzweigte und teilweise recht unübersichtliche Diskussion mit Noam Chomskys Unterscheidung von „Kompetenz“ und „Performanz“ in Verbindung gebracht.² Von „kommunikativer Kompetenz“ war zu Beginn der siebziger Jahre auch in der politischen Philosophie die Rede (Habermas 1971), in der Testpsychologie wurde vorgeschlagen, „competence“ statt „intelligence“ zu messen (McClelland 1973), und in der pädagogischen Psychologie wurde die Suche nach „sozialer Intelligenz“ erneuert, die schon den frühen Pragmatismus zu Beginn des 20. Jahrhunderts bestimmt hatte (Keating 1978). Kompetenz und Intelligenz wurden auch zum Konstrukt „kognitive Kompetenz“ zusammengefügt und waren dann in dieser Fassung forschungsleitend. In manchen Lehrplankonzeptionen der Volksschule und in daran anschließenden Ausbildungsgängen wurde sodann die Trias von Fach- und Methodenkompetenz, Sozialkompetenz und Selbstkompetenz dazu verwendet, die allgemeinen Bildungsziele einer Schule oder Ausbildung leitideenhaft zum Ausdruck zu bringen.

Der Begriff „Kompetenz“ wurde allerdings schnell ziemlich inflationär gebraucht. Fast alle pädagogischen Zielaussagen werden heute als Kompetenzen etikettiert. Zwischen einem empirisch gesicherten und einem frei verwendeten Konzept lässt sich nur noch mit Mühe unterscheiden. Zudem gibt es Evidenz, dass Kompetenz und Performanz nicht streng geschieden werden können, und dass ‚Performanz‘ in verschiedenen Hinsichten der kognitiven Entwicklung sogar grundlegender ist als Kompetenz (Sophian 1997). Das einfache Begriffspaar von Chomsky ist auf jeden Fall fraglich geworden. Eine ähnliche Entwicklung

² Noch wesentlich weiter zurück in die Geschichte der Bildung und der Psychologie reicht der Begriff der „Kompetenz“ und reichen die damit verbundenen Probleme, wenn man die Debatte(n) um die mannigfach angenommenen reziproken, transferbezogenen Wirkungen zwischen formaler und materialer Bildung einbezieht (vgl. Reusser 2001). Die Auseinandersetzung um die Bildbarkeit überfachlicher Kompetenzen (bis zur Neuauflage in der Debatte um die sog. „Schlüsselqualifikationen“ und das „Lernen lernen“), vor allem die kognitionswissenschaftlichen Subdisziplinen der Expertise- und der Transferforschung, haben diese alten Diskussionen weitergeführt. Diese Forschungsstränge müssten in eine problemgeschichtliche Diskussion des Begriffs und der Gründe seines weitläufigen rhetorischen Gebrauchs einbezogen werden.

verzeichnet auch der Begriff „Intelligenz“. Unter dem Einfluss von Howard Gardners *Frames of Mind* (1983) sind Erweiterungen des jahrzehntlang sehr eingeschränkt verwendeten (psychometrischen) Konzepts vorgenommen worden, die außerhalb der Forschung zu einem wiederum inflationären Gebrauch geführt haben. „Intrapersonale“ oder „emotionale“ Intelligenz (so sich die Konzepte überhaupt trennscharf definieren lassen) werden synonym mit Kommunikations-, Selbst- oder emotionaler ‚Kompetenz‘ verwendet, und können heute ohne Mühen in der Ratgeberliteratur nachgewiesen werden.

Am ehesten besteht Einigkeit darüber, dass sich der Kompetenzbegriff als Gegenbegriff zum klassischen Intelligenzbegriff von diesem dadurch unterscheidet, dass sich Kompetenzen nicht auf kontextfreie kognitive Dispositionen, sondern auf wissensbasierte Fähigkeiten in bestimmten kulturellen und lebensweltlichen Domänen beziehen. Kompetenzen lassen sich nicht „netto“ entwickeln, sondern sind in ihrem Erwerb auf konkrete Anforderungen und Situationen angewiesen (Reusser 2001). Das heißt, sie müssen in aktuelle Lern- und Problemlöseleistungen umgesetzt werden, was die Stiftung einer direkten Beziehung zum Handeln in einer Domäne beinhaltet. Der Kompetenzbegriff entspricht somit auch einem stärker pragmatisch-funktionalen oder aktionalen (versus stoff-materialen oder kontemplativen) Wissens- und Bildungsverständnis und er steht, was seine intrapersonale Dynamik anlangt, in Beziehung zum Grundbedürfnis des Menschen nach Kompetenz und Selbstbestimmung (Deci/Ryan 1985) in der Interaktion mit der Umwelt (White 1959). Was im Zusammenhang der Forderung nach Kompetenzorientierung kaum je diskutiert bleibt, sind die bildungstheoretischen Implikationen des zugrundeliegenden Wissensverständnisses, verbunden mit der Frage, warum und innerhalb welcher Grenzen es wichtig ist, vermehrt kompetenzorientiert im Gegensatz zu im herkömmlichen Sinne stofforientiert zu unterrichten.

Von erheblichem Einfluss auf die Entwicklung der Diskussion um den Kompetenzbegriff vor allem im deutschsprachigen Raum war die Expertise für die OECD von Weinert (1999, 2001a). Sie ist zunächst als Versuch zu verstehen, in die Diskussion über Kompetenz, die auf sehr unterschiedlichen Ebenen geführt wurde und geführt wird, Struktur und Klarheit zu bringen. Obgleich Weinert das Verständnis von Kompetenz als einer kognitiven Disposition im Zentrum behält, geht er davon aus, dass Kompetenzen erlernbar sind, und er bezieht darüber hinaus auch weitere, handlungsnähere Facetten ein. Der Literaturbericht diskutiert verschiedene Konzepte, die von

- einem allgemeinen psychologischen Konstrukt „Kompetenz“,
- einer davon unterschiedenen spezifischen Disposition für „Performanz“,

zu Facetten wie

- motivationaler und sozialer Kompetenz,
- Handlungskompetenz,
- Schlüsselkompetenzen
- sowie Metakompetenzen reichen.

In der Diskussion, in welcher diese Konzepte für Bildungsprozesse relevant sind, haben sich vor dem Hintergrund der Expertiseforschung vor allem kognitive Konzepte, die wissensbasiert sind, ergänzt durch motivationale Facetten, als theoretisch und empirisch fassbar erwiesen. Hinzu kommen Metakompetenzen und Handlungskompetenzen. Allgemeine Persönlichkeitsmerkmale, fachübergreifend bzw. wissensunabhängig gedachte intellektuelle Kompetenzen, die auch psychometrischen Messungen nur schwer zugänglich sind, werden vernachlässigt. In den Mittelpunkt des Begriffs rücken bereichsspezifisches Wissen und damit zusammenhängende Fähigkeiten unter Bezugnahme auf komplexes Problemlösen (ebd., S. 27).

Das ist ein weitreichender Schritt, wird damit doch nichts anderes als eine pädagogisch realisierbare Lösung des alten Transferproblems in Aussicht gestellt. Weinert unterscheidet im Kern zwei Theorieansätze, und zwar erstens das dem Intelligenzbegriff nahestehende „psychometric model of intellectual abilities and ability differences, and the theory of knowledge acquisition and its use in subsequent learning, problem solving and automatic behaviour“ (ebd.). Das zweite, wissensbasierte Modell (Connell/Sheridan/Gardner 2003 sprechen von „realized abilities“) liegt den Empfehlungen Weinerts zugrunde und limitiert das Konzept der Kompetenz, indem dieses nicht allgemein, d.h. bezogen auf eine hohe Generalisierbarkeit von Leistungsfacetten, sondern bezogen auf Domänen und damit auf spezifischere Anwendungsbereiche verwendet wird. Kognitive Kompetenz wird begrenzt auf gelerntes Wissen, Fähigkeiten und korrespondierendes Metawissen (ebd.). Losgelöste Schlüsselkompetenzen *ohne* genetischen Bezug zu einem langfristigen Wissensaufbau und zu Lernbereichen, wie sie in der pädagogischen und didaktischen Literatur grassieren, werden von Weinert als „inkorrekt“ zurückgewiesen (ebd.).

Die Begrenzung des Begriffs „kognitive Kompetenz“ auf Wissen und Lernen bzw. auf kontextualisierte Leistungsdispositionen hat vier Vorteile, die so gefasst werden:

- A large proportion of individual differences in intelligence is captured by learning performance, so it is not necessary to address the theoretically and methodologically difficult problem of the role of ability differences in knowledge acquisition.
- It is possible to use hierarchically and heterarchically constructed performance tests as valid indicators of competence hierarchies and competence components.
- In contrast to an ability centred concept of competence, one can directly infer the effectiveness of prior learning opportunities through knowledge-centered (or task demand-centered) competencies. This will make national and international comparison studies possible and meaningful.
- The knowledge-centered concept of competence can be applied to the large numbers of types of performance demands in daily life, school and work activities (ebd.).³

Weinert versteht Kompetenz auf dieser Linie sehr viel angenäherter an Performanz und an bereichsspezifisches Wissen – und damit an *Lernen* und Problemlösen – als Chomsky. Das heißt, es geht beim Kompetenzbegriff nicht um die genetische Basis allgemeinsten kognitiver Grundfähigkeiten, sondern um etwas Erlerntes und *Erlernbares*, mit Aebli (1980/1981): um erworbene Handlungs-, Operations- und Begriffsschemata, die Personen dazu befähigen (sollen), innerhalb bestimmter bereichsspezifischer Rahmen Probleme zu lösen. Kompetenz hat wohl eine materielle Basis in den Genen, aber sie entwickelt sich aufgrund von Lernprozessen, die kumulativ und inhaltspezifisch gestaltet werden müssen, je verschieden. Kompetenzen sind je erreichte inhaltliche Fähigkeiten und Fertigkeiten wie Lesen, Mathematik-Betreiben, naturwissenschaftlich Argumentieren-Können, die durch nachfolgendes Lernen erweitert und verändert werden können, wenngleich immer nur graduell, oft langsam und stets in Verbindung mit Fachinhalten. Das kommt einem pragmatisch-funktionalen und wissensbasiert-konstruktivistischen Konzept des Problemlösens nahe, wie es schon Dewey und in der Didaktik Aebli (1951) im Anschluss an Piaget vertreten haben, in dessen Rahmen die ständige Rekonstruktion der Erfahrung bzw.

³ Ein Nachteil dieses Konzepts gelernter Kompetenz besteht darin, dass die gegebenen allgemeinen intellektuellen Fähigkeiten bestimmen, wie gut und wie viel neues Wissen gelernt wird. Aber die Begabungsforschung zeige, so Weinert (1999, S. 28), dass verschiedene Ebenen der Intelligenz nur einen sehr generellen Maßstab darstellen, „for predicting knowledge acquisition and the level of cognitive achievement“.

der Aufbau operativer (flexibler, intelligenter) Wissens- und Denkstrukturen als Kerntätigkeit des Unterrichts und des Lernens betont wird. Fähigkeiten sind danach habituell gewordene Denkstrukturen (flexible habits), die auf intelligente Formen der Anpassung zurückzuführen sind. In der neueren Literatur ist auch von „adaptive thinking“ die Rede, also die Anpassung der Problemlösung an die sozialen Gegebenheiten (Gigerenzer 2000, S. 201ff.).

Weinert hat das Konzept „Kompetenz“ auch explizit mit der Theorie der Problemlösung zusammengebracht. Kompetenzen sind „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 2001b, S. 27f.). Wichtig ist, dass Problemlösungen zu Kompetenzen aufgebaut, die Kompetenzen aber mit neuen, nachfolgenden Lösungen auch verändert werden können. Kurz gefasst könnte man sagen: Kompetenz ist die persönlich erreichte und automatisierte Fähigkeit, in bestimmten Wissensdomänen und nach Abschluss vieler verschiedener Lernsequenzen in begrenzter Generalisierung auf neue Anforderungen hin Probleme lösen zu können.

Die vier von Weinert beschriebenen Vorteile einer wissensbasierten Konzeption kognitiver Kompetenz sind in der Entwicklung der Bildungsstandards im deutschen Sprachraum theoretisch maßgeblich geworden. Das Konzept erlaubt die Konstatierung von individuellen Unterschieden, die nicht einfach biologisch verstanden werden, sondern auf schulisch mitverantwortetes Lernen zurückführbar sind. Die jeweilige Anwendung der Fähigkeiten und Fertigkeiten in bestimmten Bereichen – das erreichte Können – lässt sich damit ebenfalls testen. Damit kann auch die Effektivität früherer Lernmöglichkeiten erfasst werden, wobei aufgrund der Annahme individueller Eingangsniveaus gestufte Ergebnisse angenommen werden können. Das Modell erlaubt schließlich auch einen Gebrauch über die Schule hinaus, weil kognitive Kompetenz als produktive Verbindung von Wissen und Können in verschiedenen Situationen nachgefragt wird und entwickelt werden muss. Die Situationen der Verwendung sind somit variabel, entsprechend sind die Kompetenzen keine festen, situations- oder kontextfreien Größen, die einfach akkumuliert werden könnten.

Das Modell der kognitiven Kompetenzen wurde in der Diskussion der Bildungsstandards auf bestimmte Lerndomänen oder schulische Fachzusammenhänge bezogen. Im Anschluss an Weinert und die Kompetenzmodelle der PISA-Studien sind Erweiterungen des Kompetenzmodells diskutiert worden, die über schulisch-fachliche Lerndomänen hinausgehen und neben kognitiven

auch allgemeine „key-competencies“ thematisieren (Rychen/Salganik 2003). Das kommt letztlich der Diskussion allgemeiner (ebd., S. 66ff.) und formaler Bildungsziele sehr nahe und übersieht neben der Bereichsspezifität des Wissens (Reusser 2001) auch Fragen der Erreichbarkeit dieser Ziele. Eine normative „gemeinsame Vision der Welt“ (ebd., S. 68) dürfte mit den Mitteln der Bildung kaum herstellbar sein. Es handelt sich um eine pädagogische Idealisierung, die weit über das hinausgeht, was sich empirisch beschreiben und pragmatisch herbeiführen lässt.⁴ Grundsätzlich ist die *Betrachtung der Ziele von Bildung und Lernen als allgemeine und bereichsspezifische Kompetenzen* vereinbar mit schulischem Unterricht. Kompetenzen spiegeln grundlegende Handlungsanforderungen, denen Schülerinnen und Schüler in einem Lernbereich, einer Domäne ausgesetzt sind. Schülerinnen und Schüler *erreichen* durch kumulatives Lernen Kompetenzen im Sinne der Fähigkeit, mit Hilfe erworbenen expliziten und impliziten Wissens mit diesen Anforderungen umzugehen. Die Leistungsverteilung und das Können – die individuellen und gruppenspezifischen Kompetenzniveaus und -profile – werden dabei nach wie vor verschieden sein. Dieser Tatbestand ist zu akzeptieren, solange gewisse Minimalziele oder Mindeststandards von allen Schülerinnen und Schülern erreicht werden.

Was früher „Fähigkeiten“ und „Fertigkeiten“ genannt wurde, kann jedoch nicht nur von den Zielen her erfasst werden, die für alle Schülerinnen und Schüler innerhalb gewisser Bandbreiten gleichsam gelten, sondern muss auch mit den unterschiedlichen *Prozessen ihrer Erreichung* verknüpft sein. Für die Schulfächer stellt sich insbesondere die Frage, wie deren Inhalte nicht nur auf Kompetenzziele hin formuliert, sondern auch *kompetenzbezogen gelehrt* werden können. Die zentrale Frage ist deshalb, wie der Prozess des kumulativen Kompetenzaufbaus durch schulischen Unterricht nicht nur besser als bisher erfasst, sondern auch individuell *gefördert* werden kann. Es ist offensichtlich, dass die für die Überprüfung von Standards konzipierten testdiagnostischen Instrumente und Verfahren dazu nicht ausreichen (vgl. Kap. 5.1.1).

Das erklärt, warum Kompetenztheorien mit zu entwickelnden neuen Verfahren der Rückmeldung, der Lernprozess- und der Förderdiagnostik – nicht zuletzt auf der Unterrichtsebene – in Verbindung gebracht werden müssen. Der gestufte Aufbau von Kompetenzen geschieht mangels ausgearbeiteter und für die

⁴ Weinert (2001a, S. 65) hat darauf verwiesen, dass allein in der deutschen Literatur über Berufsbildung mehr als 650 verschiedene „Schlüsselqualifikationen“ behauptet worden sind. Letztlich kann jedes allgemeine Ziel so genannt werden. Zudem dürfte es schwierig sein, zwischen „Schlüsselkompetenzen“ und „Kompetenzen“ sinnvoll zu unterscheiden.

Beteiligten transparenter Kompetenzmodelle in der Regel bislang ohne Vergleichsdaten und ohne Zusammenhang zwischen den Schulstufen in einem Unterricht, den die Lehrkräfte mit eigenen Interpretationen und (beim Fehlen verbindlicher Lehrmittel) auf der Basis ihres eigenen Stoff- und Lernverständnisses gestalten. Sie sind für die tatsächlich erreichten materiellen Ergebnisse des Unterrichts kaum je konkret verantwortlich, obgleich sie mehr oder weniger autonom über die Leistungen ihrer Schülerinnen und Schüler urteilen können. Kriteriumsbezogene und graduierbare Referenzstandards, die über das einzelne Klassenzimmer hinausreichen, fehlen weitgehend. Ohne fachinhaltliche Vergleichsmaßstäbe, die sich auf normative Erwartungen hinsichtlich minimal zu erreichender Ziele beziehen, ist es schwierig bis unmöglich, die Qualität von Schule und Unterricht zu bewerten. Und es ist dann auch schwierig, für die Prozesse des Lehrens und Lernens in öffentlichen Institutionen der Bildung Verantwortung zu übernehmen. *School Accountability* – verstanden als Rechenschaftslegung und als (begrenzte) Haftbarkeit für erreichte Lernleistungen – ist daher nicht zufällig zu einem Schlüsselthema geworden, das nicht nur in der amerikanischen Diskussion (O'Reilly 1996; Evers/Walberg 2002) bildungspolitischen Rang gewonnen hat.

2.3 Evaluationen und Accountability

Das Kernproblem der schulischen Qualitätssicherung ist in der internationalen Diskussion, wie gesagt, mit der Unterscheidung zwischen „Inputsteuerung“ und „Outputsteuerung“ gefasst worden. Der Aufbau von Erwartungen durch Bildungsprogramme und Bildungsziele, die Bereitstellung von Ressourcen und die Instrumentierung der Praxis sagen noch nichts über das Erreichen der Ziele und die Nutzung der Ressourcen oder Instrumente aus. Die erste Dimension lässt sich als Input verstehen, die zweite als Output. Neu ist die präzisere und folgenreiche Orientierung *auch* am Resultat und nicht nur an den Zielen oder Absichten.⁵ Es geht verstärkt um den erkennbaren Nutzen der Maßnahmen und des Einsatzes der Ressourcen, eine Frage, die dem idealistischen deutschen Bildungsdiskurs – nicht jedoch der Schuladministration und der Praxis – tradi-

⁵ „Der Output wird somit zum entscheidenden Bezugspunkt für die Beurteilung des Schulsystems und für Maßnahmen zur Verbesserung und Weiterentwicklung [...]. Nicht mehr durch detaillierte Richtlinien und Regelungen, sondern durch Definition von Zielen, deren Einhaltung auch tatsächlich überprüft wird, sorgt der Staat für Qualität“ (Klieme/Avenarius/Blum et al. 2003, S. 7).

tionell sehr fremd ist. Der Übergang von einer reinen Input- zu einer Input-Output-Steuerung ist nicht nur mit einer veränderten Sicht von Bildung verbunden, sondern er bedeutet auch „einen epochalen Schub an Objektivierung des schulischen Lernens“ (Messner 2004, S. 29).

Mit der Erweiterung der Perspektive sind verschiedene Fragen verbunden,

- wie Ziele in Prozesse und Effekte umgesetzt werden, also wie die tatsächliche Bildungsqualität zustande kommt,
- was die Schule im Unterschied zu anderen Einflussgrößen dazu beiträgt,
- wo ihre besondere Verantwortung liegt und
- wie sie aus Rückmeldungen darüber lernen kann, was sie selbst bewirkt.

Das Resultat wäre an den Fähigkeiten und Fertigkeiten – *Kompetenzen* – abzulesen, die die Schülerinnen und Schüler am Ende ihrer Schulzeit oder nach Abschluss gewisser Stufen tatsächlich erreicht haben. „Tatsächlich erreicht“ bezieht sich auf das individuelle Können in bestimmten fachlichen und überfachlichen Wissensbereichen nach längeren, schulisch beeinflussten Lernprozessen. Ziele und Effekte werden so in ein überprüfbares Verhältnis gesetzt, für das die Schule mit ihrem Auftrag zugleich Verantwortung übernimmt.

Die Frage der *accountability* im Sinne einer (Selbst-)Verpflichtung zu Bilanzierung, Rückmeldung von Daten und Rechenschaftslegung ist nicht grundsätzlich neu; sie ist jedoch schärfer zu formulieren, was auch heißt, die Ziele der Schulen präziser zu fassen. Diese waren im deutschsprachigen Bildungsraum bisher tendenziell vage gehalten und dienten allenfalls rhetorischen Zwecken. Damit ist auch gesagt, dass die Ziele selten auf externe und nur sehr bedingt auf interne Überprüfungen hin angelegt sind. Verantwortung im Sinne von *accountability* aber muss konkrete Gestalt annehmen und setzt Nachweise voraus. Auch und vornehmlich aus diesem Grunde sind Evaluationen der Schulqualität zu einer der stärksten bildungspolitischen Forderungen der letzten Jahre geworden.

Hier ist wiederum an historische Erfahrungen zu erinnern. Studien zur Evaluation von einzelnen Schulen oder Schuldistrikten sowie zu verschiedenen Verfahren der Evaluation finden sich in der amerikanischen Literatur seit dem Ende des 19. Jahrhunderts durchgehend bis heute. Die Beispiele reichen etwa von Arthur MacDonalds (1897/1898) psychometrischen Messungen der Schulleistungen über den kriteriengeleiteten Vergleich progressiver Ansätze öffentlicher Schulen in den zwanziger (Washburne/Stearns 1928) bis hin zu Evaluationen mit Testverfahren in den fünfziger Jahren (Sachs Adams/Torgerson 1956). Es gibt auch deutsche Beispiele früher Evaluationen des Wissensstandes

von Schulentlassenen, die darauf abzielten, über die Einzelbenotung hinaus die Qualität der Schulen abschätzen zu können (Müller-Jung 1928).

Auch die Organisation von Evaluationen ist kein neues Phänomen. Amerikanische Evaluationsagenturen wie die *National Study of School Evaluation* (NSSE) gibt es seit den dreißiger Jahren.⁶ Zusammenschlüsse von Privatschulen wie die 1947 gegründete *New York State Association of Independent Schools* widmen sich seit den frühen sechziger Jahren Aufgaben der Schulevaluation. Internationale Vergleichsstudien sind ebenfalls seit den sechziger Jahren des 20. Jahrhunderts durchgeführt worden.⁷ Daneben müssen auch die älteren Formen der Schulinspektionen beachtet werden, die in regelmäßigen externen Kontrollen der einzelnen Schulen bestehen. In England wurden die ersten Schulinspektorate 1839 gegründet,⁸ ein ähnliches System ist auch in deutschsprachigen Ländern entwickelt worden.⁹ Auch in diesen Formen der Evaluation ging es immer um die Frage, wie die Bildungsziele umgesetzt wurden und welche Ressourcen zu welchen Resultaten geführt hatten.

Das Problem der Überprüfung der Schulqualität ist also historisch vielfach bearbeitet worden, ohne dass einheitliche Modelle dafür verwendet worden wären. Aber eine entsprechende Praxis ist seit dem 19. Jahrhundert in vielen nationalen Bildungssystemen nachweisbar. Das wirft die Frage auf, warum die historischen Formen der Evaluation sich als wenig wirksam erwiesen haben. Eine erste Antwort bezieht sich auf die Punktualität solcher Unternehmungen; externe Evaluationen von Schulen waren oft Einzelprojekte, die nicht regelmäßig wiederholt wurden. Eine zweite Antwort bezieht sich auf die kaum elaborierten Methoden der Evaluation und eine dritte auf die fehlende Rückkopplung mit dem Praxisfeld. Schließlich kann als Erklärung der vergleichsweise geringen Wirksamkeit noch die oft fehlende Neutralität ins Feld geführt werden. Vielfach waren Schulevaluationen Bestätigungen von Reformervorstellungen und keine wirklichen Stärke-Schwäche-Analysen, oder sie wurden, wie in vielen Fällen

⁶ Die NSSE wurde 1933 als *Cooperative Study of Secondary School Standards* gegründet. Sie war ein Zusammenschluss von sechs regionalen Kommissionen zur Schulakkreditierung.

⁷ Die *International Association for the Evaluation of Educational Achievement* (IEA) geht zurück auf eine Tagung im Hamburger *UNESCO Institute for Education* im Jahre 1958. 1967 wurde aus dem lockeren Zusammenschluss von Forschern eine Rechtseinheit.

⁸ Zur Geschichte der englischen Schulinspektorate: Edmonds (1999).

⁹ Ein aus den Quellen gearbeitetes Beispiel ist Chantal Müllers Züricher Dissertation (2006) über die Entwicklung der Schulinspektorate im Schweizer Kanton Graubünden.

die Inspektorate, mit Loyalitätsbindungen verknüpft, die kein wirklich unabhängiges Urteil erlaubten.

Die Idee der Rückkopplung von Daten, der Rechenschaftslegung und damit der Übernahme von Verantwortung (*accountability*) hat der amerikanische Bildungsökonom Henry M. Levin 1974 als einer der Ersten in einem Aufsatz für die Zeitschrift *School Review* dargelegt. Die Verantwortung der Schule wird hier modellhaft mit der gezielten Evaluation ihrer Anstrengungen und der dabei je erzielten Resultate in Verbindung gebracht. Grundlegend heißt es bei Levin:

„Essentially, an accountability system is a closed loop reflecting a chain of responses to perceived needs or demands; an activity or set of activities that emerges to fill those demands; outcomes that result from those activities; and feedback on outcomes to the source of the demands“ (Levin 1974, S. 375).

Das Feedback kann neue Ansprüche (demands) erzeugen oder die bestehenden regenerieren. In beiden Fällen können die vorgängigen Aktivitäten bestätigt werden oder sich verändern, was auch für neue Aktivitäten gilt. Nach dem Feedback können neue oder alternative Ergebnisse (outcomes) erreicht werden. Der Loop ist vollendet, wenn das Feedback die „Quelle der Ansprüche“ (source of demands) erreicht, also den Auftraggeber oder den Abnehmer.

„The result is that an accountability system is continuous and dynamic when the linkages are tight and information is generated and transmitted freely“ (ebd.).

Seitdem ist das Thema politisch stark aufgewertet worden, ohne am grundlegenden Modell von Levin viel zu verändern. Es beschreibt, was Qualitätssicherung im Kern ausmacht und wozu sie in der Praxis dienlich ist. Die Verantwortung der Schule kann nicht abstrakt definiert werden, sondern verlangt die Beobachtung des Prozesses und so eine Datenbasis. Andernfalls kann jeder Akteur Verantwortung reklamieren, ohne zur Rechenschaft verpflichtet zu sein. Oder es kann jeder Verantwortung bestreiten, obwohl die Zuständigkeit gegeben ist. Die Rechenschaft über das Erreichte und der Nutzen des bisherigen Prozesses machen die zentrale Begründung hinter dem Modell aus.

Das Modell umfasst vier Komponenten, die sich so reformulieren lassen:

- *Ziele*: erreichbare Endzustände eines Prozesses;
- *Mittel*: Maßnahmen, die sicherstellen, dass der begonnene Prozess in Richtung Ziele verläuft;
- *Feedbackschleifen*: regelmäßige Überprüfung der erreichten Qualität;
- *Folgen*: Anpassung des Prozesses in Richtung Feedback-Daten.

Das ist die Basisidee hinter dem, was heute „Evaluation“ genannt wird. Evaluieren werden können die Leistungen der Schülerinnen und Schüler, die Kompetenzen der Lehrkräfte oder die Qualität der Schulen. Zugrunde gelegt wird sinngemäß oft das Modell von Levin. Qualität wird mit Hilfe von Feedbackschleifen im Prozess beschrieben. Damit wird es theoretisch möglich, Lernen nicht nur auf Ziele auszurichten, sondern den fortlaufenden Prozess zu beeinflussen. Es ist im Sinne dieses Prozessmodells nicht hinreichend, aus allgemeinen Lernzielen operative abzuleiten und dann dabei stehen zu bleiben; entscheidend ist, die tatsächlichen Erfahrungen einschätzen und gegebenenfalls auch korrigieren zu können. Die Schule übernimmt so Verantwortung für ihren eigenen Lernprozess.

Levin unterscheidet „goals“ und „objectives“. *Goals* sind allgemeine Bildungsziele, die nach Prioritäten geordnet und auf mehreren Ebenen abgestimmt sind. Sie reagieren auf politische Prozesse und stellen ein System von gesellschaftlichen Anforderungen dar, auf die das Bildungssystem reagieren muss, ohne dass zwischen diesen Anforderungen und dem, was Schulen leisten, ein direkter Zusammenhang besteht. *Objectives* sind explizite und überprüfbare Teilziele, die im Laufe eines Prozesses erreicht oder verfehlt werden können. Diese Ziele beschreiben den gewünschten Erfolg in verschiedenen Lerndomänen, aber sie dürfen nicht wie eine feste, letzte Größe betrachtet werden. Nicht nur die Ziele müssen transparent kommuniziert werden, sondern auch die Erträge. Das fällt dem Erziehungssystem schwer. Die Rhetorik der Ziele steht in keinem Verhältnis zur Diskussion der Resultate, die in offener Form gar nicht stattfindet (ebd., S. 380f.).

Soll sich das ändern, sind verschiedene Maßnahmen nötig. Was Levin in diesem Zusammenhang vorschlug, kommt dem ziemlich nahe, was die Systementwicklung anschließend bestimmt hat. Im Kern werden Schulen wie Unternehmen betrachtet, die Ziele erreichen und Bilanz ziehen müssen. Zu diesem Zweck sollen besondere Agenturen (educational services) eingerichtet werden, die Anreize schaffen, damit die gewünschten Resultate erreicht werden können, und die den Weg der Zielerreichung kontrollieren. Bislang – aus der Sicht von 1974 gesprochen – ist das nicht annähernd der Fall.

„The result is that schools now are held accountable not for explicit educational outcomes, but for explicit educational processes and inputs“ (ebd., S. 383).

Aber auch wenn Schulleitungen und Behörden (managers of school) nicht bewusst versuchen, Ziele zu erreichen und Leistungen zu steigern, produzieren Schulen doch Resultate, und dies unablässig. „The implementation of specific processes and inputs will create outcomes“ (ebd.). Outcomes können unmittelbar solche der Schule sein oder mittelbar solche, die der Gesellschaft zugute kommen oder in ihr Kosten verursachen.¹⁰ In jedem Fall, gewollt oder nicht gewollt, erzielt die Schule Effekte, die in den einen oder anderen Bereich fallen, wobei auch zwischen kurzfristigen und langfristigen Effekten unterschieden werden kann (ebd., S. 387).

Wenn nun evaluiert werden soll, ob diese Effekte den Zielen des Unterrichts entsprechen oder nicht, müssen sie gemessen werden (ebd., S. 383). Und die Daten der Evaluation müssen an die Verantwortlichen zurückgemeldet werden, damit bessere Entscheidungssituationen zustande kommen. „Before the various constituencies and managers of education can decide whether modifications are in order, they need good information on outcomes“ (ebd., S. 383f.). Das können Daten aus Einzelevaluationen, aber auch aus Longitudinalstudien sein, die ausdrücklich erwähnt werden (ebd., S. 385). Gesteuert werden damit Ziele, Ressourcen sowie Belohnungen und Sanktionen (ebd., S. 386). Das System lernt von dem, was es an Resultaten produziert, also durch verschiedene Formen von Feedback.

Was auf diesem Weg angestrebt wird, ist ein *Bildungsmonitoring*, ohne das nationale Bildungssysteme kaum wirksam gesteuert werden können. Auch die dabei auftretenden Probleme wurden 1974 bereits klar benannt:

„The fact that educational objectives of a specific nature rarely emerge from the political process and the lack of incentives for measuring outcomes lead to a situation where educational outputs *are not commonly monitored*. Moreover, there exist measurement problems in assessing outcomes, and there is a particularly severe evaluation problem in ascertaining the unique contribution of the edu-

¹⁰ „Educational outcomes can be thought of in two ways: proximate and ultimate. The proximate outcomes are those effects of schooling such as school induced increases in cognitive proficiencies and alterations of attitudes that are the immediate outputs of the educational process. The ultimate outcomes refer to the social effects of these changes in terms of their benefits or costs to society“ (Levin 1974, S. 383).

cational sector to particular proficiencies and behaviors“ (ebd., S. 387; Hervorhebung d. Verf.).

Grundlegend im Prozess der Qualitätssicherung ist gemäss diesem Modell die Orientierung am Resultat. „Resultate“ hat *jeder* Unterricht, wie auch jeder Unterricht mit Leistungsbewertung zu tun hat, Tests werden in jeder Schule geschrieben, und sie dienen immer der Feststellung, was gelernt wurde und was nicht (Weinert 2001b). Die Frage ist, was explizit formulierte Bildungsstandards an dieser Praxis ändern, welchen Gewinn sie mit sich bringen und welche Risiken mit ihnen verbunden sind. Zudem fragt sich, wie nationale oder regionale Bildungsstandards überprüft werden sollen und welche Rolle dabei den Leistungstests zukommt. Wenn ein Monitoring aufgebaut wird, dann muss auch eine Gewähr bestehen, dass die Daten an das System zurückvermittelt werden.

Die Evaluation von Schulen gemäß Leistungserwartungen oder Standards ist eine historisch gewachsene Praxis, die in das 18. Jahrhundert zurückreicht. Die Praxis der öffentlichen Prüfungen gehört in diesen Zusammenhang, ebenso die erwähnten Inspektorate oder auch Rekrutenprüfungen, die in der Schweiz bis heute üblich sind. Im Wesentlichen waren dies Maßnahmen der Verwaltung, die – wie gezeigt – im 20. Jahrhundert entweder stark zurückgenommen wurden oder kaum noch Erfolg hatten, vor allem weil sie der Profession entgegenstanden (Ravitch 2002). In den letzten zwanzig Jahre hat aber eine politische Entwicklung eingesetzt, die aus verstreuten und eher schwachen Konzepten der Überprüfung ein Modell hat werden lassen, das wohl nationale Anpassung verlangt, aber in bestimmten Grundelementen weitgehende Übereinstimmung zeigt. Das gilt im geschichtlichen Längsschnitt ebenso wie im internationalen Vergleich. Der Stand der Entwicklung lässt sich etwa an dem von Thomas Kellaghan (2002) herausgegebenen *International Handbook of Educational Evaluation* ablesen.

Bildungsstandards sind in diesem Modell ein zentraler Teil der Qualitätssicherung. Sie sollen entwickelt und eingeführt werden, um den angesprochenen Prozess der Outputsteuerung zu intensivieren und zu beschleunigen. Ohne Standards wären auch Rückmeldesysteme nur individuell möglich. Die heutige Diskussion bezieht den Begriff „Bildungsstandards“ überwiegend auf die Inhalte (contents) bestimmter Fächer (domains) in bestimmten Klassenstufen (grades), und dies mit Blick auf einen messbaren Output. Bildungsstandards im engeren Sinne beziehen sich auf die Zielerreichung auf verschiedenen Stufen oder Kompetenzniveaus, die regelmäßig und fortlaufend getestet werden soll. Die curricularen Gehalte der Schulfächer erhalten dabei eine neue Form, näm-

lich als Aufgabenkulturen, die sich auf den curricularen Kern eines Faches oder eines Lernfeldes beziehen.

Der Ausdruck „Standards“ wird aber auch in einem weiteren Sinne gebraucht, etwa im Blick auf Schul- und Unterrichtsqualität, Ausstattung, Ressourcennutzung oder auch Anforderungen an das schulische Personal. Standards sind also nicht nur curriculare Gehalte. Der Begriff muss über das hinausgehen, was bisher in Lehrplänen festgeschrieben war. Standards sollen Lehrpläne nicht ersetzen, aber selbst elaborierte Neufassungen von Lehrplänen sind für sich genommen kaum geeignet, für eine Verbesserung der Unterrichtspraxis zu sorgen. Lehrpläne orientieren sich an dem tatsächlichen Unterricht nur sehr begrenzt. Sie erfüllen einen anderen Zweck, indem sie nämlich das didaktische Feld ordnen, den Zusammenhang der verschiedenen Fächer oder Lernbereiche definieren und die Verteilung der Ressourcen begründen.

2.4 Konzepte von Bildungsstandards

Die amerikanische Historikerin Diane Ravitch hat vor mehr als zehn Jahren einen Vorschlag vorgebracht, wie sich der Ausdruck „Bildungsstandards“ begrifflich differenzieren lässt. Der Vorschlag hat sich als nützlich erwiesen. In der deutschsprachigen Literatur wird darauf nur selten explizit zurückgegriffen, implizit aber werden ähnliche Unterscheidungen vorgenommen. Für den angelsächsischen Bereich fasst Ravitch zusammen und expliziert, was sich im bildungspolitischen Diskurs seit den siebziger Jahren entwickelt hat. Ihre dreiteilige Theorie der Standards ist zur Problembestimmung hilfreich und sorgt für die notwendigen Differenzierungen.

Allgemein wird der Ausdruck „Standard“ mit einer doppelten Bedeutung gebraucht, als Ziel und als Maß der Zielerreichung.

„A standard is both a *goal* (what should be done) and a *measure* of progress toward that goal (how well it was to be done)“ (Ravitch 1995, S. 7).

Der Ausdruck „Standards“ wird in der heutigen pädagogischen Kommunikation ganz unterschiedlich gebraucht. Aber wenn der Ausdruck Sinn ergeben soll, muss er eine bestimmte Minimalbedingung erfüllen: „A Standard is not useful or meaningful unless there is some way to measure whether it is reached“ (ebd., S. 11). Das unterscheidet Standards von den bisherigen Zielen. Für sie war wichtig, wie sie zustande kamen und auf welchen Ebenen sie formuliert wurden. Ob sie erreicht wurden oder nicht, war vielleicht eine Frage, aber es

gab nur sehr begrenzt Verfahren, diese Frage auch zu beantworten. Vor allem konnte nicht gesagt werden, ob über die einzelne Klasse und die jeweiligen Noten hinaus Ziele der allgemeinen Bildung in der Breite erreicht wurden oder nicht.

Ravitch unterscheidet drei zusammenhängende Typen von Standards:

- Festlegung der Gehalte: *content standards*;
- Festlegung der Leistungsniveaus: *performance standards*;
- Festlegung der Ressourcen, von denen die Möglichkeiten des Lernens bestimmt werden: *opportunity-to-learn standards*.

Was gelernt werden soll, muss möglichst klar und eindeutig bestimmt sein, damit sich die Erwartungen auf beiden Seiten, bei den Lehrenden wie den Lernenden, auf einen gehaltvollen und effektiv genutzten Prozess einstellen können. Das ist nicht trivial, legt man zugrunde, dass in vielen Schulen Absprachen zwischen den Lehrkräften kaum bestehen, Lehrpläne nicht den Unterricht bestimmen und Inhalte oft sehr individuell zustande kommen. Auch fehlen vielfach Prioritäten. Damit aber werden die Lernerwartungen unspezifisch oder diffus, während es darauf ankommt, sie möglichst präzise zu formulieren und zu kommunizieren.

Diane Ravitch schreibt dazu:

„Content standards (or curriculum standards) describe what teachers are supposed to teach and students are expected to learn. They provide clear, specific descriptions of the skills and knowledge that should be taught to students“ (ebd., S. 12).

Die curricularen Standards oder die Inhalte des Unterrichts müssen so spezifisch gefasst werden, dass sie von Lehrkräften, den Schülerinnen und Schülern, den Eltern und anderen möglichst schnell und ohne Komplikationen verstanden werden. Und die Standards müssen so klar formuliert sein, dass die Lehrkräfte wissen, was von den Schülerinnen und Schülern erwartet wird, und sie ihren Unterricht so einrichten, dass er den Schülerinnen und Schülern hilft, auch tatsächlich lernen zu können, was von ihnen erwartet wird. Ansonsten kommt der Lernerfolg mehr oder weniger zufällig zustande und wird sehr stark von der sozialen Herkunft bestimmt (ebd.). In diesem Sinne reagieren klare Standards auf die Forderung der Chancengleichheit.

Aber die Präzisierung der inhaltlichen Standards und damit der Lernerwartungen ist nicht genug. Es müssen auch Niveaus unterschieden werden, also

das, was minimal und maximal erreicht werden kann. Die Differenzierung von Leistungsstufen oder Niveaus der Performanz ist ein entscheidender Schritt. Hinter den traditionellen Lehrplänen stand die Erwartung, dass alle Schülerinnen und Schüler die gleichen Ziele erreichen sollen oder müssen. Das ist nicht möglich, aber nur das sieht ein Lehrplan vor. Erst eine graduierte Zielerreichung und mithin am Ende gestufte Kompetenzen schaffen auch einen realistischen Erwartungshorizont. Und nur unter dieser Voraussetzung sind auch gezielte, auf eine heterogene Schülerschaft abgestimmte Maßnahmen einer adaptiven Förderung möglich. Es ist sicher kein Zufall, dass die Angaben zur domänen- oder fachspezifischen Förderung in heutigen Lehrplänen höchst vage gehalten werden, falls sie überhaupt vorkommen.

Allgemein schreibt Diane Ravitch zu diesem Punkt:

„Performance standards define degrees of mastery or levels of attainment. They answer the question: ‘How good is good enough?’ Performance standards describe what kind of performance represents inadequate, acceptable, or outstanding accomplishment“ (ebd., S. 12f.).

Dieser Vorschlag zur Stufung ist wiederum nicht neu, er kommt in verschiedenen Papieren der neunziger Jahre vor, auf die Ravitch sich bezieht, und die Idee der gestuften Leistung steht ja auch hinter dem historischen Notenschema, das immerhin seit dem 17. Jahrhundert in Gebrauch ist. Neu sind auch hier die Präzisierung und das Streben nach vergleichbaren Leistungsanforderungen. Die einzelnen Stufen der Leistung müssen wie die inhaltlichen Standards klar beschrieben und verständlich kommuniziert werden, die Lernenden müssen vorher wissen, was sie erreichen können und was dabei jeweils verlangt wird. Die heutige Notengebung legt in aller Regel erst nach dem Vollzug der Leistung die Kriterien der Bewertung an, und oft sind diese Kriterien implizit, werden also nicht *mit* der Note kommuniziert.

Zentral ist schließlich auch die Frage der Ressourcen. Sie bestimmen, was Diane Ravitch die *Möglichkeit*, Lernen zu können, nennt.

„Opportunity-to-learn standards define the availability of programs, staff, and other resources that schools, districts, and states provide so that students are able to meet challenging content and performance standards“ (ebd., S. 13).

Hinter dieser Formulierung steht die politische Forderung, dass die Erwartungen an die Lehrenden und Lernenden nicht erhöht werden können, ohne zugleich die Ressourcen zu verbessern.¹¹ Das gilt unabhängig von der Tatsache, dass einzelne Schulen mit geringen Ressourcen und hohen sozialen Nachteilen erfolgreich sein können. Es ist ein wenig beachtetes Resultat der ersten PISA-Studie, dass die vor Ort vorhandenen *Möglichkeiten* des Lernens, also die schulischen Ressourcen, tatsächlich den gewichtigsten internen Faktor ausmachen, um die Leistungen der Schülerinnen und Schüler zu beeinflussen (OECD 2001, S. 241ff.).

Ressourcen sind zunächst Finanzmittel, aber sie schlagen sich nicht nur in Budgets nieder. *Alles*, was zum Lernen genutzt werden kann, ist eine Ressource, die Kompetenz der Lehrkräfte ebenso wie die Qualität der verfügbaren Lehrmittel und Medien, die Fortbildungs- und Unterstützungssysteme, die Lern- und Arbeitsbedingungen oder auch das Interesse und die Wertschätzung, die sich mit einem Fach oder einem Lehrgang verbinden. Ressourcen dürfen also nicht lediglich materiell verstanden werden, oft sind symbolische Ressourcen für den Lernerfolg ebenso wichtig wie die materielle Ausstattung. Und auch symbolische Ressourcen lassen sich standardisieren, man denke an Logos, didaktische Markierungen oder Leitbilder. Dieser weite Begriff von Ressourcen ist inzwischen auch in der Unterrichtsforschung genutzt worden (Cohen/Raudenbush/Lowenberg Ball 2002).

Unter „Standards“ lassen sich also mit Diane Ravitch Festlegungen von Inhalten, Zielen oder Maßstäben und Ressourcen verstehen, mit denen die Lernprozesse von Schülerinnen und Schülern beeinflusst und verbessert werden sollen. Es handelt sich um Fixierungen, die im gesamten Geltungsbereich von Bildungsprogrammen Anwendung finden sollen. Die Festlegungen können je nach Verlauf der Erfahrungen verändert werden, dies jedoch möglichst unter der Voraussetzung von Evaluationsdaten. Damit ist ein bestimmter Schluss verbunden: Was nicht in einem Lernprogramm enthalten ist, nicht abverlangt wird und sich der Kontrolle entziehen kann, wird mit hoher Wahrscheinlichkeit nicht gelernt. Alle drei Kriterien sind zu beachten: die Aufnahme von Aufgaben und Leistungen in ein Programm, das Abverlangen der Leistung und die Überprüfung der Zielerreichung.

¹¹ Ursprünglich hieß dieser Bereich bei den Kongressdebatten unter der Administration Bush sen. „school delivery standards“. In der Administration Clinton wurden daraus „opportunity-to-learn standards“ (Ravitch 1995, S. 13).

Die drei Dimensionen der Standards spielen zusammen und sind untereinander verknüpft:

„Content standards without performance standards are meaningless. Content standards define what is to be taught and learned; performance standards describe how well it has been learned. Similarly, opportunity-to-learn standards cannot stand on their own, without content and performance standards“ (Ravitch 1995, S. 13).

Auch Ravitch schlägt Tests vor, die sich aber nicht einfach auf Schulwissen beziehen sollen, das wie im konventionellen Unterricht abgefragt wird, sondern mit denen die Kompetenzen zur Problemlösung gemessen werden. Das geht auf einen Vorschlag von Robert Glaser (1981) zurück und setzt ein aktives, kognitives Konzept von Lernen voraus. Schülerinnen und Schüler „empfangen“ im Unterricht nicht einfach Wissen, sondern *konstruieren* es auf der Basis ihrer eigenen Erfahrungen, ihres Vorwissens und ihres Verstehenshorizontes. Schulen können also nicht „aus dem Nichts“ Wissen aufbauen, sondern nur die gegebenen Erfahrungen und das vorhandene Wissen erweitern (Ravitch 1995, S. 103).

Von kritischer Bedeutung ist das Vorwissen der Schülerinnen und Schüler: „The more they know, the more they can learn“ (ebd., S. 104). Am Besten lernen sie, wenn sie Wissen, Fertigkeiten (skills) oder Problemlösungen auf reale Situationen anwenden können und wenn sie dabei mit anderen kooperieren. Das Lernen ist domänenspezifisch, jeder größere Lernbereich hat eigene kognitive Strategien, die nicht allgemeingültig sind. Wer zum Beispiel ein Postulat wie „kritisches Denken“ zum Bildungsziel erhebt, so 1985 die American Federation of Teachers (AFT), kann dieses Ziel nicht allgemein verwirklichen. „Knowledge of the subject matter is necessary“ (ebd.). Auch dieser Befund steht im Widerspruch zu vielen rhetorischen Zielen, denen sich die Schule gegenübersteht und die ihr politisches Umfeld bestimmen, die sie aber mit dem, was sie kann und leistet, nie wirklich anstrebt.

Was getestet wird, sind die Verstehenskompetenz und die Anwendung des Gelernten in eigenen Problemlösungen. „Assessment must be designed both to find out whether students *understand* what they have learned, and to enable them to demonstrate what they know“ (ebd.). Damit ist ein enger Zusammenhang zwischen Unterricht, Curriculum und Tests gegeben, den zuerst wohl Resnick/Resnick (1983) beschrieben haben. Auch Formen der Leistungsüberprüfung durch konventionelle Tests dienen der Evaluation, aber sie wirken nicht zurück auf das Curriculum und den Unterricht. Ähnlich argumentiert

Glaser (1986). Der Zweck des Tests ist die Verbesserung des Unterrichts und des Lernens, nicht einfach die Erreichung von Graden oder die Einführung von Rankings (Ravitch 1995, S. 104). Im Idealfall erreichen die Testergebnisse den Unterricht und tragen dazu bei, bestimmte Konzeptionen des Lehrens und Lernens vom Resultat her zu verbessern.

Damit wird die amerikanische Diskussion bis Mitte der neunziger Jahre aus einer bestimmten Sicht zusammengefasst, wobei schon zu diesem Zeitpunkt die Frage des Testens, also die empirische Überprüfung der Standard-Erreichung, in den Mittelpunkt rückte. Ein wichtiger Ausgangspunkt für die gesamte Standards-Bewegung war der Report *A Nation at Risk* (National Commission on Excellence in Education 1983), der 1983 die verbreitete Krisenstimmung zusammenfasste und die neuerliche Diskussion über Standardisierung des amerikanischen Bildungswesens in Gang brachte. Die Terminologie von *A Nation at Risk* erinnert stark an Levin und die Vorschläge der Bildungsökonomie aus den siebziger Jahren. Seitdem ist intensiv versucht worden, für Schulfächer auf verschiedenen Stufen Standards zu entwickeln und die Ergebnisse zu testen.

Das maßgebende deutsche Gutachten *Zur Entwicklung nationaler Bildungsstandards* (Klieme/Avenarius/Blum et al. 2003) geht ebenfalls von einem inhaltlichen Begriff der Bildungsstandards aus. Das Gutachten reagiert auf die für Deutschland unvorteilhaften Ergebnisse der TIMS-Studie und der ersten PISA-Studie mit Empfehlungen, die darauf abzielen, die Curricula zentraler Fächer stärker als bislang festzulegen und die Ergebnisse des Lernens zu überprüfen. In dieser Hinsicht entspricht der Ansatz den amerikanischen Konzepten. Hier wie dort geht es um eine Verstärkung der Outputsteuerung und so der Orientierung am Resultat. Konzeptionell orientiert sich das Gutachten stark am Konzept der kognitiven Kompetenz.

Allgemein heißt es zum Begriff:

„Bildungsstandards formulieren Anforderungen an das Lehren und Lernen in der Schule. Sie benennen Ziele für die pädagogische Arbeit, ausgedrückt als erwünschte Lernergebnisse der Schülerinnen und Schüler. Damit konkretisieren Standards den Bildungsauftrag, den allgemein bildende Schulen zu erfüllen haben“ (ebd., S. 19).

„Konkretisieren“ kann im Sinne von Diane Ravitch und anderen *Präzisieren* und *Spezifizieren* heißen, damit „Ziele“ als Leistungserwartungen erkennbar werden. Die Erwartungen richten sich auf Kompetenzen, die die Schule ihren Schülerinnen und Schülern vermitteln muss, „damit bestimmte zentrale Bildungsziele erreicht werden“ (ebd.). Unterschieden wird noch nicht zwischen fachlichen

und überfachlichen Kompetenzen, und es wird auf dieser Stufe der Abstraktion auch keine Einschränkung der Geltung der Aussagen etwa nur für bestimmte Fächer oder Lernbereiche gemacht. Generell wird festgehalten:

„Die Bildungsstandards legen fest, welche Kompetenzen die Kinder oder Jugendlichen bis zu einer bestimmten Jahrgangsstufe erworben haben sollen. Die Kompetenzen werden so konkret beschrieben, dass sie in Aufgabenstellungen umgesetzt und prinzipiell mit Hilfe von *Testverfahren* erfasst werden können“ (ebd.).

Damit hat der Begriff „Bildungsstandards“ drei Dimensionen, die sich von den Typen bei Ravitch unterscheiden. Die Dimensionen sind:

- Bildungsziele,
- Kompetenzen und Aufgabenstellungen sowie
- deren Überprüfung durch Tests.

Das wird im Einzelnen so dargestellt: Bildungsstandards orientieren sich grundsätzlich an gesellschaftlichen Bildungszielen, die sie in konkrete Anforderungen umsetzen. Bildungsziele formulieren allgemeine Erwartungen mit „relativ allgemein gehaltenen Aussagen“, die einen verpflichtenden Rahmen darstellen (ebd., S. 20).

„Die Ebene der Bildungsziele ist handlungsleitend für die Entwicklung des Schulsystems, für die Profilierung einzelner Schulen, für die Gestaltung von Unterricht durch die Lehrerinnen und Lehrer und nicht zuletzt für die Erstellung von Stoffplänen und didaktischen Ansätzen zur Umsetzung einzelner Themen. Auf dieser Ebene drückt sich der gesellschaftliche Anspruch von Schule aus, hier ist die Ebene des politisch Wünschbaren, begründet in einem gesellschaftlichen Entwicklungskonzept. Dies ist auch die Ebene, auf der Innovationen ansetzen und grundlegende Kurskorrekturen vorgenommen werden“ (ebd.).

Für die Umsetzung dieser Ziele oder allgemeinen Erwartungen ist ein eigenes „Medium“ erforderlich, „in dem sich die Ziele spezifizieren und definieren lassen“. Dieses Medium sind traditionell die Lehrpläne, die durch Kompetenzmodelle und Aufgabekulturen ergänzt werden sollen (ebd., S. 21). Bildungsstandards konkretisieren daher die allgemeinen Ziele „in Form von Kompetenzanforderungen“ (ebd.).

„Sie legen fest, über welche Kompetenzen ein Schüler, eine Schülerin verfügen muss, wenn wichtige Ziele der Schule als erreicht gelten sollen. Systematisch geordnet werden diese Anforderungen in Kompetenzmodellen, die Aspekte, Abstufungen und Entwicklungsverläufe von Kompetenzen darstellen“ (ebd.).

Grundlegend für die Entwicklung von Bildungsstandards sind also Kompetenzmodelle, die fachbezogen verstanden werden. Diese Strategie knüpft *nicht* an die populäre Unterscheidung von Sach-, Selbst- und Sozialkompetenz an, wie sie in vielen heutigen Lehrplänen üblich ist. Bildungsstandards werden mit Hilfe von Kompetenzmodellen als Leistungsdispositionen für einzelne *Fächer* oder *Lernbereiche* entwickelt.

„Ein solches Kompetenzmodell unterscheidet Teildimensionen innerhalb einer Domäne (also z.B. Rezeption und Produktion von Texten, mündlichen und schriftlichen Sprachgebrauch), und es beschreibt jeweils unterschiedliche Niveaustufen auf solchen Dimensionen. Jede Kompetenzstufe ist durch kognitive Prozesse und Handlungen von bestimmter Qualität spezifiziert, die Schüler auf dieser Stufe bewältigen können, nicht aber Schüler auf niedrigeren Stufen. Zum Bildungsstandard gehört, dass für einzelne Jahrgänge festgelegt wird, welche Stufen die Schülerinnen und Schüler erreichen sollen“ (ebd., S. 22).

Die Bildungsstandards ihrerseits konkretisieren sich in Aufgabenstellungen, deren Bearbeitung anzeigt, welche Kompetenzstufe erreicht wurde. Das Ergebnis wird mit Hilfe von Leistungstests erfasst. Testaufgaben prüfen, ob die Lernenden „das angestrebte Ergebnis oder Handlungspotential entwickelt“ haben oder nicht (ebd., S. 23). Zu diesem Zweck müssen Aufgaben generiert und auf ihre Validität hin geprüft werden. Die Prüfung allein reicht aber nicht aus. Die Ergebnisse müssen Schule und Unterricht erreichen, das *Feedback* wird daher – wie von Henry M. Levin vorhergesagt – zu einem Schlüsselproblem der Systementwicklung.

„Feedback über den erreichten Kompetenzstand ist [...] ein unverzichtbarer Bestandteil einer kontinuierlichen, systematischen Qualitätsentwicklung. Tests können Auskunft geben über die Erreichung von Bildungszielen. Diese Information kann für die Entwicklung von Schule, Unterricht und Didaktik genutzt werden“ (ebd., S. 24).

Geeignete sind von weniger geeigneten Bildungsstandards zu unterscheiden. Hier stellt sich ein Qualitätsproblem eigener Art, wie auch schon die Kritik an

den Tests dargelegt hat. Merkmale guter Bildungsstandards werden in sieben Dimensionen beschrieben:

- (1) **Fachlichkeit**
Bildungsstandards sind jeweils auf einen bestimmten Lernbereich bezogen und arbeiten die Grundprinzipien der Disziplin bzw. des Unterrichtsfaches klar heraus.
- (2) **Fokussierung**
Die Standards decken nicht die gesamte Breite des Lernbereiches bzw. des Faches in allen Verästelungen ab, sondern konzentrieren sich auf einen Kernbereich.
- (3) **Kumulativität**
Bildungsstandards beziehen sich auf die Kompetenzen, die bis zu einem bestimmten Zeitpunkt im Verlaufe der Lerngeschichte aufgebaut worden sind. Damit zielen sie auf kumulatives, systematisch vernetztes Lernen.
- (4) **Verbindlichkeit für alle**
Standards drücken die Mindestvoraussetzungen aus, die von allen Lernenden erwartet werden. Diese Mindeststandards müssen schulformübergreifend für alle Schülerinnen und Schüler gelten.
- (5) **Differenzierung**
Die Standards legen nicht nur eine „Messlatte“ an, sondern differenzieren auch zwischen Kompetenzstufen, die über und unter bzw. vor und nach dem Erreichen des Mindestniveaus liegen. Sie machen so Lernentwicklungen verstehbar und ermöglichen weitere Abstufungen und Profilbildungen, die ergänzende Anforderungen in einem Land, einer Schule oder einer Schulform darstellen.
- (6) **Verständlichkeit**
Die Bildungsstandards sind klar, knapp und nachvollziehbar formuliert.
- (7) **Realisierbarkeit**
Die mit den Bildungsstandards gegebenen Anforderungen stellen eine Herausforderung für die Lehrenden und Lernenden dar, sind aber mit realistischem Aufwand erreichbar (ebd., S. 24f.).

Die Funktion von Bildungsstandards lässt sich auf dieser Basis mit einem Satz beschreiben: „*Sie arbeiten in klarer und konzentrierter Form heraus, worauf es in unserem Schulsystem ankommt.* Sie konzentrieren den pädagogischen Auftrag der Schule in zentralen Lernfeldern und richten somit das Lehren und Lernen auf gemeinsame Ziele aus“ (ebd., S. 47). Es sind Standards, die nicht für einen bestimmten

Schultyp gelten, sondern die übergreifend von allen Schülerinnen und Schülern einer bestimmten Stufe erreicht werden können.

Generell bestätigt die Expertise die Annahme, dass eine neue Stufe der Qualitätssicherung nur mit einem neuen Verhältnis von Input- und Outputorientierung erreicht werden kann. Allerdings geht es in der Expertise nicht um ein „radikales Modell“ einer am Output orientierten Steuerung, die sich allein und ausschließlich vom Ergebnis leiten ließe (ebd., S. 93). Empfohlen wird eine Mischung von Bildungsstandards, Kerncurricula und Kanonorientierung (ebd., S. 95ff.). Andererseits sind Bildungsstandards nicht einfach Ableitungen aus Curricula und Lehrplänen. Standards bestimmen „konkrete Kompetenzanforderungen“ und beinhalten somit „einen Maßstab zur Bewertung von Lernergebnissen“ (ebd., S. 99).

Vor allem diese Idee liegt den politischen Erwartungen zugrunde. „Bildungsstandards definieren Maßstäbe, nach denen der ‚Output‘ erfasst und bewertet werden kann“ (ebd.). In der Expertise geht es um „Feedback über Lernergebnisse“, die sich auf drei Ebenen darstellen lassen sollen. Die Untersuchung von Schülerleistungen mit Hilfe von Tests kann dazu dienen,

- generell, regelmäßig und auf nationaler Ebene das Erreichen von Bildungsstandards zu kontrollieren (*Bildungsmonitoring*);
- Rückschlüsse auf den Erfolg schulischer Programme oder unterrichtlicher Maßnahmen zu ziehen und damit eine Grundlage für die Verbesserung von Schule zu haben (*Schulevaluation*);
- etwas über den einzelnen Schüler zu erfahren, über seine Stärken und Defizite, damit gezielte Fördermaßnahmen getroffen werden können (*Individualdiagnostik*) (ebd.).

Diese drei Ebenen bestimmen die Erwartung der Outputsteuerung. Tests sollen nicht nur der Datenerhebung dienen, vielmehr sollen sich die Ergebnisse ins System rückübersetzen lassen. Daher ist auch hier Feedback grundlegend (vgl. Kap. 5.1.1); Tests sind so Teil einer umfassenden Qualitätssicherung, nicht deren einzige Größe.

2.5 Implementation und Schulentwicklung

Die Implementation von Bildungsstandards verlangt mehr als eine kohärente Politik. Die sprachliche Formulierung der Standards und ihrer Ausführungsbedingungen ist interpretationsabhängig (Hill 2001). Was politisch intendiert ist, muss mit den stets auch von eigenen Deutungen ausgehenden Akteuren ko-

konstruiert und auf Verständigung hin kommuniziert werden. Dies setzt einen sozialen Prozess voraus. Die Sprache der Politik ist oft missverständlich, wie es auf der anderen Seite auch ein Misstrauen der Akteure gibt. Besonders schwierig ist es, die positiven Seiten notorisch umstrittener Konzepte zu vermitteln. Allein der Ausdruck „Standards“ löst bei nicht wenigen Akteuren Reaktionen der Abwehr aus, die von einem Konzept der Implementation nicht einfach ignoriert werden können.

Aber es geht nicht nur um das Profil der Kommunikation. Wer Standards im System verankern will, muss soziale Kontexte, Ebenen und Prozesse unterscheiden. In der politischen Theorie ist es seit langem üblich, zwischen Mikro- und Makroprozessen zu unterscheiden, also nicht von einem einheitlichen und geschlossenen System auszugehen (Eulau 1969). Es ist auch seit längerem klar, dass die Kontrolle von Bildungsprozessen verschiedene Instanzen kennt, die bei der Implementation eine unterschiedliche Rolle spielen. Die lokale Ebene ist immer unterschieden von derjenigen der übergeordneten Behörden oder der staatlichen Instanzen (Koerner 1968). In der neueren politischen Literatur zur Nachhaltigkeit ist darauf verwiesen worden, dass dem Konzept nur dann praktische Relevanz zukommt, „if it can accommodate local preferences and capabilities“ (Victor 2006, S. 99).

„Implementation“ ist ursprünglich ein Begriff aus den Ingenieurwissenschaften, der später von der Computertechnologie übernommen wurde. Hier bezeichnet der Ausdruck die praktische Anwendung einer Methodologie oder eines Algorithmus. Der Ausdruck wird seit längerem auch in der politischen Wissenschaft verwendet, etwa im Blick auf die Implementation von Politikstrategien oder neuen Wegen politischer Entscheidungen. Allgemein lässt sich darunter die Umsetzung von Zielen oder Strukturvorgaben in einem differenzierten und sich entwickelnden sozialen System verstehen, dessen Rahmenbedingungen beachtet werden müssen. Etablierte soziale Systeme wie Schulen reagieren auf neue Implementationsstrategien mit einschlägigen Erfahrungen, welche ebenfalls berücksichtigt werden müssen. Hier ist oft die Innovation beweispflichtig, dann nämlich, wenn die Akteure Analogien zu schlechten oder misslungenen Strategien der Vergangenheit ziehen können.

Im Bildungsbereich sind Strategien politischer Implementationen wiederum nichts Neues. Das amerikanische Curriculum am Ende des 19. Jahrhunderts war geradezu „a definer of standards“ (Resnick/Resnick 1985, S. 6), also der Garant dafür, dass überhaupt Festlegungen über einzelne Schulen oder Distrikte hinaus kommuniziert werden konnten und auch Akzeptanz fanden. Durchgesetzt wurden die Standards etwa der sich entwickelnden High School nicht auf dem Wege der Verordnung, sondern durch sehr verschiedene Wege

der Implementation, die von der Gestaltung der Lehrmittel über die Ordnung der Zeitstruktur bis zur Ausbildung der Lehrkräfte reichten. Grundfragen blieben immer offen (ebd., S. 7), und eine einheitliche Standardisierung hat sich selbst auf lokalem Niveau nie erreichen lassen. Standards sind keine starren Größen, die überall gleich realisiert werden. Schuleingangstests etwa waren von Schule zu Schule verschieden.

Die Implementation komplexer schulischer Neuerungen lässt sich so auch als gebührende Anpassung neuer Aufgaben oder Ziele verstehen, die nicht nur mit Transfer zu tun hat, sondern auch mit Problemen der Übersetzung – verstanden als Entwicklungsprozess – rechnen muss. Der Prozess setzt immer lokale Theorien darüber voraus, ob und wie der Wandel vollzogen werden muss (Spillane 2002a). Letztlich erfolgt die Implementation von Bildungsstandards auch nicht einfach in Regelsystemen, sondern in Mentalitäten, die nur zum Teil professionell geformt sind. Adressat von Implementationsstrategien sind nicht nur die Lehrkräfte und die Bildungsverwaltung, sondern auch die Eltern, die Schülerinnen und Schüler, die Medien und nicht zuletzt die lokale Öffentlichkeit. Die Kommunikation mit diesen verschiedenen Akteursgruppen bringt divergierende Interpretationen mit sich, die sich auf ein bestimmtes Konzept von Bildungsstandards beziehen müssen. Dieses Konzept muss überzeugen und kann nicht einfach verordnet werden.

Die Entwicklung der Implementation von bildungspolitischen Programmen lässt sich als Lernprozess eigener Art verstehen. Im Blick auf die amerikanischen Erfahrungen seit etwa 1960 sind zu Beginn der neunziger Jahre drei Phasen unterschieden worden, die einen allmählichen Wechsel von großflächigen Interventionsprogrammen hin zu lokaler Entwicklungsarbeit anzeigen (Odden 1991). „Großflächig“ heißt, dass die frühen Programme ohne Beachtung der lokalen Unterschiede und mit linearen Strategien der Intervention arbeiteten. Förderprojekte wie *Head Start* waren so konzipiert, dass überall möglichst gleiche Effekte erzielt werden sollten. Entsprechend wurde eine zentrale Interventionsstrategie gewählt, die implizit voraussetzte, das Bildungssystem reagiere an allen vergleichbaren Stellen in etwa gleich, was historisch nie der Fall war.

Die frühen Programme der sechziger und siebziger Jahre zeigen also bei allen Unterschieden in den Inhalten und Zielsetzungen einige Gemeinsamkeiten im Hinblick auf die Strategien und Effekte der Implementation. Regierungsprogramme hatten keinen oder nur einen unzureichenden bis feindlichen (hostilen) Kontakt zur lokalen Ebene (ebd., S. 5). Die lokalen Behörden verfügten weder über die Kapazitäten noch waren sie willens, den Initiativen zu folgen, die höhere Stellen für sie konzipiert hatten (ebd.). Oft wurden die Fonds zweckentfremdet verwendet, der Kontrollaufwand stieg, ohne den eigentlich

beabsichtigten Effekt zu erreichen, nämlich Schule und Unterricht zu verbessern.

„Early implementation research showed that implementation problems not only emerged from faulty program designs but also, and even more importantly, from the policy’s relationship to the local institutional setting“ (ebd.).

Ein Ergebnis der frühen Implementationsforschung lautete daher: „Higher level government programs simply did not work, and [...] local governments would never implement them faithfully“ (ebd.). Die aus diesen Befunden resultierende Kritik der staatlichen Bildungsreformen (z.B. Boyd 1987) legte Alternativen der Forschung nahe, die sich nicht auf die *anfängliche* Implementation eines Programms richteten, sondern auf dessen tatsächliche Entwicklung, also die *abgeschlossene* Implementation. Auf diesem Wege konnte gezeigt werden, dass keineswegs alle Regierungsprogramme scheitern, sondern dass unter bestimmten Voraussetzungen sehr wohl gesagt werden kann: *federalism works* (Peterson/Rabe/Wong 1986). Die Aussage trifft zu, wenn Entwicklungsstufen beachtet werden und die lokalen Gegebenheiten eine Rolle spielen.

Aber eine erfolgreiche Implementation, also die Verankerung eines Programms oder einer Politik in einer Organisation, ist nicht genug. Damit verbunden wurde ein naheliegender Einwand:

„Claiming that programs get implemented [...] is not the same as claiming that they are effective, i.e. that they solve the problem for what they were created“ (Odden 1991, S. 8).

Die dritte Phase der neueren Implementationsforschung hat bis zu Beginn der neunziger Jahre zwei verschiedene Positionen hervorgebracht. Die eine geht davon aus, dass vor allem die mit einem neuen Programm verbundenen Instrumente ausschlaggebend seien für den Erfolg (so schon McDonnell/Elmore 1987), die andere sieht die Qualität und die Wirksamkeit (impact) eines Programms abhängig von der lokalen Gestaltung der Programme sowie den Verbindungen zwischen Mikro- und Makroebenen der Implementation. Die zentrale Bedeutung der „local characteristics“ für den Wandel der Bildungssysteme hatte Michael Fullan schon 1982 hervorgehoben (Fullan/Stiegelbauer 1991, S. 68ff.).

Zwischen beiden Strategien besteht keine Unvereinbarkeit. Neue und wirksame Instrumente sind für die Implementation bildungspolitischer Programme genauso wichtig wie die überzeugende Anbindung an die lokalen Verhältnisse.

Zwischen den Ebenen der Implementation muss es Verknüpfungen geben, die den Prozess unterstützen und ihm nicht entgegenarbeiten. Die Verknüpfungen müssen sich einspielen, sie lassen sich nicht einfach verordnen. Andererseits hängen der Erfolg und das Ansehen eines Projekts auch davon ab, inwiefern die eingesetzten Instrumente sich als tauglich herausstellen. Wenn die Rückmeldesysteme nicht funktionieren, ist schnell das gesamte Projekt der Bildungsstandards gefährdet.

Mit dieser Optik der lokalen Verknüpfung entstanden verschiedene Studien, die die politischen Programme, die mit Bildungsreformen und Implementationsstrategien in einzelnen amerikanischen Bundesstaaten umgesetzt werden sollten, untersucht haben (Odden/Marsh 1989 und diverse andere; vgl. die Beiträge in Odden 1991). Nach den Ergebnissen dieser Regionalstudien sind Bildungsreformen abhängig von den Entwicklungen vor Ort. Sie sind dann erfolgreich, wenn ihr Boden bereitet ist. In diesem Sinne sind sie „reinforced activities in which local districts already were engaged“ (Odden 1991, S. 10). Eher erfolglos sind „redistributive programs“, die in irgendeiner Hinsicht eine Neuverteilung anstreben. „Thus, implementation should be relatively swift, relatively uncontentious, and strongly linked to the extant local priorities“ (ebd.).

Das gilt auch für den schwierigen und anhaltenden Prozess der Veränderung des Curriculums, der ohne „professional expertise of local teachers and administrators“ nicht durchgeführt werden kann (ebd., S. 11). Netzwerke von reformorientierten Lehrkräften sind hier gegenüber Top-Down-Verfahren die aussichtsreichere Implementationsstrategie. Allerdings gilt das nur dann, wenn Akzeptanz vorhanden ist und sich in der Lehrerschaft oder ihren Verbänden kein grundsätzlicher Widerstand regt (ebd.). Ist das der Fall, wird die Implementation blockiert oder subversiv behandelt. Rein als politische Vision wird wenig bewirkt, die Projektziele müssen Verständnis finden, Interesse wecken und in der Lehrerschaft verankert sein (ebd.). Zudem ist die angemessene Instrumentierung der Reformziele eine unabdingbare Bedingung. Ohne geeignete Tools, die zur Schule passen, wird wenig bewegt.

Aus diesem Grunde sind Probleme der *Implementation* zugleich Fragen einer letztlich die Unterrichtsebene erreichenden pädagogischen *Qualitätsentwicklung*. Die Qualität von Schule bemisst sich letztlich danach, was die Schülerinnen und Schüler tatsächlich lernen, also welche Kompetenzen in welchen Domänen sie aufbauen können. Die bloße Rhetorik von Zielen gibt darüber keine Auskunft. Wieweit es gelingt, Schule für möglichst alle Schülerinnen und Schüler zu einem produktiven Lern- und Entwicklungsraum werden zu lassen, hängt von mehreren voneinander abhängigen Steuerungsebenen und Segmenten des Bildungssystems ab – letztendlich davon, inwieweit die Reform zu erneuerten

Strukturen der Schule und ihrer Unterstützungssysteme führt und inwieweit sie dazu beiträgt, die Kompetenzen, Einstellungen, Identitäten und Praktiken ihrer Akteure zu verändern. Dabei entscheidet nicht eine Ebene oder ein Segment allein, sondern deren Zusammenspiel, das mehr oder weniger produktiv sein kann.

Die Bedingungen und Handlungsmöglichkeiten der Produktivität von Bildungssystemen lassen sich im Sinne eines *Angebots-Nutzungs-Modells* (Fend 2000) verstehen. Danach wird Schulleistung von einem als „kollektiver Akteur“ verstandenen System sowie den heranwachsenden Kindern und Jugendlichen durch gemeinsames Gestaltungshandeln ko-produziert oder ko-konstruiert. Qualität kommt immer nur durch eine optimale „Synchronisierung von Angebotsmerkmalen und Nutzungsmöglichkeiten“ zustande, wobei Stützsysteme vorausgesetzt werden müssen, die auch außerhalb des Unterrichts angesiedelt sind (ebd., S. 57).

Möglichkeiten zur Normierung der Bildungsqualität gibt es dabei grundsätzlich auf allen Ebenen des Systems sowie auf der Nutzungsseite schulischer Bildung:

- Auf der *bildungspolitischen Strukturebene*, die keine einheitliche nationale sein muss, sondern sich in föderale und regionale Ebenen, in denen ein eigenes „Bildungsmanagement“ stattfindet, aufgliedern kann, lässt sich die Qualität des Systems, wenn auch eher schwach, durch die Schulaufsicht, die Ermittlung von Übertrittsquoten oder die Ergebnisse von Abschlüssen überprüfen. Neu hinzukommen sollen bei der „neuen Rechenschaftslegung“ objektifizierbare Outputmerkmale in Gestalt von regelmäßig überprüften Leistungen einer Schule bzw. empirisch gemessenen Lernergebnissen der Schülerinnen und Schüler auf der Basis eines nationalen Konsenses über die „Kompetenzen und Standards, die ein Bildungswesen in seinen verschiedenen Teilsektoren bei Kindern und Jugendlichen erreichen soll“ (Fend 2005, S. 15).
- Auf der Ebene der *Schule als pädagogische Handlungseinheit* gibt es mittlerweile eine Vielzahl von Instrumenten, die darauf abzielen, die pädagogische Qualität einzelner Schulen zu erfassen (Verfahren der externen und internen Evaluation) und zu stärken (Verfahren der Schulentwicklung). Die Berücksichtigung der Einzelschule als intermediäre Ebene der Qualitätssicherung und -entwicklung von Bildungssystemen hat ihren Ursprung in der internationalen Schuleffektivitätsforschung, die gezeigt hat, dass Schulen mit einer optimalen Konfiguration von Dimensionen wie Kooperations- und Führungskultur, Lehrereinstellungen und -werten, Schulklima, Curriculum und

Didaktik die Qualität des Lernens und seiner Ergebnisse nachhaltiger beeinflussen als es strukturelle Systemmerkmale und administrative Vorgaben vermögen.

- Auf der *Mikroebene des Lehrerhandelns und des Unterrichts* sind es die Lehrkräfte, von deren pädagogischen, fachlichen und didaktischen Fähigkeiten die Qualität des Angebots an Lerngelegenheiten für die Schülerinnen und Schüler abhängt. Auch wenn wesentliche Vorentscheidungen durch das Vorhandensein von Lehrplänen und Lehrmitteln, die Selektion von Schülerinnen und Schülern in Schulformen und Jahrgangsklassen sowie durch Stundentafeln und die Schulausstattung bereits zu Beginn der Beschulung gefallen sind, sind es die Lehrpersonen, denen vielfältige Normierungen durch die Gestaltung produktiver Lehr-Lernumgebungen obliegen.

Soweit lässt sich die durch die Angebotsseite bestimmte Qualität des Systems fassen. Schwieriger ist die optimale Abstimmung dieses Angebots im Hinblick auf die in der Regel heterogenen Nutzungsfähigkeiten und -voraussetzungen der Schülerinnen und Schüler. Auch eine professionelle Gestaltung des Angebots auf allen Systemebenen garantiert noch nicht, dass diese Angebote auch *effektiv genutzt* werden und Lerngelegenheiten sich in Lernerträgen von Schülerinnen und Schülern niederschlagen.

Die Nutzung des Angebotes kann angeleitet, aber nicht aufgezwungen werden. Daher steht es nur bedingt unter der Kontrolle von Schulen und Lehrkräften, wie bereitwillig, intensiv und motiviert ihre Bildungsangebote aufgenommen und verarbeitet werden. Das heißt, *wie ertragreich die Angebote der Schule sind, hängt von deren Nutzung durch die Schülerinnen und Schüler ab*. Auf den Aufbau der kognitiven, motivationalen und sozio-emotionalen Nutzungsqualitäten von Bildungsangeboten hat die Schule jedoch nur begrenzten Einfluss. Diese Qualitäten werden zu einem guten Teil, wenn nicht überwiegend, in familialen Milieus, Peer-Groups sowie weiteren außerschulischen Stützsystemen entwickelt und stabilisiert.

Damit sich Schulen verantwortlich fühlen können für Lernergebnisse, die ihnen nach Durchführung von nationalen oder regionalen Tests und als Ergebnis von der Evaluationen rückgemeldet werden, müssen sie über entsprechende Rahmenbedingungen verfügen, innerhalb derer sie ihre Bildungsangebote im Rahmen ihres professionellen Selbstverständnisses und unter Würdigung der eigenen situationalen Voraussetzungen selber gestalten können. Das heißt, die Philosophie einer „neuen Rechenschaftslegung“ verlangt die Koppelung von verbindlichen, orientierenden Vorgaben mit (neuen) *Spielräumen von Schulautonomie* und daran geknüpften *Ressourcen und Werkzeugen der lokalen und regionalen*

Schulentwicklung. Jegliche Form von Evaluation ergibt nur Sinn, wenn im Rahmen geöffneter Handlungsmöglichkeiten auch produktiv auf die Ergebnisse reagiert werden kann.

Auf der Ebene der nationalen und regionalen Politik bedeutet dies eine *Reform des Verwaltungshandelns*, verbunden mit der Allokation von Ressourcen. Wenn Schulen für die Ergebnisse ihres Handelns verantwortlich gemacht werden sollen, so müssen sie in wesentlich größerem Umfang als bisher auch Entscheidungen, die vormals administrativ gefällt wurden, zum Beispiel in Bezug auf die Gestaltung der Schulverfassung, die Ausgestaltung von Schulleitung, die Gestaltung von individuellen Förderangeboten, selber treffen können. Auf der Ebene der pädagogischen Arbeit der Einzelschulen bedeutet dies die Etablierung von Modellen einer permanenten Sicherung und Weiterentwicklung der Qualität des eigenen professionellen Handelns.

2.6 Kritik an Standards und Tests

Da Bildungsstandards in mancherlei Hinsicht einem Paradigmenwechsel bezüglich Qualitätspolitik und Qualitätssicherung nahe kommen und gleichzeitig das Selbstverständnis von Lehrpersonen als Ergebnisverantwortliche schulischer Bildungsprozesse tangieren, ist es nicht verwunderlich, dass der bildungspolitische Diskurs über Standards und deren Implementation auch auf zum Teil heftige Kritik gestoßen ist – auch im deutschen Sprachraum (vgl. zusammenfassend Uhl 2006). Ein Teil der Kritik ist technischer Natur und bezieht sich auf die Qualität der bisher ausgearbeiteten Standards und der darauf bezogenen Leistungstests, soweit solche vorliegen. Andere Kritiken betreffen den Kern des Steuerungsmodells, also die Strategie der Steuerung vom Output her, deren Vorteile bestritten und als illusorisch angesehen werden. Das betrifft auch den Zusammenhang von Tests und Rückmeldesystemen.

Vorwiegend im deutschen Sprachraum wird der Begriff „Standards“ im Zusammenhang mit Bildung zumeist als eine Zumutung für pädagogisches Denken empfunden (Klieme 2005b). Man begegnet dem Begriff mit einer besonderen Skepsis, insinuiert er für viele doch das Gegenteil dessen, was gemeinhin mit pädagogischem Denken verbunden wird: die förderliche Pflege von Eigensinn und Autonomie sowie das Lernen auf eigenen Wegen unter der Leitidee der Entfaltung der Persönlichkeit. Demgegenüber lässt der Begriff „Standards“ die Assoziation aufkommen, dass Schülerinnen und Schüler und ihr Lernen künftig produktemäßig „standardisiert“, deren Eigenes eingeebnet und deren Lernen in ein vorgegebenes Korsett gezwängt werden sollen (Spinner 2005).

Allgemein wird sodann befürchtet, dass der Diskurs über Schulqualität künftig nur mehr in Kategorien von Bildungsstandards geführt werde (Herzog 2006), was zu einer Verengung des Qualitätsdenkens und zu einer Gefährdung pädagogisch notwendiger Spielräume im Handeln von Schulen und Lehrpersonen führen werde. Die Schule bestehe nicht nur aus Output und Bildung lasse sich nicht bloß „quantitativ“ betrachten und aus einer Kosten-Nutzen-Optik bewerten (ebd.; Herrmann 2003). „Technokratische“, von ökonomischen Denkmustern und einem „reduktionistischen Bildungsbegriff“ geprägte Gängelung und ausufernde Steuerungsphantasien würden im Kontext eines „erpressten Reformismus“ zu keiner Hebung von Bildungsqualität beitragen (Frankfurter Erklärung vom 10. Oktober 2005). Eher sei mit einer „Verödung des Lernens“ im Sinne eines unweigerlichen „Verlustes an Sensibilität für individuelle, situative und inhaltspezifische Weltaneignungen in ihrem Eigensinn“ zu rechnen (Rumpf 2005, S. 6).

Im Anschluss an Luhmann/Schorr (1979) und unter Bezugnahme auf Literatur zur Schulentwicklung wird zudem ins Feld geführt, dass pädagogische Prozesse grundsätzlich nicht technologisierbar seien, Bildung sich nicht kalibrieren lasse und Output-Prüfungen sowohl in Bezug auf Einzelschüler als auch in Bezug auf ganze Schulsysteme deshalb nicht rational bewerkstelligt werden könnten (vgl. z.B. Herzog 2006). Herzog hält zwar, wie viele der Skeptiker, fest, „nicht gegen Bildungsstandards Partei ergreifen“ zu wollen, kritisiert diese unter der Optik „Was droht, vergessen zu gehen?“ dennoch radikal: „Schulen bilden komplexe Systeme, die einer eigenen Logik folgen, die von außen zwar irritierbar, nicht aber beherrschbar sind.“ Ausgehend von der Behauptung, Bildungsstandards würde (zwingend?) ein Wirkungsmodell von Bildungsqualität zugrunde liegen, das einem industriellen Fertigungsprozess entspreche, hält er „die Vorstellung, mittels Standards ließen sich Bildungsprozesse auf ein Niveau der Verlässlichkeit anheben“ wie wir es eben von solchen „Fertigungsprozessen gewohnt sind (für) absurd“ (S. 9f.).

Was die Ebene des Lernens angeht wird befürchtet, Bildungsstandards würden – vor allem wenn sie auch als Selektionsinstrumente verwendet würden – zu einem für Schülerinnen und Schüler erhöhten Leistungsdruck und zu einem als negativ zu qualifizierenden „teaching to the test“ führen. Dadurch würde nicht zuletzt das politische Postulat der Chancengerechtigkeit unterminiert, anstatt, dass diese durch Qualitätsentwicklungsmaßnahmen gestärkt würde (Böttcher 2003; Ratke 2006) – ein Argument, das auch in der US-amerikanischen Diskussion um den „No Child Left Behind Act“ (NCLB; dort im Zusammenhang mit einem im deutschsprachigen Raum nicht vorgesehenen „High stakes

Testing“ mit unmittelbaren Konsequenzen für die Schulen) zu hören ist (vgl. Linn/Baker 2002; Meier/Wood 2004).

Die Mehrzahl der Kritiken an Bildungsstandards ist jedoch nur teilweise pädagogischer, sondern vielmehr und vor allem *politischer* Art, wobei sich vor allem in der angelsächsischen Pädagogik unversöhnliche Lager gegenüberstehen, die gegensätzliche Philosophien verfolgen. Dabei ist die eben kurz dargestellte, vor allem im deutschen Sprachraum vorgebrachte Kritik nicht annähernd so entwickelt wie etwa in den Vereinigten Staaten. Bestimmte Grundzüge der Kritik lassen sich aber quer zu den nationalen Systemen ausmachen. Auf sie gehen wir im Folgenden näher ein.

Zuerst einmal sind Standards zweifelsohne keine Wundermittel, und schon gar keine (industrielle Fertigungs-)Technologie, welche Bildungserfolg „herstellbar“ macht. Zwischen der politischen Rhetorik und den tatsächlichen Möglichkeiten der Leistungsmessung ist schon früh ein Missverhältnis konstatiert worden (Shavelson/Baxter/Pine 1992). Aber nicht nur hier, im externen, immer wieder auch beargwöhnten, sicher nicht zeitlos unproblematischen Verhältnis von Forschung und Bildungspolitik, liegt ein Problem. Diane Ravitch etwa hat auch interne Konfliktlinien beschrieben, die sich weniger aus der Idee der Bildungsstandards oder der Unterscheidung von drei Typen ergeben, sondern die vielmehr mit der Art und Weise ihrer *Implementation* zu tun haben:

„What is most controversial about these three types of standards is how they will be enforced and by whom. Will such standards be mandatory, voluntary, or de facto?“ (Ravitch 1995, S. 14).

Die Kritik an der amerikanischen Standards-Bewegung hat behauptet, dass es sich dabei bloß um rhetorische Fragen handle, die das eigentliche Problem verbergen. Bildungsstandards seien lediglich der *Slogan* der Bewegung, mit dem verdeckt werde, was die Praxis ausmache, nämlich zunehmende Ungleichheit, finanzielle Unterversorgung und wachsende Armut in vielen Schulbezirken. Die Rhetorik von Slogans überlagere die Realität in den Schulen, weil anderes – die Implementation von Bildungsstandards – vordringlich erscheine.

Michael Apple hat diese These 1992 in seiner Analyse der curricularen Vorschläge des *National Council of Teachers of Mathematics* entwickelt. Er teilt drei der allgemeinen Ziele beider Expertisen, nämlich erstens ein nicht-elitäres Curriculum in Mathematik zu entwickeln, das nicht von Begabungen (talents) ausgeht, sondern für alle Schülerinnen und Schüler geeignet ist. Zweitens wird der Ansatz positiv hervorgehoben, mathematisches Lernen von „real-world problems“ ausgehen zu lassen. Und drittens wird die Bindung der Leistungsmessung

(assessment) an das Curriculum und die örtlichen Ressourcen begrüßt. Nur so ist Fairness gegenüber den ganz unterschiedlichen Lernvoraussetzungen möglich (Apple 1992, S. 412f.). Eine andere Frage ist, wie die rhetorische Funktion der Standards zu verstehen ist.

Diese Frage untersucht Apple mit einem Analyseinstrument, das die beiden Philosophen Paul Komisar und James McClellan 1961 entwickelt haben. Das Instrument beschreibt die Rolle von Slogans in Erziehungsdiskursen. Slogans dienen dazu, für bestimmte Bewegungen Anhänger zu rekrutieren. Sie müssen attraktiv formuliert werden und den Kern der Bewegung in einem treffenden Ausdruck zusammenfassen, wie etwa die „Pädagogik vom Kinde aus“ des frühen 20. Jahrhunderts, die gleichermaßen Programm wie Schlachtruf einer sehr heterogenen „Bewegung“ war. In einer ganz kurzen Botschaft muss alles gesagt sein, was zur Formung einer Überzeugung und so zur Anhängerschaft nötig ist. Apple überträgt diese Erfahrung der Reformpädagogik auf die sich formierende Standards-Bewegung, die erneut sehr verschiedene Gruppen und Interessen umfasst und doch wie eine Einheit erscheinen kann, die sich um einen Slogan schart.

Slogans können nur dann erfolgreich sein, wenn sie bestimmte Bedingungen erfüllen. Zum einen müssen Slogans Innovationen versprechen, dabei aber hinreichend vage sein, so dass verschiedene Gruppen sich hinter das Konzept stellen können, die ansonsten nicht zueinander finden würden. Gleichzeitig darf die Vagheit nicht übertrieben werden. Das Slogan-System funktioniert nur, wenn es hinreichend spezifiziert ist, um auch für Praktiker interessant zu sein. Drittens muss das System attraktiv sein, „Charme entfalten“, wie Apple sagt. Der Stil muss ad hoc überzeugen und die Botschaft muss zum Handeln auffordern.

„It offers us a sense of imaginative possibility and in doing so generates a call, and a claim for, action“ (Apple 1992, S. 413f.).

Im Sinne dieser Kriterien waren Semantik und politische Rhetorik von Standards erfolgreich. Aber der Erfolg verdeckt laut Apple die Probleme. Gegenüber der Art und Weise, wie die Gesellschaft Ungleichheit produziert, sind die Bildungsstandards „relativ naiv“ (ebd., S. 418). Nachteile durch soziale Herkunft, Rasse oder Geschlecht werden wohl berücksichtigt, aber das kann die vorherrschenden Einstellungen der Gesellschaft, etwa zur „natürlichen“ Begabung, kaum sehr weit korrigieren (ebd., S. 419). Und die Zentralität der Mathematik als „high status knowledge“ mit entsprechender Selektivität ist damit auch nicht außer Kraft gesetzt (ebd., S. 420ff.). Die Frage ist, „whose know-

ledge and whose problems“ gelernt werden sollen (ebd., S. 422), eine Frage, die sich nicht mit Tests beantworten lässt. Alles in allem genommen macht die Reform enorme Versprechungen, aber die Gefahr ist groß, dass sie an der Gesellschaft scheitert, wie schon viele Erziehungsreformen zuvor (ebd., S. 428; vgl. Sarason 1990 und diverse andere).

Thomas Green (1980) hat darauf verwiesen, dass Bildungssysteme wie das amerikanische immer zwei Ziele zugleich verfolgen, die unvereinbar sind. Sie sollen für Chancengleichheit sorgen und zugleich die individuellen Bedürfnisse der Lernenden ernst nehmen. Im ersten Fall werden Standards gesetzt, die für alle gelten und die mindestens in einem hohen Minimum auch von allen erreicht werden sollten. Im zweiten Fall werden Standards möglichst vermieden, damit nach eigenem Bedürfnis gelernt werden kann. Das erste Ziel gilt aus der Sicht des zweiten als Bürokratisierung, das zweite erscheint aus der Sicht des ersten als undemokratisch. Alle großen Ideologien im Bildungssystem erwachsen aus diesem Gegensatz, der bislang unaufgelöst ist. Kritiken auf dieser Linie an den Standards sind hundertfach nachzuweisen.

Chancengleichheit kann als Paradox der öffentlichen Schule bezeichnet werden (so schon Komisar 1966), das macht egalitäre Argumente um nichts unattraktiver. Auf der anderen Seite können Ansätze des Lernens nach Bedürfnis als schulfremd hingestellt werden, weil die Schule einem mehr oder weniger verbindlichen Curriculum folgt, ohne dass dadurch das Argument der Kindzentrierung an Einfluss verlöre. Apple gehen die egalitären Standards nicht weit genug, für andere Kritiker sind sie an sich schon eine Zumutung. Von dieser vorhersagbaren Struktur der bildungspolitischen *Diskussion*, die Green vorschnell mit dem *Verhalten* des Systems in Verbindung bringt, sind Entwicklungen der Praxis zu unterscheiden. Slogans sagen zwar etwas über einen wie auch immer brüchigen Grundkonsens innerhalb einer Bewegung aus, jedoch nichts über reale Implementationen, die unterhalb der ideologischen Schwelle vom System akzeptiert werden.

Eine weitere Kritik an den Standards betrifft diese Frage: Staatliche Bildungspolitik, das zeigt das englische Beispiel, reglementiert die Praxis mit Gesetzen und so mit Macht. Kontrollmechanismen wie Testsysteme werden aufgebaut und haben dann Folgen, ohne dass genau klar wäre, was gemessen wird oder was mit den Daten angefangen werden soll. Es kann so etwas wie eine *Audit-Society* (Power 1997) entstehen, eine Gesellschaft, die auf allen Ebenen ständig Rechenschaft (audit) abverlangt, ohne davon besonders viel Nutzen zu haben. „Audits work because organisations have literally been made auditable“ (ebd., S. 91). Rechenschaft verkommt zum Ritual, einfach weil die Umwelten

der Organisationen darauf eingerichtet sind und formal Bilanzierungen abverlangen, die intern leer laufen (ebd.).

Es ist fraglich, so die Kritik, ob das grundsätzliche Versprechen, die Dinge „besser“ oder „effizienter“ zu machen, mit Hilfe von Standards und Evaluationen eingelöst werden kann. Die *ökonomische Rhetorik* löst den Verdacht aus, es handle sich dabei lediglich um „a messianic belief“ (Wolf 2000, S. 29), mit dem letztlich nur Unheil angerichtet werde. Diese Kritik wird häufig vertreten, einhergehend mit dem Verdacht, die Einführung von Bildungsstandards führe zu einem verengten Curriculum, das von der Wirtschaft angetrieben werde (economy driven) und zur Uniformierung des Unterrichts führe. Dadurch sei die pädagogische Freiheit gefährdet und werde der Bürokratisierung Vorschub geleistet (Miller 1995 und diverse andere; vgl. die „Frankfurter Erklärung“).

Aber das deckt sich nicht mit generalisierbaren Erfahrungen und stellt keine Erfassung der komplexen Wirklichkeit dar, die z.B. auch in England keineswegs nur negativ ist. Eine Kritik am Weißbuch der englischen Regierung von 1997 (Department for Education and Employment (DfEE) 1997), das eine positive Bilanz der Implementation von Bildungsstandards zieht,¹² geht denn auch eher dahin, die technischen Probleme des Testens ernst zu nehmen, besonders des Testens über längere Zeiträume, und dabei die Relativität der Aussagen vor Augen zu haben. „There is no absolute criterion which determines what pupils will be able to do“ (Plewis/Goldstein 1998, S. 19). Das Lernen reagiert nicht einfach auf gesetzte „benchmarks“ (ebd.), und es dürfte schwer sein, die Ziele ohne deren „shifting“ über die Zeit konstant zu halten (ebd.).

Eine Richtung in der amerikanischen Kritik verweist auf die Gefahr, bei hohen Investitionen in die Beschreibung der Ergebnisse die Inhalte zu vernachlässigen und die Prozesse des Lernens einzig auf das Ergebnis zuzuschneiden. Beachtet werden dann nur noch Testresultate, einschließlich der dazu nötigen Motivierungstechniken. Damit werden kurzfristige Erfolge gesucht und der langfristige Aufbau der Kompetenz in schulischen Lernbereichen (domains) vernachlässigt. Je mehr man auf diese Weise testet, desto weniger stabil sind die Resultate. In Metaanalysen zeigt sich allerdings, dass diese Befürchtungen übertrieben sind, oft mit schlechten Tests zu tun haben und gelegentlich auch nur Schutzbehauptungen darstellen. Mindestens übersehen sie die positiven Möglichkeiten der Überprüfung des Unterrichts mit Hilfe von Evaluationen und Rückmeldesystemen (Evers/Walberg 2004).

¹² Daten zur Qualitätssteigerung finden sich auch in: Arbeitsgruppe Internationale Vergleichsstudie (2003, S. 167ff.).

Der Sinn des Testens ist die fortlaufende Kontrolle der Resultate. Damit verbunden ist folgende Erwartungshaltung:

- (1) You know what you want;
- (2) you measure it well;
- (3) you make reasonable choices based on the interpretations of results;
- (4) everyone knows the consequences of the system;
- (5) management accounts for key system variables; and
- (6) the outcomes improve (Baker 2004b, S. 158).

Wer das Erreichen von Standards messen will, muss, was immer sie sind, Standards präzise und akkurat formulieren, sie mit den richtigen Mitteln versehen und auf verschiedene Typen von Lernern ausrichten. Die bisherigen Erfahrungen zeigen, dass viele Lehrmittel und nicht wenige Tests entweder auf die besten oder die durchschnittlichen Schülerinnen und Schüler ausgerichtet sind. Die schwierigste Gruppe, die leistungsschwachen Schülerinnen und Schüler, werden vielfach nicht erfasst. Schwache Schülerinnen und Schüler wechseln in den Vereinigten Staaten oft die Schule, haben begrenzte Kenntnisse der Unterrichtssprache, erhalten vielfach schlechtere Lehrer und werden mit geringeren Ressourcen unterrichtet. Wer das Schulsystem ernsthaft von den Resultaten aus entwickeln will und zugleich Förderabsichten vertritt, muss dieses Problem lösen können (ebd., S. 159). Die Lehrkräfte und die Schülerinnen und Schüler richten sich kognitiv und vor allem motivational darauf ein, dass und wie getestet wird (O'Neil/Abedi/Miyoshi et al. 2004).¹³ Tests sind nicht einfach nur Leistungsmessungen, sondern selbst Lernobjekt, das von den Schülerinnen und Schülern als für ihr Lernen mitverantwortliche Lernsubjekte mehr oder weniger ernst genommen wird. Sie lernen, sich regelmäßig auf Tests einzustellen, so wie sie gelernt haben, regelmäßig Klassenarbeiten zu schreiben und am Ende eine Note entgegenzunehmen. Ohne klare Zweckbestimmung und ohne effektive Rückmeldesysteme besteht die Gefahr einer Ritualisierung, die kaum die erwünschten Auswirkungen auf das Lernen im Unterricht hat.

Erfahrungen vor allem aus den USA zeigen, dass der Zusammenhang zwischen Standardisierung und Testing nicht unproblematisch ist. Rigide Testprogramme können die soziale Ungleichheit verstärken und die ohnehin gegebene Benachteiligung bestimmter Gruppen erhöhen. Gute Testprogramme sind dagegen aufwändig in der Entwicklung, die Rückmeldung der Resultate an die

¹³ Vermutet wird ein „senioritis-effect“, also eine Abnahme der Motivation am Ende der Schulzeit (O'Neil/Abedi/Miyoshi et al. 2004, S. 199).

Schulen ist ein andauerndes Problem (Orfield/Kornhaber 2001, S. 149f.) und der Aufwand ergibt nur Sinn, wenn die Ergebnisse bei politischen Entscheidungsprozessen auch genutzt werden. Datenbezogene Bildungspolitik ist aber immer noch die Ausnahme (ebd., S. 155). Letztlich stellt also diese Art von Qualitätssicherung selbst ein Qualitätsproblem dar.¹⁴

Wenn regelmäßige Tests die Entwicklung steuern sollen, dann müssen sie mit klaren inhaltlichen Anforderungen verknüpft sein, kognitive Ansprüche stellen und sich auf Leistungen beziehen. Das verlangt, dass sie für den tatsächlich geleisteten Unterricht sensibel sind, technisch eine hohe Qualität aufweisen, auf klare Ziele/Standards ausgerichtet sind und; dass Evidenz besteht, dass sie sich für verschiedene Zwecke und Subgruppen von Schülerinnen und Schülern eignen (Baker 2004b, S. 162). Ohne Berücksichtigung des sozialen und kulturellen Hintergrundes der Schülerinnen und Schüler geht die Testkonstruktion an den Realitäten vorbei (so schon Porter 1993). Nicht selten werden Tests verwendet, die gleich hohe Maßstäbe für alle anlegen, statt hohe, aber differenzierte Standards zu testen (Linn 2000, 2003). Die Prämisse in der Anlage vieler Tests ist nicht, dass es individuelle Unterschiede im Entwicklungsstand der Schülerinnen und Schüler gibt, sondern, dass sie alle das gleiche Ziel hätten erreichen können, was nicht der Fall ist und sein kann (Lam 2004, S. 141). Daher ist *Fairness* des Vergleichs eine zentrale Anforderung an alle Leistungstests (Nachtigall/Kröhne 2006).

Eine weitere Variante der Kritik betrifft die politische Rhetorik hinter den Bildungsstandards (Baker 2004b, S. 153ff.). Wenn in der amerikanischen Administration behauptet wird, es sei möglich, die öffentliche Bildung so auszurichten und zu reformieren, dass „kein Kind zurückbleibt“, so ist damit zunächst nicht mehr erfasst als die Kommunikation eines politischen Slogans. In der Praxis erreichen nie alle Kinder gleiche Ziele, schon weil die Schule soziale Nachteile nur begrenzt ausgleichen kann (Meier/Wood 2004). Slogans sind nie Ziele, aber sie werden vielfach genauso verstanden, als Eröffnung von Möglichkeiten, die oft praktisch ausgeschlossen sind. Andererseits ist es nur mit Hilfe von Standards möglich, die Linie zu bestimmen, die auch schwächere Schülerinnen und Schüler erreichen müssen. Förderprogramme setzen, wenn sie wirksam werden sollen, Standards und darauf bezogene Ressourcen voraus. Grundsätzlich wird in den meisten Beiträgen nicht die Erweiterung der Perspektive hin zu den Resultaten von Schule und Unterricht in Frage gestellt, sondern eher die Rhetorik der Implementation oder die technischen wie die po-

¹⁴ Siehe auch Sacks (1999) oder McNeil (2000).

litischen Probleme des Testens. Aus der Kritik folgt, die Entwicklung sehr genau zu beobachten und tatsächlich von den Ergebnissen auszugehen, in deren Licht sich bestimmte Zielsetzungen beibehalten oder auch korrigieren lassen. Die Reform muss als anhaltender Prozess des Systemumbaus betrachtet werden, der sich nicht linear vollzieht und in dessen Verlauf eine hohe Fehler-sensibilität entwickelt werden muss. Kritik ist dabei unverzichtbar. Zudem sollte klar sein, dass bestimmte Probleme des jeweiligen Bildungssystems auch nach der Implementation von Bildungsstandards bestehen bleiben. Die Politik sollte schließlich sehr genau bestimmen, was mit den Bildungsstandards erreicht werden soll und wie sich die Erfahrungen verwenden lassen.

Auf der anderen Seite steht eine Kritik, die von den politischen Problemen der Implementation ausgeht und beklagt, dass und wie die Standards unterlaufen werden. Basis dieser Kritik ist ein Vergleich der tatsächlichen Standardentwicklung in den amerikanischen Bundesstaaten (Finn/Petrelli 2000). Die Probleme der Implementation werden so gefasst: „It’s difficult to reach agreement on standards, hard to get the assessments properly aligned with those standards, and painful – mainly for political reasons – to impose meaningful consequences on students, teachers, and schools“ (Finn 2002, S. 32). Einige Staaten wie Texas haben sich diesen Herausforderungen gestellt, viele jedoch nicht. Standards, so der Befund, haben nur dann einen positiven Effekt, wenn Konsens besteht, die Formen der Überprüfung klar sind und Konsequenzen gezogen werden können (ebd.). Das Unterlaufen der Standards ist auf vielfältige und oft kaum sichtbare Weise möglich. Standards können stark oder schwach sein, leicht oder anforderungsreich, was oft aus den Formulierungen gar nicht hervorgeht. Die Qualität und Strenge der Tests stimmt nicht mit den publizierten Standards überein. Benchmarks können gesetzt werden. Auch wenn der Test aus qualitativ guten Fragen besteht, können viele Schülerinnen und Schüler den Test bestehen, obwohl sie viele Fragen *nicht* beantwortet haben. Wenn die Bestehensrate ein Indikator für den Erfolg oder die Qualität einer Schule ist, wird die Rate steigen, sofern die Länder oder die Distrikte Einfluss nehmen können.

„With so many ways to let standards slip, it is little wonder that we find many states reporting far larger fractions of their students (or schools) being deemed proficient on the states’ own measures than are judged proficient according to the standards set by the National Assessment Governing Board (NAGB) for reporting results on the National Assessment of Educational Progress (NAEP). And even when a state holds firm in its standards, it may falter in the politically ticklish task of attaching rewards and sanctions to student, school, and teacher performance“ (ebd., S. 32/33).

3 Stand der Forschung und Entwicklung in vier ausgewählten Ländern

Im dritten Teil der Expertise wird ein Überblick über Erfahrungen mit der Entwicklung und Implementation von Bildungsstandards in vier ausgewählten Ländern gegeben. In diesen Ländern – Vereinigte Staaten von Amerika, Niederlande, Schweden, England – ist vor Jahren und Jahrzehnten damit begonnen worden, das nationale Bildungssystem stärker in Richtung einer Output-Orientierung umzusteuern. Hier liegen also Erfahrungen und Forschungsergebnisse vor, die bei der Abschätzung hilfreich sein können, wie sich der „Paradigmenwechsel“ in den Auftragsländern vollziehen lässt. Dabei kann es nicht darum gehen, direkte Übertragungen vorzunehmen. Aber gewisse Problemlagen kommen in allen Bildungssystemen vor, wenn sie auf eine Steuerung stärker vom Resultat her eingestellt werden. Bestimmte Maßnahmen sind erfolgversprechender als andere, wobei man von den einschlägigen Erfahrungen auch dann lernen kann, wenn die Systemkonfigurationen unterschiedlich sind. Das entspricht der historischen Erfahrung punktueller Adaptionen und der Anpassung fremder Modelle.

Die Wahl der Länder begründet sich einerseits mit ihren Erfahrungen, andererseits aber auch aufgrund ihrer unterschiedlichen Voraussetzungen. Organisation und Strukturen des Bildungswesens sind verschieden, der soziokulturelle Kontext ist grob vergleichbar. Gemeinsam ist den Ländern, dass sie im Bildungsbereich zum Teil massive Reformanstrengungen unternommen und dabei weitreichende Innovationen umgesetzt haben. Materiell weichen diese Reformen aber stark voneinander ab. Das gilt schon für die Definition von „Bildungsstandards“, bezieht sich auf den Umgang damit und kommt auch hinsichtlich des Aufbaus von Evaluationsstrukturen zum Tragen.

Vor dem Hintergrund der Differenzen in den Voraussetzungen, und dies bezüglich der aktuellen Systemkonfiguration wie hinsichtlich normativer historisch-kultureller Traditionen und Kontexte, wurden verschiedene System- und Reformkomponenten bei den Ausführungen zu den einzelnen Ländern zum Teil unterschiedlich gewichtet. Tatsächlich lassen sich am Beispiel der vier Länder mithin je unterschiedliche Problemlagen, Reaktionsweisen und Korrekturmaßnahmen innerhalb der Reformen exemplarisch betrachten. Man kann auch sagen: Anhand der Beispiele USA, England, Schweden und Niederlande lässt sich je Unterschiedliches lernen. Zudem übertrifft die Forschungslage zu Prozessen der Implementation von standardbasierten Bildungsreformen in den USA die europäischen Länder bei weitem; dieser Vorsprung betrifft sowohl Erfahrungen und wissenschaftliche Ergebnisse aus Begleitstudien und Evaluationen konkreter Implementationsvorhaben ebenso wie die Ebene der Theoriebildung. Implementation erfolgt in diesem föderalistischen Staat unter dem Vorzeichen eines zunehmenden Engagements von Seiten des Bundes – verknüpft mit erhöhten Accountability-Anforderungen. Umgekehrt ist die Stärkung einer Output-Steuerung in Schweden und den Niederlanden begleitet von einer weitgehenden Dezentralisierung – in Schweden vor allem als Kommunalisierung und zwar im Anschluss an eine lange Tradition zentralstaatlicher, bürokratisch

verfahrender Organisation und Steuerung; entstanden ist hier ein komplexes Gefüge der Ebenen im Kontext von Maßnahmen der Deregulierung und der Orientierung an zentral vorgegebenen Zielen. Dies steht wiederum ganz im Gegensatz zu der langen Autonomie-Tradition in der Geschichte des niederländischen Bildungswesens. Entsprechend anders stellen sich dort vor dem Hintergrund der Umsteuerung auf der lokalen Ebene Probleme wie dasjenige der Akzeptanz der kommunalen Akteure beim Lehr- und Schulleitungspersonal und der Professionalisierung angesichts neuer Aufgaben und Zuständigkeiten auf beiden Seiten. Solche Fragen sowie befürchtete Veränderungen auf Kosten der verankerten Maxime der Gleichheit in Schweden scheinen aktuell dazu zu führen, dass das Pendel eher wieder zurückschlägt in Richtung zentraler Vorgaben und Kontrollen.

Lernen aus teilweise unerwünschten Effekten der Steuerungsreformen sowie der zwischenzeitlich, im Anschluss an den letzten Regierungswechsel, eingeleiteten Korrekturen lässt sich auch anhand des Beispiels England. Hier ging die Reform zuerst insoweit in umgekehrte Richtung, als die lokale bzw. regionale Ebene der Bezirke, die Local Education Authorities, gleichzeitig mit der Einführung eines nationalen Curriculums vom Staat Ende der 80er-Jahre gewissermaßen entmachtet wurde. Nach erfolgten Kurskorrekturen im letzten Jahrzehnt ist das Beispiel im Ergebnis paradigmatisch für eine im europäischen Vergleich weitgehende Deregulierung des Schulangebots bei einem gleichzeitig engmaschigen Regime zentraler Zielvorgaben und -kontrollen.

Die ausgewählten Länder und ihre Entwicklungen lassen sich folgendermaßen näher charakterisieren:

Vereinigte Staaten

Zuständig für das Schulwesen sind in den USA grundsätzlich die einzelnen Staaten; die Schulverwaltung ist dezentral in den Distrikten angesiedelt. In den letzten beiden Jahrzehnten zeichnet sich ein zunehmender Regulierungswille auf Seiten der Bundesregierung, aber auch bei den Ministerien der Bundesstaaten ab. Artikuliert wird dieser Wille im Diskurs über Standards, mit dem sich neue Politikstrategien verbinden, die in allen fünfzig Bundesstaaten, wenngleich unterschiedlich, verfolgt werden. Zusätzlich tritt der Bund als zunehmend starker Akteur auf, der mit seiner Gesetzgebung und der damit verknüpften Ressourcenzuteilung inzwischen massiven Einfluss ausübt.

Die europäische Standard-Debatte wurde stark von der amerikanischen angeregt, die bis in die siebziger Jahre zurückgeht und eigentlich noch älter ist. Entsprechend groß sind Forschungsaufkommen und Erfahrungen, die auch für

die Entwicklung im deutschen Sprachraum von Interesse sind. Bereits heute wird ständig auf amerikanische Befunde Bezug genommen. Wer etwa auf negative Konsequenzen der Einführung von Bildungsstandards, Leistungsüberprüfungen, externen Evaluationen oder High-Stakes-Testing hinweisen will, zitiert gewöhnlich amerikanische Studien. Die Breite der Erfahrungen wird dabei oft nicht erfasst, auch nicht die Spezifik der politischen Diskussion. Die Rezeption ist also nicht ganz einfach, was auch mit der Größe und Beschaffenheit des Forschungsaufkommens zu tun hat.

Die Vereinigten Staaten werden in diesem Kapitel insbesondere im Blick auf Probleme und Theorien der Implementation herangezogen. Hier hat sich seit etwa zwanzig Jahren ein eigenes Forschungsgebiet herausgebildet, das aus einer bestimmten Beobachtung hervorging. Wenn Projekte der Schulentwicklung keine verbesserten Resultate mit sich brachten und vom System ohne großen Effekt absorbiert wurden, dann konnte das häufig nicht auf eine mangelhafte Konzeption der Innovation zurückgeführt werden, sondern auf eine unvollständige Implementation. Dies betrifft etwa die Implementation groß angelegter sogenannter systemischer Standardsreformen (*Standards-Based Systemic Reform*) sowie die Implementation und Verbreitung umfassender Schulreform-Modelle (*Comprehensive School Reform Designs*), wie sie von Seiten des Bundes oder auch verschiedener Staaten propagiert und subventioniert werden. An ihnen hat sich die Implementationsforschung entwickelt.

Das Problem der suboptimalen Implementation ist auffällig, wurde früher jedoch kaum beachtet. Projekte wurden politisch diskutiert, finanziert, in Gang gesetzt, jedoch nicht mit einer aussichtsreichen Implementationsstrategie versehen. Das hat sich inzwischen geändert, weil sichtbar wurde, welche Probleme sich stellen, wenn man die große Zahl an politischen, Aufsichts- und Verwaltungsebenen (Bund, Bundesstaat, Distrikt, Einzelschule, School Board) mit je eigenen Politiken, Projekten und Interessen betrachtet und dann noch private und nationale Initiativen in Rechnung stellt, an denen sich Schulen, Distrikte und Staaten beteiligen können.

Niederlande

Dominierend im niederländischen Bildungssystem ist die in der Verfassung festgeschriebene „Freiheit der Bildung“, die dazu geführt hat, dass sehr viele, zumeist konfessionelle, Privatschulen bestehen, die die Eltern wählen können und die vom Staat alimentiert werden. Beide, staatliche wie private Schulen, waren seit Beginn des 20. Jahrhunderts immer sehr selbständig. Seit zwei Jahrzehnten wird eine Bildungspolitik verfolgt, die von einem weiteren Ausbau der

schulischen Autonomie ausgeht und Dezentralisierung befördert. In der Konsequenz kommt es zu einer Verlagerung der Verantwortung für die Qualität der Bildung auf die lokale Ebene.

Schon aus historischen Gründen stehen die Niederlande bezüglich Dezentralisierung international an der Spitze. Um Qualität und Vergleichbarkeit dennoch zu sichern, wurden für die Grundschule sogenannte Kernziele eingeführt, die jedoch sehr allgemein gehalten sind. Es sind keine Standards im angelsächsischen Verständnis, wenngleich sie eine durchaus ähnliche Funktion übernehmen. Am Ende der Sekundarstufe stehen zentrale, standardisierte Abschlussprüfungen, auf die hin gelernt werden muss. Die Kopplung mit den nationalen Lernzielen ist allerdings bislang nur recht lose.

Es gibt ein sehr umfangreiches und weit gefächertes System schullaufbahnbegleitender Tests, das die Schulen gemäß ihren Bedürfnissen einsetzen können. Obligatorisch ist lediglich die zentrale Abschlussprüfung. Neben den nationalen Kernzielen, den Abschlussprüfungen und einem gewissen Wettbewerb bildet das staatliche Inspektorat das Gegengewicht zur hohen Autonomie der Schulen. Das Inspektorat ist auch zuständig für die Erstellung des jährlichen Schulberichts, und es veröffentlicht jährlich im Internet für jede Schule eine sogenannte „Qualitätskarte“, die den Eltern als Grundlage für die Schulwahl dienen kann.

Während im Bereich Leistungsmessung der Ausdruck „Standards“ bisher nicht gebräuchlich war, spielen Standards für die Schulqualität, wie sie in den Rahmenrichtlinien für die Inspektion festgelegt sind, in den Niederlanden eine große Rolle.¹⁵ Auf Anregung des nationalen Bildungsrats sollen Standards im Sinn von Minimalzielen, deren Erreichung überprüft wird, entwickelt werden. Auch hier nimmt also die Steuerung zu, wenngleich sie nicht rigide ist und auf die Profession setzt. Rückmeldungen sind entscheidend, nicht Incentives oder Sanktionen.

¹⁵ Für die Untersuchung der Qualität des niederländischen Bildungswesens hat die Schulinspektion 13 Standards und elf Substandards, für das Sekundarschulwesen sowie die Hochschulbildung vierzehn Qualitätsstandards und die Berufs- und Erwachsenenbildung zehn Standards definiert, die einrichtungsbezogenen Kontext, Prozess- und Wirkungsqualitäten beschreiben. Untersetzt sind diese Standards jeweils mit einer Vielzahl von Indikatoren (im Primärbereich sind es insgesamt 99), die von den Inspektoren standardisiert erfasst und auf einer vierteiligen Skala bewertet werden.

Schweden

Im Gegensatz zu den Niederlanden fand die Dezentralisierung des schwedischen Bildungswesens in den letzten 15 Jahren vor dem Hintergrund eines traditionell zentralisierten, bürokratischen Systems statt, das gemäß wohlfahrtsstaatlicher Maxime auf Gleichheit setzt. Detaillierte Richtlinien und Lehrpläne bildeten die Vorgaben für den Unterricht. Seit Beginn der neunziger Jahre wurden die Curricula drastisch reduziert, das Schulwesen kommunalisiert, die Selbstständigkeit der Einzelschule deutlich gesteigert und als Gegengewicht ein komplexes System der schulischen Evaluation entwickelt. Eine Neuheit bilden die freie Schulwahl sowie inzwischen deutliche Elemente von Wettbewerb zwischen den Schulen.

Auf allen Ebenen des Bildungssystems besteht eine enge Koppelung von „steering by goals“ und „steering by results“: Ziele werden in den nationalen Lehrplänen, in den individuellen kommunalen Schulplänen oder auch in einzel-schulischen Arbeitsplänen gesetzt. Über die Umsetzung und die von den Schülerinnen und Schülern erreichten Leistungen wird im kommunalen und schulischen Qualitätsbericht Rechenschaft zuhanden der nationalen Schulbehörde *Skolverket* bzw. der Gemeinde abgelegt. Hinzu kommen Aufsichtsuntersuchungen und Evaluationen durch die jeweils übergeordneten Ebenen. Den unteren Ebenen (Kommunen, Schulen) wird gleichzeitig ein hohes Maß an Eigenverantwortung und Professionalität für schulische Evaluation und Selbstevaluation zugesprochen, welches jedoch auch mit Problemen behaftet ist. Aufgrund dessen lässt sich in Schweden gegenwärtig eine Pendelbewegung zurück in Richtung einer neuerlichen Stärkung der zentralen Steuerung und Kontrolle feststellen.

In Schweden wird der Begriff „Standards“ für Leistungserwartungen eher gemieden. Faktisch sind jedoch durch die allgemein gehaltenen „goals“ im nationalen Curriculum und in den nationalen Fachlehrplänen sowie durch definierte Kriterien der Leistungsbewertung Standards festgelegt. Nationale Tests werden, wie in den Niederlanden, in der 5. Klasse freiwillig und in der 9. Klasse obligatorisch durchgeführt. Ebenfalls wie in den Niederlanden kann man in Schweden von einer eher losen Kopplung zwischen Standards, Curricula und Tests sprechen. Ähnliches gilt für die verschiedenen Systemebenen. Diese Situation schafft Ambiguitäten, wird aber gleichzeitig als gute Voraussetzung für Lehrerprofessionalität und lokale Entwicklung angesehen.

England

England war bis 1988 ein traditionell lokal verwaltetes Bildungssystem mit hoher Autonomie der Einzelschule. Mit dem Education Reform Act von 1988 erfolgte eine Umgestaltung des Bildungssystems in kurzer Zeit und in radikaler Weise. Eingeführt wurde erstmalig ein nationales Curriculum sowie ein neues Steuerungsmodell, das auf Leistungsmessung und externe Evaluation setzt. Die Schulwahl ist den Eltern freigestellt. Dadurch entstand, politisch gewollt, ein erhöhter Informationsbedarf im Blick auf die Leistungen und die Qualität der Schulen. Die Folge waren regelmäßige Inspektionen der Bildungseinrichtungen und der lokalen Verwaltung einerseits, ein dichtes nationales Test- und Prüfungssystem andererseits, welches sich stark an den Vorgaben des nationalen Curriculums orientiert.

Das 1988 eingeführte National Curriculum enthält klare Vorgaben zu den Unterrichtsinhalten und den erwarteten Leistungen, welche im Sinne von *Content Standards* und *Performance Standards* zu interpretieren sind. Während die Eigenständigkeit der Schulen gegenüber der lokalen Schulbehörde zugenommen hat, gilt dies angesichts der Einführung des National Curriculum nicht für die Lehrpersonen. Gleichzeitig sind die Schulen vor dem Hintergrund ihres erweiterten Gestaltungsspielraumes im Zeichen der Rechenschaftslegung zunehmend zu Objekten staatlicher Evaluation geworden. In Reaktion auf die Schockwirkung im System, die der konservative „Pressure“-Ansatz ausgelöst hatte, zeichnete sich nach dem Regierungswechsel von 1997 ein Trend zu mehr Unterstützung und Qualitätsförderung statt Kontrolle und Sanktionen ab. Die grundlegenden Merkmale der Steuerung blieben aber auch unter der Labour-Regierung erhalten.

3.1 Vereinigte Staaten¹⁶

Die lange Vorgeschichte der Entwicklung von Standards und des Gebrauchs von Leistungstests im amerikanischen Bildungssystem reicht bis in das 19.

¹⁶ Die folgenden Ausführungen sind in Anlehnung an eine Literaturrecherche verfasst, die im Auftrag der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren (EDK) erstellt worden ist: *Literaturanalyse zu Entwicklung, Anwendung und insbesondere Implementation von Standards in Schulsystemen: Nordamerika*. Im Auftrag der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren EDK. Durchgeführt von Esther Berner und Stefanie Stolz, am Lehrstuhl Allgemeine Pädagogik, Pädagogisches Institut der Universität Zürich, unter der Leitung von Jürgen Oelkers.

Jahrhundert zurück. Kein anderes Land verfügt über ähnliche Erfahrungen, aber kein anderes Land hat auch ähnliche Kontroversen ausgebildet. Grosse Debatten über die Qualität der öffentlichen Bildung und über die Struktur des Curriculums sowie die damit verbundenen Leistungserwartungen finden regelmäßig statt, allein seit Ende des Zweiten Weltkriegs lassen sich verschiedene Diskussionswellen nachweisen, die immer auch die Bildungspolitik beeinflusst haben, vor allem weil sie mit Annahmen einer tiefen Krise verbunden waren (Spring 1989; speziell für die fünfziger Jahre und die Naturwissenschaften Rudolph 2002).

Kennzeichnend für die Vereinigten Staaten ist auch, dass die Kontroversen mit gegensätzlichen politischen Lagern verbunden sind und unversöhnlich anmuten. Diane Ravitch hat auf der Basis dieses Konfliktes zwei Paradigmen unterschieden, die die Zukunft der amerikanischen Bildung beeinflussen:

„The *professional education paradigm*, which deeply believes that the profession should be insulated from public pressure for accountability and which is deeply suspicious of the intervention of policymakers; and the *policymaker paradigm*, which insists that the public school system must be subject to some incentives and sanctions based on its performance as are other large scale organizations“ (Ravitch 2002, S. 21; Hervorhebungen d. Verf.).

Tatsächlich lässt sich eine Veränderung der Politik in den Bundesstaaten sowie auf nationaler Ebene seit Beginn der achtziger Jahre nachweisen. Diese Veränderung hat zu tun mit einem stärkeren Einfluss der Bildungsökonomie und einer Diskussion, was den Einfluss des Staates auf die öffentliche Bildung ausmachen soll. „Progressive“ und „konservative“ Konzepte stehen sich seitdem unvermittelbar gegenüber und sorgen für eine anhaltende Lagerbildung. Die Legitimität des einen wird vom je anderen Lager bestritten, wobei Standardargumente ausgetauscht werden, die mit dem je dazu passenden Datenmaterial versehen sind.

3.1.1 Der Wandel der Bildungspolitik

Den Ausgangspunkt der neueren Geschichte der amerikanischen Debatte über Bildungsstandards stellt die Publikation von *A Nation at Risk* dar (National Commission on Excellence in Education 1983). Dieser Bericht zog hohe medi-

ale Aufmerksamkeit auf sich und forderte auf dramatische Weise¹⁷ eine grundlegende Reform des amerikanischen Schulsystems. Mit *A Nation at Risk* wurde die aktuelle Diskussion über Standards eingeleitet. Auch sie setzt die Annahme einer tiefen, epochalen Krise voraus. Der Bericht empfahl unter anderem höhere Anforderungen an die Schulabschlüsse, anspruchsvollere Inhalte der Lehrpläne, strengere Standards des Unterrichts sowie höhere Erwartungen an die Schülerinnen und Schüler einschließlich längerer Schultage und weniger Ferien. Auch die Lehrerbildung war Objekt der Kritik, ebenso die Personalentwicklung und die Ressourcen der Schulen. *A Nation at Risk* stieß trotz – oder wegen – der kaum überzeugenden Datenlage (Berliner/Biddle, 1995) auf große öffentliche Zustimmung und setzte einen massiven politischen Prozess in Gang, der sich zunehmend auf die Idee der Bildungsstandards zentrierte.

In den frühen neunziger Jahren erreichte die Kampagne eines standardbasierten Bildungssystems in den USA ihren ersten Höhepunkt. Ein bedeutender Impuls für diese Bewegung ging vom *National Education Summit* des Jahres 1989 aus. Bei dieser ersten Bildungskonferenz verabschiedeten der damalige amerikanische Präsident George H. Bush und die Gouverneure der einzelnen Bundesstaaten sechs nationale Erziehungsziele (*National Education Goals 2000*) die bis zum Jahr 2000 verwirklicht werden sollten. Das dritte und das vierte dieser Erziehungsziele, die 1991 unter dem Titel *The National Education Goals Report: Building a Nation of Learners* veröffentlicht wurden, beziehen sich ausdrücklich auf die schulischen Leistungen. Das dritte Ziel beinhaltet bereits wichtige Bestandteile eines Standard- und Accountability-basierten Schulsystems. Es impliziert einen nationalen Konsens über die Unterrichtsinhalte in Mathematik, Englisch, Naturwissenschaften, Geschichte und Geografie (Content Standards) und es fordert einheitliche Leistungsniveaus in den Jahrgangsstufen 4, 8 und 12 (Performance Standards). Die ambitionierte Formulierung dieser Ziele war Absicht, um die Aufmerksamkeit der amerikanischen Gesellschaft auf die Dringlichkeit einer verbesserten Ausbildung aller Schülerinnen und Schüler zu konzentrieren.

¹⁷ „If an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war. As it stands, we have allowed this to happen to ourselves. We have even squandered the gains in student achievement made in the wake of the Sputnik challenge. Moreover, we have dismantled essential support systems which helped make those gains possible. We have, in effect, been committing an act of unthinking, unilateral educational disarmament“ (National Commission on Excellence in Education 1983).

Im gleichen Jahr 1991 setzte der amerikanische Kongress das *National Council on Education Standards and Testing* (NCEST) ein. Der Rat, geleitet von zwei Gouverneuren,¹⁸ hatte den Auftrag, Verfahren für die Entwicklung von curricularen Standards und Tests vorzulegen, die zwischen den Bundesstaaten und der Administration in Washington konsensfähig sind. Vorarbeiten hatte das *New Standards Projekt* geleistet,¹⁹ an dem seit 1990 gearbeitet wurde. Das NCEST lieferte seinen Bericht dem Kongress am 24. Januar 1992 ab. In dem Bericht wurde die Bedeutung sowohl von nationalen Standards für die fachlichen Curricula als auch des Aufbaus eines nationalen Assessments betont, mit dem sich die Erreichung der Standards kontrollieren lässt.

1994 wurde unter Präsident Clinton der *Improving America's Schools Act* (IASA) erlassen, der gedacht war als Reautorisierung des *Elementary and Secondary Education Act* (ESEA) von 1965. Durch Title I der IASA wurden die Bundesstaaten dazu verpflichtet, Content und Performance Standards sowie ein System der Leistungserhebung und -bewertung (Assessment) einzuführen, das alle Schülerinnen und Schüler mit einbeziehen sollte. Zwei Jahre später erklärten sich auf dem zweiten *National Education Summit* mehr als vierzig Gouverneure und eine Vielzahl von amerikanischen Unternehmern bereit, die Entwicklung von transparenten und eindeutigen schulischen Standards in den Kernfächern zu unterstützen. Das dritte *National Education Summit* von 1999 bestätigte schließlich die bisherige Entwicklung in dem Sinne, dass erneut die Verantwortungs- und Rechenschaftspflicht der öffentlichen Schulen gefordert wurde.

Eine zentrale Zielsetzung der Entwicklung und Etablierung von Standards war, die im internationalen Vergleich fortwährend schlechten Leistungen der amerikanischen Schülerinnen und Schüler signifikant zu erhöhen und damit verbunden die Qualität des bestehenden Bildungssystems zu überdenken. Auch sollten durch Standards die beharrlich großen Leistungsdiskrepanzen von Schülerinnen und Schülern verschiedener Nationalitäten, unterschiedlicher Ethnien und sozial schwächerer Gruppen reduziert werden, um die Chancengleichheit

¹⁸ Die beiden Gouverneure waren der Demokrat Roy Romer (geb. 1928) (Colorado) und der Republikaner Caroll A. Campbell Jr. (1940-2005) (South-Carolina).

¹⁹ Das Projekt wurde durchgeführt vom *National Center of Education and the Economy* in Washington, D.C. und vom *Learning and Development Research Center* der University of Pittsburgh. Das National Center ist ein 1989 gegründetes Non-Profit-Unternehmen, das aus Spenden finanziert wird. Das Center stellt seit 1998 Schulen ein standardbasiertes Reformprogramm zur Verfügung, das nach eigenen Angaben von mehr als 500 Schulen in 15 Bundesstaaten implementiert wurde, vgl. <http://www.ncee.org/> (Stand 15.08.06).

zu verbessern. Weiterhin sollte durch die Implementation von Standards eine bis dahin nicht vorhandene klare Festlegung erfolgen, was Schülerinnen und Schüler lernen sollen. Nur so würden sich die Leistungen objektiv beurteilen lassen.

Im Jahre 2000 wurden in den Vereinigten Staaten mehr als 422 Mia. US-Dollar für die Elementar- und Sekundarbildung ausgegeben. Die Schulleistungen werden durch die NAEP-Tests²⁰ auf nationaler Ebene und bezogen auf Jahrgänge gemessen. Die jährlichen Berichte *NAEP at a glance* beziehen sich auf die Testergebnisse am Ende der vierten, der achten und der zwölften Klasse (grades). Es gibt drei Kompetenzstufen, die erreicht werden können, nämlich basic, proficient und advanced. Einige Ergebnisse zeigen das Problem: Etwa 60% der High School Seniors verfügen nicht über Basiskenntnisse der amerikanischen Geschichte, mehr als die Hälfte der Viertklässler aus einkommensschwachen Familien können nicht auf Basisniveau lesen. Im internationalen Vergleich belegen amerikanische Schülerinnen und Schüler hintere Plätze. In der TIMSS-R-Auswertung von 1999 waren dies ein 19. Rang von 38 Ländern in Mathematik und ein 18. Rang in den Naturwissenschaften.

Die Politik hat auf diese Situation reagiert. Ein von den Bundesstaaten zunehmend genutztes Mittel sind Voucher-Programme, die zu mehr Möglichkeiten der Schulwahl führen sollen. *School-Choice* ist ein zentrales Thema in den Vereinigten Staaten, das auch von Bildungsexperten unterstützt wird. Die Koret Task Force on K-12 Education, ein Beratungsgremium, das von der Hoover Institution zusammen mit der Koret Foundation eingesetzt wurde,²¹ empfiehlt ausdrücklich den Ausbau von Wahlmöglichkeiten und Stipendien- oder Voucher-Programmen (Peterson 2003). Ein Grund ist, den ärmeren Familien Alternativen zu ihren örtlichen Schulen zu bieten, die zumeist „low-performing schools“ sind. Ein weiterer Grund ist die Auslösung von Wettbewerb. Auch Eltern reagieren auf die gegebene Situation. Im Schuljahr 2001/02 wurden rund zwei Millionen Kinder in der Form des Homeschooling unterrichtet, die Zahl nimmt ständig zu (Clowes 2003).

²⁰ National Assessment for Educational Progress. Die Tests sind 1969 eingeführt worden und betreffen folgende Lernbereiche bzw. Fächer: reading, mathematics, science, writing, U.S. history, civics, geography, arts. Vgl. <http://nces.ed.gov/nationsreportcard/about/> (Stand 08.07.06). Hier sind auch die Leistungsdaten einsehbar.

²¹ Der derzeitige Leiter des Gremiums ist Chester E. Finn Jr. Mitglieder sind unter anderem Diane Ravitch, Herbert Walberg und Caroline Hoxby.

Der Nationale Rat der amerikanischen Mathematiklehrer (NCTM) hatte 1989 als Antwort auf *A Nation at Risk* und parallel zum ersten Bildungsgipfel einen Entwurf für Standards des Curriculums und der Evaluation der Lernergebnisse vorgelegt, an dem sich die Diskussion sehr nachhaltig orientierte.²² Die Erarbeitung der *Curriculum and Evaluation Standards for School Mathematics* erfolgte in enger Zusammenarbeit mit Lehrpersonen, führenden Mathematikprofessoren und Fachdidaktikern – eine Initiative, der bald auch andere professionelle Organisationen folgten. 1991 veröffentlichte der NCTM professionelle Standards für die Praxis, also den Unterricht in Mathematik,²³ die auch auf die Ausbildung der Lehrkräfte abzielten. Ein bestimmtes Kompetenzniveau sollte in den Anforderungen an Lehrberufe nicht unterschritten werden dürfen. Aus Vorstößen wie diesem entstanden schließlich Standards für die amerikanische Lehrerbildung, die von großen Evaluationsagenturen entwickelt wurden. Diese Standards betrafen die Curricula und die Organisation der Lehrerbildung.

Im Jahr 2000 hat der NCTM erneut überarbeitete *Principles and Standards* als Rahmenkonzept eines modernen Mathematikunterrichts vorgelegt. Die stark beachteten Richtlinien konkretisieren nicht nur die Lerninhalte (als Content Standards), sondern verknüpfen diese mit prozess- und methodenbezogenen Opportunity-to-Learn-Standards sowie mit allgemein didaktischen Prinzipien eines curricular bedeutsamen und verstehensorientierten Mathematikunterrichts. Diese Vorgabe ist als konkrete Umsetzung und paradigmatische Materialisierung der Theorie der Bildungsstandards zu verstehen, der andere Fächer gefolgt sind.

Im Laufe der neunziger Jahre wurden in vielen Schulfächern detaillierte Vorschläge für nationale Standards²⁴ entwickelt, die sich aber gegenüber den Standards der einzelnen Bundesstaaten nicht durchsetzen konnten. Zumindest entstand kein einheitliches nationales Curriculum, das zu entwickeln ursprünglich als politische Botschaft kommuniziert worden war.²⁵ Zwar wurde dem *National Education Standards and Improvement Council* (NESIC), welches im Rahmen

²² National Council of Teachers of Mathematics: *Curriculum and Evaluation Standards for School Mathematics* (Reston, VA: NCTM 1989).

²³ National Council of Teachers of Mathematics: *Professional Standards for Teaching Mathematics* (Reston, VA: NCTM 1991).

²⁴ Wie *National Standards for History*: <http://nchs.ucla.edu/standards/> (Stand 12.08.06).

²⁵ Die Idee geht zurück auf eine Konferenz in Charlottesville, VA im Jahre 1989. Der damalige Präsident George H. Bush und die Gouverneure der fünfzig amerikanischen Bundesstaaten vereinbarten *National Education Goals*. Ein Ziel betraf die Festlegung der Kernfächer des Curriculums, nämlich Englisch, Mathematik, Naturwissenschaften, Geschichte und Geografie.

der *Goals 2000: Educate America Act* (1994) gegründet worden war, die Aufgabe zugesprochen, die Entwicklung nationaler Standards zu beaufsichtigen und die von den Bundesstaaten entwickelten Standards zu zertifizieren. Dieser Versuch, nur die Kontrolle über (freiwillige) nationale sowie die staatlichen Standards einer nationalen Agentur zu übergeben, scheiterte unter der Regierung Clinton am Widerstand der republikanischen Mehrheit im Kongress.

Die öffentliche Zustimmung in den neunziger Jahren zum gesamten Standardprojekt war hoch. In verschiedenen Umfragen wurde die Meinung unterstützt, dass höhere Erwartungen bessere Leistungen bewirken würden. Auch die Schülerinnen und Schüler waren in Umfragen der Meinung, dass höhere Leistungsstandards sie veranlassen würden besser zu lernen. Das deckte sich schließlich mit den Erwartungen der Lehrkräfte, die durch höhere Standards vor allem die Leistungen in den akademischen Fächern gefördert sehen wollten. Offenbar hat die Diskussion nicht nur über Standards, sondern über *höhere* Standards die Erwartungen der Öffentlichkeit gesteigert, was auch mit verbreiteten Krisenannahmen zu tun hat, die in diesem Ausmaß spezifisch sind für amerikanische Diskussionen über öffentliche Bildung (vgl. die demoskopischen Studien von Farkas/Friedman/Boese et al. 1994; Friedman/Duffet 1997).²⁶

Nachdem kein all-amerikanisches Curriculum entstand, sind für die Steuerung des Bildungssystems nationale Anreizsysteme und monetäre Steuerungsimperative ausgebildet worden, die die zweite Administration Bush auf den Weg brachte. Am 8. Januar 2002 unterzeichnete Präsident George W. Bush den *No Child Left Behind Act of 2001*²⁷, der seitdem die amerikanische Entwicklung sehr nachhaltig bestimmt. In einer Regierungserklärung vom Januar 2001 hatte Bush den öffentlichen Schulen sein „tiefes Vertrauen“ ausgesprochen, aber gleichzeitig seiner Beunruhigung Ausdruck verliehen, dass „too many of our neediest children are being left behind“, obwohl seit 1965²⁸ mehr als 200 Billionen US-Dollar allein aus Bundesmitteln in die öffentliche Bildung geflossen seien. Das Gesetz *No Child Left Behind* (NCLB) trägt dieser Frage Rechnung:

²⁶ Diese und ähnliche Studien wurden von der New Yorker Agentur *Public Agenda* durchgeführt. Die Agentur ist „a non partisan opinion research and civic engagement organisation“ (vgl. <http://www.publicagenda.org/>, Stand 08.08.06). Sie wurde 1975 von Daniel Yankelovich und Cyrus Vance gegründet und ist unabhängig. Der Auftrag bezieht sich auf Themen von öffentlichem Interesse.

²⁷ U.S. Department of Education: *No Child Left Behind. Law – P.L. 107-110 (h:R.1)* (vgl. <http://www.ed.gov/offices/OESE/esea/>, Stand 08.08.06).

²⁸ 1965 war das Jahr der Verabschiedung des *Elementary and Secondary Education Act* (ESEA).

Wie können die Mittel effizienter so eingesetzt werden, dass auch und vor allem die Kinder profitieren, die die Schule bislang am wenigsten nutzen konnten?

Das Gesetz stellt vier Strategien in den Mittelpunkt, die allesamt auf die Forderung nach mehr Effizienz reagieren:

- *Increased Accountability*: Verbesserte Leistungsbilanz und Steuerung des Systems vom Resultat aus;
- *More Choices for Parents and Students*: Verbesserte Schulwahl und Einsatz von Bildungsgutscheinen;
- *Greater Flexibility for States, School Districts, and Schools*: Verbesserter Mitteleinsatz und flexiblere Reaktion auf Probleme vor Ort;
- *Putting Reading First*: Verbesserung der Elementarbildung mit dem Vorrang der Leseförderung.

Die Leistungsbilanzen von Schulen müssen sich an Standards orientieren, die erreichte Qualität wird regelmäßig getestet und zwischen den Schulen soll es Wettbewerb geben. Die je erreichte Qualität wird öffentlich gemacht. Die Schulwahl und der Einsatz von Bildungsgutscheinen sollen Eltern und Schülerinnen und Schülern die Chance geben, Schulen, die als schlecht identifiziert wurden, verlassen zu können, ohne dadurch Nachteile zu erleiden. Die lokalen Schulbehörden können dafür Bundesmittel flexibel einsetzen, ebenso wie für die gezielte Verbesserung der schlechten Schulen. Bei dem ganzen Programm spielt die Anhebung basaler Kompetenzen, vor allem Lesen, von Kindern aus bildungsfernen Milieus eine zentrale Rolle. Das Gesetz regelt auch, dass nur noch solche Methoden des Unterrichts zugelassen sind, die mit empirischen Daten nachweisen können, dass sie Erfolg haben.

Alle amerikanischen Bundesstaaten müssen die Schülerinnen und Schüler regelmäßig testen, um den Nachweis führen zu können, dass sie ein bestimmtes Minimum erreicht haben. Das betrifft allerdings nur die öffentlichen Schulen, nicht die privaten. Das NCLB-Gesetz sieht weiterhin vor, dass der *Adequate Yearly Progress* beschrieben wird. Damit geben die Schulen Rechenschaft über die jährlichen Lernfortschritte ihrer Schülerinnen und Schüler, die erreicht werden *müssen* und also einen entsprechenden Einsatz der Ressourcen verlangen. Vor allem in diesem Sinne soll niemand zurückgelassen werden. Die Entwicklung legt konkrete Verantwortung fest und verpflichtet zur Transparenz. Daher ziehen die Befürworter der Standardsbewegung eine positive Bilanz des Gesetzes (vgl. Chubb 2005).

Der *No Child Left Behind Act* hatte noch einen weiteren Effekt, der offenbar auch bildungspolitische Absicht war. Durch dieses Gesetz fand eine bemerk-

kenswerte Verlagerung der Zuständigkeiten hin zur nationalen Bildungspolitik statt, welche vorher nur für die Finanzierung einkommensschwacher Schülerinnen und Schüler verantwortlich war und nun maßgeblich für die Ziele, Ergebnisse und Folgen von Bildung mitverantwortlich zeichnet, und zwar unter angestrengten Bedingungen der Accountability der Staaten und einzelnen Schulen. Verbunden mit der Novellierung des Elementary and Secondary Education Act (1965) stellt dem *No Child Left Behind Act* einen sehr umfassenden Eingriff in den traditionell dezentral organisierten Bildungsbereich dar.

Die Bundesmittel sind auch deswegen so bedeutsam, weil ein Grossteil der Budgets öffentlicher Schulen aus der lokalen Einkommenssteuer gespiesen wird. Die Ressourcen der einzelnen Schulen sind entsprechend nicht nur sehr unterschiedlich, sondern oft unvergleichbar. Wenn schon zwischen benachbarten Schuldistrikten aufgrund des Steueraufkommens die Klassengröße variiert, ist leicht vorstellbar, dass öffentliche Schulen umso angesehener sind, je wohlhabender ihr Standort ist. Bundesmittel sind daher gerade für die ärmeren Distrikte ein hoher Anreiz, sich in die Richtungen zu entwickeln, die das NCLB-Gesetz vorgegeben hat.

Im Jahre 2000, also vor Verabschiedung des NCLB-Gesetzes, resümierte die *Mid-Continent Research for Education and Learning* (McREL) die bestehende Situation zum standardbasierten Unterricht und die damit verbundenen Herausforderungen folgendermaßen:

- Es gibt nur wenig Übereinstimmung darüber, wie standardbasierte Bildung aussieht bzw. aussehen könnte. Dementsprechend spiegeln Forschungsberichte und Evaluationsstudien eine große Differenz in der Herangehensweise wider.
- Das Drängen auf *High-Stakes-Accountability* tendiert dazu, die Rahmenbedingungen für standardbasierte Lehrpläne und Ausbildungen einzuengen.
- Es besteht kein eindeutiges, umfassendes und kohärentes System bezüglich Standards und Benchmarks; lokale, staatliche und nationale Standards unterscheiden sich zum Teil erheblich.
- Lehrkräfte haben oft keine hochwertigen Unterrichtsmittel, die auf die lokalen Standards abgestimmt sind.
- Es gibt kaum Angaben über eine standardbasierte Vorgehensweise, die sich an die Lernbedürfnisse aller Schülerinnen und Schüler richtet und insbesondere Lernende mit einbezieht, die die festgelegten Standards *nicht* erreichen konnten.

- Notwendig sind mehr Informationen zur Finanzierung – und Finanzierbarkeit – bezüglich der Implementation von Bildungsstandards (Mid-continent Research for Education and Learning (McREL) 2000).

Rückblickend fand in den letzten zwei Jahrzehnten auf nationaler und insbesondere auf staatlicher und regionaler Ebene eine intensive Diskussion um die Entwicklung von Standards statt, die zur Implementation von standardbasierten Curricula in den jeweiligen Bundesstaaten führte. Bedingt durch den *No Child Left Behind Act* haben die Bundesstaaten in den letzten Jahren mit unterschiedlicher Geschwindigkeit und unterschiedlichem Erfolg Verfahrensweisen eingeführt, die den Anforderungen des Gesetzes entsprechen sollen. Der *No Child Left Behind Act* forciert die staatlichen Bemühungen in den Bereichen Assessment, Accountability und Teacher Quality mit zum Teil einschneidenden Auswirkungen bis hin zur Qualitätsüberprüfung der Leistungen der Lehrkräfte. Dieses Jahr, 2007, steht die Reauthorisierung des Gesetzes durch den Kongress an – eine der Hauptprioritäten im Regierungsjahr für Georg W. Bush, für den denn auch das Kernanliegen der Accountability als Instrument der Leistungssteigerung nicht in Frage steht. Auch die breite Öffentlichkeit unterstützt die Hauptprinzipien von NCLB, wie die aktuellen Auseinandersetzungen zum Thema zeigen, wenngleich Korrekturen gefordert werden (vgl. z.B. die Beiträge im *American Educational Research Journal*, 44 (2007)). Dies betrifft neben der Forderung nach adäquater Finanzierungshilfe durch den Bund, um die aufgestellten Ziele erreichen zu können, vor allem auch die Methode der Accountability und damit der Tests zur Messung des geforderten AYP. Repräsentativ sind etwa die Verlautbarungen der National Education Association (NEA), der größten nationalen Berufsorganisation von Lehr- und Erziehungspersonen. Hauptanliegen sind die Verwendung von zusätzlich anderen statt lediglich Ergebnissen aus standardisierten Tests für die Messung der Leistungen von Schülerinnen und Schülern und Schulen, dabei die konsequente Verwendung von Modellen des Leistungszuwachses, die Reduktion der Klassengrößen sowie eine Hebung der Lehrerqualifikationen.²⁹ Das Anliegen einer neuen, fairen Ausgestaltung von Accountability wird im Hinblick auf die Erneuerung von NCLB auch vom Forum on Educational Accountability vorgebracht. Insgesamt schlägt das Forum vor, auf Capacity-Building zur Unterstützung systemischer Reformen mit dem Ziel des „school improvement“ zu setzen, anstelle des gegenwärtigen bestrafenden Systems der Rechenschaftslegung. „The new law

²⁹ Vgl. <http://www.nea.org/lac/esea/index.html> (Stand 04.05.07).

should hold schools, districts and states accountable for implementing the essential systemic changes rather than continuing to hold them accountable for failing to reach arbitrary Adequate Yearly Progress (AYP) testing targets“ (Forum on Educational Accountability 2007, S. 1).

Insgesamt fallen die Resümees betreffend die Entwicklung der Schülerleistungen unter NCLB, wie sie jetzt vor Erneuerung des Gesetzes überall publiziert werden, wie immer, je nach Studie unterschiedlich und zum Teil widersprüchlich aus.

3.1.2 Verständnis und Bedeutung von Standards

Der Begriff „Standards“ wird vielfältig gebraucht und von unterschiedlichen Interessengruppen im Bildungssystem eingesetzt. Der Ausdruck entfaltet so spezifische Konnotationen, die die jeweiligen Dispositionen und Erfahrungen sehr verschiedener Akteure zum Ausdruck bringen.³⁰ Eine zusammenfassende Definition hat der Nationale Forschungsrat im Jahre 2001 vorgelegt. Sie lautet wie folgt:

„For many members of the general public and for the education policy community, ‚standards‘ focus on outcomes and imply a mechanism by which to hold schools accountable for what students learn. In such cases, specific levels of performance relative to standards are defined, and assessments are designed to measure student progress toward attaining those standards. Assessment results may then be used as part of an accountability system, as a professional development tool to provide feedback to teachers, or to inform policy decisions [...]. To many educators, a ‚standard‘ is a statement describing what a person should know or be able to do. That use of ‚standard‘ is often called a ‚content standard‘“ (National Research Council 2001).

³⁰ „Education researchers, policy researchers, policymakers, politicians, administrators, and teachers are all talking about standards. Yet, each group and individuals within the groups have experiences and dispositions that influence their thinking about the central questions: „What is a standard?“ Clearly the word „standard“ has different meanings and purposes for the various actors within in the education system. That is part of the appeal of the word“ (Tate 2004; vgl. auch Jofus/Berman 1998 oder Resnick/Nolan 1995).

Entsprechende Definitionen sind seit längerem in Gebrauch.³¹ Nach der von Diane Ravitch kodifizierten Unterscheidung müssen neben Standards, die die Inhalte und Ziele des Unterrichts (Content Standards, Curriculum Standards) beschreiben, auch Standards für die Leistungen oder die Niveaus des Lernens (Performance Standards) und Standards für die Ausstattung oder die Ressourcen (Opportunity-to-Learn Standards) bestimmt werden. Zusätzlich muss entschieden werden, auf welchen Ebenen die Standards eingesetzt werden sollen. So gibt es National Standards der professionellen Fachorganisationen, State Standards und auch District Standards. Oft rivalisieren diese Standards oder zwingen zu aufwändigen Anpassungen. Gerade im Zusammenhang mit Fragen der Implementation wird so immer wieder die Frage nach dem Alignment der Standards aufgeworfen (vgl. Baker 2004a).

Inhaltliche Bildungsstandards für die Schulfächer sind auf allen drei Ebenen entwickelt worden. Entwürfe für nationale Curricula liegen wie erwähnt für Mathematik durch das NCTM und die Naturwissenschaften durch das NRC vor, daneben gibt es entsprechende Curricula der Bundesstaaten und auch mancher Distrikte. Sie, die Distrikte, spielen vor allem dann eine Rolle, wenn die Bundesstaaten sehr groß sind und eine hohe Bevölkerungsdichte aufweisen. Traditionell hat New York City auf die Entwicklung von Lehrmitteln und anderen Standards immer einen hohen Einfluss gehabt. Hier gibt es 1200 öffentliche Schulen mit mehr als einer Million Schülerinnen und Schüler.

Die Content Standards konzentrierten sich zunächst auf die Kernfächer. Die Folge der rivalisierenden Konstruktion war und ist eine breite Vielfalt von Standards. „Some states set standards for grade clusters, like the National Science Education Standards, while others set standards for each grade. Some focus on a few big ideas, while others are quite extensive“ (National Research Council 1999). Einige Sets von Content Standards können direkt für den Unterricht eingesetzt werden, andere müssen dazu zuerst in einem Prozess der Curriculumentwicklung (Curriculum Frameworks und weitere Curriculum materials) aufbereitet werden.

³¹ „Content standards specify what students should know and be able to do. In shorthand, they involve the knowledge and skills essential to a discipline that students are expected to learn. Those ‚skills‘ include the ways of thinking, working, communicating, reasoning, and investigating that characterize each discipline. That ‚knowledge‘ includes the most important and enduring ideas, concepts, issues, dilemmas, and information of the discipline“ (National Education Goals Panel 1993, S. 9).

Die Bestimmung der Inhalte war oft heterogenen Gremien überlassen, in denen sehr unterschiedliche Interessen vertreten waren. Fachdidaktiker, Pädagogen, Vertreter der Gemeinden und Politiker hatten sehr unterschiedliche Erwartungen an das Wissen und Können von Schülerinnen und Schülern. Die Bandbreite reichte von *readiness for the workplace* über *knowledge and skills young people need for effective citizenship* bis hin zu *students need to understand an increasingly multicultural society*. Auch das erklärt, warum Mathematik und Naturwissenschaften fachliche Standards mit einem weit größeren Konsens als andere Fächer vorlegen konnten.

Ein zentrales Element des standardbasierten Bildungssystems sind *Performance Standards*, die komplementär zu den Content Standards zu verstehen sind. Performance Standards beziehen sich auf die Leistungen und bringen zum Ausdruck, was die Schülerinnen und Schüler mit bestimmten Anstrengungen erreichen können (National Research Council 1999). Sie bestimmen den Grad der Beherrschung oder das Niveau der Kenntnisse, wobei Stufen der Erreichbarkeit unterschieden werden. Eine bestimmte Leistung kann so für eine unzureichende, eine zufriedenstellende oder eine hervorragende Fähigkeit stehen.

„Performance standards specify ‚how good is good enough‘. In shorthand, they indicate how adept or competent a student demonstration must be to indicate attainment of the content standards. They involve judgments of what distinguishes an adequate from an outstanding level of performance. Performance standards are not the skills and modes of reasoning referred to in the content standards. Rather they indicate both the nature of the evidence (such as an essay, mathematical proof, scientific experiment, project, examination, or combination of these) required to demonstrate that the content standards have been met and the quality of student performance that will be deemed acceptable“ (National Education Goals Panel 1993, S. 22).

Ferner nehmen Performance Standards eine wichtige anleitende Funktion ein. Durch anschauliche und substanzielle Beispiele sollen sie Schülerinnen und Schüler, Eltern und Lehrkräfte bei der Anhebung der Leistungen anleiten. Nach Hansche (1998) bestehen Performance Standards aus vier Elementen:

- *Performance Levels*: Schülerleistungen werden nach verschiedenen Kategorien klassifiziert (bspw. below partially proficient – partially proficient – proficient – advanced);

- *Performance Descriptors*: Beschreibungen der Schülerleistungen für jede einzelne Leistungskategorie;³²
- *Exemplars*: Anschauliche Beispiele für Schülerleistungen auf jedem Niveau;
- *Cut scores*: Punktbewertungen, die die verschiedenen performance levels voneinander trennen (Hansche 1998).

Die Etablierung eines standardbasierten Schulsystems erfordert neben der inhaltlichen Reform auch eine Form der materiellen und finanziellen Unterstützung. Dementsprechend bezeichnen Opportunity-to-Learn Standards (oder auch School Delivery Standards) die Faktoren, die festlegen, welche Bedingungen bzw. welche Umgebung Voraussetzung ist, um den Schülerinnen und Schülern die Erreichung der Standards zu ermöglichen.

„According to the Goals 2000: Educate America Act, an opportunity-to-learn (OTL) standard means the criteria for, and the basis of, assessing the sufficiency or quality of the resources, practices, and conditions necessary at each level of the education system (schools, local educational agencies, and states) to provide all students with an opportunity to learn the material in voluntary national content standards or state content standards“ (Porter 1995, S. 21).

Leicht abweichend von dieser Differenzierung unterscheidet etwa die *Education Week* folgende Typen von Standards:

„*Academic standards* describe what students should know and be able to do in the core academic subjects at each grade level.

Content standards describe basic agreement about the body of education knowledge that all students should know.

Performance standards describe what level of performance is good enough for students to be described as advanced, proficient, below basic, or by some other performance level“ (Education Week on the Web 2005).

³² Beispiele für allgemeine Performance Descriptors des NAEP:

Basic: This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.

Proficient: This level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.

Advanced: This level signifies superior performance. (Hansche 1998, S. 16f.)

McREL empfiehlt eine detailliertere Unterscheidung zwischen Content und Curriculum Standards:

„Content standards describe the knowledge and skills that students should attain. Curriculum standards, on the other hand, can describe overarching goals, or ways in which the curriculum should be orchestrated to achieve a desired result. In either case, curriculum standards do not explicitly describe student knowledge and skills as do content standards. Mixing both types of standards as if there were no significant difference creates unnecessary confusion“ (Mid-continent Research for Education and Learning (McREL) 2006).

Die Entwicklung der Standards hat zu tun mit dem Aufbau neuartiger Formen von Leistungsmessung, die mit der Frage verknüpft sind, welche Verantwortung Schulen wie Lehrkräfte beim Zustandekommen der Leistungen ihrer Schülerinnen und Schüler übernehmen. Die Frage der Verantwortung ist ihrerseits verbunden worden mit Rechenschaftslegung (Accountability) und möglichen positiven wie negativen Sanktionen.

3.1.3 Überprüfung von Standards: Tests und Accountability

Leistungsüberprüfung, der Einsatz von Testsystemen und die Zuschreibung von Verantwortung waren bei den Reformbemühungen seit dem Zweiten Weltkrieg immer von Bedeutung, entweder als Mittelpunkt von Kontroversen oder als „Lieblinge“ der politischen Entscheidungsträger (Madaus 1985). Aber erst seit Beginn der neunziger Jahre ist daraus mehr geworden als politische Semantik, die mit den unterschiedlichsten Konzepten kompatibel schien. Die Verantwortung der Schulen für die Leistung ihrer Schülerinnen und Schüler, dementsprechend für die Qualität der Bildung, ist zusammen mit den neuen oder erweiterten Formen der Leistungsüberprüfung zu einem der bedeutendsten Konzepte der standardbasierten Schulreform avanciert. „Accountability for student performance is one of the two or three – if not the most – prominent issues in a policy at the state and local levels right now“ (Olson 1999).

Die Schulen sollen und müssen nachweisen, was sie leisten und dafür Verantwortung übernehmen. Aber auch die Schülerinnen und Schüler tragen Verantwortung für das, was sie leisten oder nicht leisten.

„*School and district accountability* systems typically include efforts to assess and rate schools or districts based on student performance and other indicators, to publicly report on school or district performance, and to provide rewards and sanc-

tions for schools or districts based on performance or improvement over time. *Student accountability* generally refers to efforts to hold students responsible of their own performance by requiring students to pass a test to be promoted from grade to grade or to pass an exam to graduate from high school“ (Education Week 2005b).

In den Vereinigten Staaten existieren drei Hauptvarianten von Accountability-Systemen. Die einen bewerten die Schulen nach dem „value added“, das heißt aufgrund der Leistungszunahme von bestimmten Schülergruppen über die Zeit; die anderen bewerten die Schulen auf der Basis eines Prozentsatzes an Schülerinnen und Schülern, die über einer bestimmten Marke abschneiden; im dritten Fall werden die Schulen aufgrund der Zu- und Abnahme ihrer Test-Scores bewertet. Damit verlässliche Aussagen über Leistungszunahmen gemacht werden können, ist in jedem Fall eine längere Entwicklung notwendig, während die Politik eher an kurzfristigen Resultaten interessiert ist.

Bis weit in die achtziger Jahre hinein überprüften die Bundesstaaten allein die Inputs der öffentlichen Schulbildungen und schenkten den eigentlichen Leistungen der Schülerinnen und Schüler wenig Aufmerksamkeit. Ausschlaggebend für die Bewertung waren Indikatoren wie die Anzahl der Bücher in der Schulbibliothek oder die Anzahl der Computer in den Klassenräumen. Ausschlaggebend für die Entwicklung von Accountability und damit für die Überprüfung von Outputs wie Testergebnissen und Schulabschlussraten war ein Vorschlag der Gouverneure, den Schulen größere Handlungs- und Organisationsspielräume einzuräumen sowie zusätzliche finanzielle Mittel, wenn im Gegenzug Schulen und Lehrkräfte mehr Verantwortung für die Leistung der Schülerinnen und Schüler übernehmen würden (Education Week 2006b).

Im Zusammenhang mit den 1994 erlassenen *Goals 2000: Educate America Act*, dem *Improving America's Schools Act* (IASA) und der damit verbundenen Revision des *Elementary and Secondary Education Act* (ESEA) sowie der Title-I-Gesetzgebung erhielt die „new accountability“³³ eine gesetzliche Basis. In diesem Rahmen vollzog sich innerhalb weniger Jahre ein deutlicher Wandel von der Input-Orientierung hin zu einer auf den Output bezogenen Bewertung des Schulsystems, die sich konzentriert auf

³³ Die Philosophie der *New Accountability* betont Schülerleistungen gegenüber System-Input, fokussiert eher die einzelnen Schulen statt die Schuldistrikte als Einheiten der Verbesserung, publiziert Schüler-Outcomes und reagiert mit Belohnungen und Sanktionen, um Schulen zu motivieren, ihr Curriculum und ihren Unterricht zu verändern (Fuhrman 1999).

- die Leistungen der Schülerinnen und Schüler,
- die einzelnen Schulen als die Orte der Entstehung und Verbesserung der Leistungen und
- die öffentliche Berichterstattung über die Ergebnisse und die mit den Leistungen verbundenen Konsequenzen (Fuhrman 1999).

Durch das NCLB-Gesetz wurden, wie erwähnt, alle Schulen verpflichtet, die Leistungen der Schülerinnen und Schüler zu testen (Kornhaber 2004). Damit verbunden waren entscheidende Bestimmungen hinsichtlich der Accountability und der Leistungsbewertung, die vielen Staaten tiefgreifende Veränderungen auferlegten. Das gilt nicht nur für die Frage, wie getestet werden soll. Eine wesentliche Veränderung, die in der europäischen Diskussion bislang kaum beachtet wurde, betrifft den *Adequate Yearly Progress* (AYP), mit dem bestimmte Maßnahmen³⁴ wie die Mittelzuteilung oder positive und negative Sanktionen verknüpft werden (Linn/Baker/Betebenner 2002; Erpenpach/Forte 2005). Bezüglich des AYP werden in der NCLB den Bundesstaaten verbindliche Vorgaben gemacht:

- Die Staaten müssen messbare AYP-Ziele entwickeln für die Verbesserung der Leistungsfähigkeit, gültig für alle Schülerinnen und Schüler und insbesondere im Hinblick auf spezielle Gruppen: finanziell benachteiligte Schülerinnen und Schüler, Schülerinnen aus Minderheitengruppen, Schüler mit Körperbehinderungen und Schüler mit begrenzten Englischkenntnissen.
- Die Zielsetzungen müssen so ausgerichtet sein, dass alle Schülerinnen und Schüler den Proficient Level oder ein noch höheres Niveau innerhalb von zwölf Jahren nach Beginn der Messung erreicht haben.
- Der AYP muss hauptsächlich auf dem staatlichen Assessment beruhen, aber zudem noch einen zusätzlichen schulischen Indikator beinhalten.

³⁴ „Under NCLB, scores will be used to determine whether adequate yearly progress has been met and what sort of consequences will ensue. For example, schools that receive Title I funding but fall short of adequate yearly progress for 2 consecutive years must provide students with other public school choice. Those that fall short for 3 consecutive years must provide students with additional services, such as tutoring. Such functions will further stretch the already-thin resources of schools whose populations include a large percentage of low-scoring students“ (Kornhaber 2004, S. 49).

- Die AYP-Zielsetzungen müssen auf der Ebene der Schule bewertet werden. Schulen, die ihre AYP-Ziele in zwei aufeinanderfolgenden Jahren nicht erreicht haben, werden auf eine bestimmte Form der Verbesserung festgelegt.
- Die Ergebnisse des AYP müssen separat für jede Schülergruppe aufgeführt werden, um nachvollziehen zu können, ob jede Gruppe die AYP-Ziele erfüllt hat.
- Aus jeder Schülergruppe müssen mindestens 95% an staatlichen Assessments teilnehmen (Linn/Baker/Betebenner 2002).

Alle Bundesstaaten verfügen inzwischen über Accountability-Systeme. Seit dem NCLB-Gesetz sind sie bemüht, die Vorgaben des Gesetzes umzusetzen (Erpenbach/Forte 2005). Dennoch variieren die Konzepte von Staat zu Staat. So hat die Mehrheit der Staaten zwar Performance Standards eingeführt, diese weichen aber bezüglich der Niveaus der Standards (wie basic oder proficient levels), der Prozentsätze der Schülerinnen und Schüler, die die gesetzten Standards erreichen sollen, sowie der zeitlichen Fristen, die sich die Schulen gesetzt haben, um die Ziele zu erfüllen, erheblich voneinander ab (Fuhrman 2004). Die Effekte sind derzeit noch schwer abschätzbar. Angesichts der schnellen bildungspolitischen Umsetzungen fehlen derzeit größere Studien zur Wirksamkeit der von dem NCLB-Gesetz geforderten Accountability (Müller/Silver 2006). Neuerliche Hinweise sind vorhanden, dass besonders die leistungsschwachen Schulen bzw. deren Lehrpersonal negativ auf eine forcierte Rechenschaftslegung reagieren (Finnigan/Gross 2007).

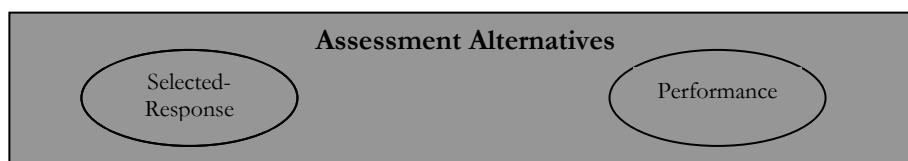
Mit der zunehmenden Output-Orientierung und erhöhter Accountability gehen Veränderungen in den Assessment-Verfahren einher. Diese Veränderungen besagen nicht, dass Inputs unbedeutend geworden sind, jedoch wird gemäß dem oben beschriebenen Politikwechsel den Lernresultaten seit Anfang der neunziger Jahre die relevantere Position im Bildungssystem zugesprochen (Mid-continent Research for Education and Learning (McREL) 2000). Allgemein versteht man unter Assessment³⁵ ein Verfahren zur Erhebung und Bewertung von Schülerleistungen.

³⁵ Als „Assessment“ wird allgemein das Sammeln von Informationen über Individuen, Gruppen oder Systeme bezeichnet, das auf einer Mehrzahl von Instrumenten beruht, wozu auch ein Test gehören kann. „Assessment“ ist also ein umfassenderer Begriff beziehungsweise Prozess als „Testen“ (Council of Chief State School Officers 2003).

„Assessment is a process that uses tests and/or other instruments and procedures to collect information that is for use in making appropriate inferences about student learning. Included are tests or examinations used at all levels of the educational process, including assessment at the state, district, school, and classroom levels“ (La Marca/Redfield/Winter et al. 2000, S. 1).

In einem standardbasierten Schulsystem geben Assessments Aufschluss über die Leistungen der Schülerinnen und Schüler in Bezug auf die vorher festgelegten Standards. Die möglichen Formen von Assessment im amerikanischen Bildungswesen sind sehr vielfältig (Abb. 2). Sie reichen von „standardized tests“ bis hin zu verschiedenen Arten des „authentic assessment“³⁶, zu denen etwa Schüler-Portfolios zählen.

Abbildung 2: Kontinuum der Assessment-Verfahren



matching	true	multiple-	fill-in-	conscon-	individual/	interviews,	demonstrations,
	false	choice	the-blank	structured	group	role-	essays,
				response,	projects,	playing	art productions,
				short-answer,	logs,		portfolios
				extended	journals		
				response			

Quelle: Hansche 1998, S. 24

Während sich einerseits die Diskussion bezüglich höherer Standards und Formen des Testings thematisch ausdehnte, hat sich andererseits die Debatte über die relevanten Formen des Assessments stark verengt. In der Kritik wird daher festgehalten: „The political climate now favoring test-based assessment and accountability has obscured the nature of educational assessment“ (Kornhaber 2004, S. 50). Die politische Orientierung an Testverfahren als der bevorzugten Methode des Assessment beruht teilweise auf der Annahme, dass dadurch pädagogische

³⁶ *Authentic assessment* beabsichtigt die Fähigkeiten, Kenntnisse und Leistungen, die tatsächlich in einem Fach entwickelt werden können, besser wiederzugeben, als sie im Kontext eines Testes erfasst werden (Kornhaber 2004).

dagogische Problembereiche erfasst werden, die ansonsten zu wenig Beachtung finden. Dazu zählen niedrige Standards, schwache Motivation, minimale Erwartungen, unangemessene Lehrpläne und unzureichender Unterricht. Ein testbasiertes Assessment bietet daneben noch einige andere Vorteile:

- Das Verfahren ist relativ kostengünstig, vergleicht man es mit strukturellen Veränderungen wie Reduktion der Klassengröße, Intensivierung der Ausbildung oder beruflicher Weiterbildung.
- Assessment und Testing können extern vergeben werden. Die zentrale Organisation und Durchführung auf der Ebene von Distrikten oder Staaten ist einfacher und unkomplizierter als eine individuelle Leistungserhebung und -erfassung der jeweiligen Schule.
- Änderungen im Test- bzw. Assessment-Verfahren können schnell verwirklicht werden.
- Die Ergebnisse von Tests sind sichtbar und können der Öffentlichkeit vermittelt werden (Linn 2000).

Nach dem Erlass des NCLB-Gesetzes war eine Anpassung der Verfahren zur Leistungsbeurteilung in den Bundesstaaten erforderlich. Dazu gehören neben dem oben erwähnten *Adequate Yearly Progress* (AYP) die jedes zweite Jahr verpflichtende Teilnahme am staatlich organisierten *National Assessment of Educational Progress* (NAEP) in Lesen und Mathematik in der vierten und achten Jahrgangsstufe. Tests gibt es seit 2004 in allen Bundesstaaten. Bereits im Schuljahr 2003/04 hatten 20 Staaten und der Distrikt Columbia in Englisch und Mathematik Tests in allen geforderten Jahrgangsstufen durchgeführt. Auch haben 15 der Bundesstaaten die gleichen standardbasierten Test-Programme in jeder Jahrgangsstufe verwendet, um dadurch einen Vergleich der Ergebnisse zwischen den Klassen zu ermöglichen.

Insgesamt sind sowohl für den nationalen NAEP-Test wie auch für die staatlichen Tests in den vergangenen Jahren Leistungszunahmen konstatiert worden. Auch der Leistungsunterschied zwischen Weißen und Schwarzen sowie Hispanics zeigt zumindest Tendenzen hin zu einer Abnahme. Gleichzeitig werden auch negative Auswirkungen auf die Curricula und den Unterricht festgestellt (vgl. Müller/Silver 2006), beispielsweise Fälle von Ausschlüssen gewisser Schülergruppen, etwa solchen mit limitierten Englischkenntnissen. In bestimmten Staaten wie Texas oder North Carolina wird die Leistungsverbesserung mit der erfolgreichen Etablierung von Bildungsstandards (Content Standards) in Verbindung gebracht (Finn 2002, S. 31). Hier besteht offenbar ein po-

sitiver Zusammenhang zwischen Tests, Standards und Leistungsverhalten der Schülerinnen und Schüler.

Allerdings wird die Entwicklung verschieden gedeutet, je nachdem, welchem politischen Lager die Autoren angehören und welche Daten sie konsultieren. Es erweist sich insgesamt als schwierig, Outcomes bei Schülerinnen und Schülern und Lehrpersonen auf die Wirkung der Einführung von Standards zurückzuführen. „For example, it is impossible to know from some studies of systemic reform to what degree the outcomes are due to the standards framework, the curriculum, the assessment, the accountability policy, the professional development, the school governance, or more than one of these elements“ (Lauer/Snow/Martin-Glenn et al. 2005, S. 99).

Wenn Bundesstaaten verglichen werden und die Entwicklung des Leistungsverhaltens innerhalb einer Dekade zur Darstellung kommt, lässt sich etwa für Texas ein positiver Schluss ziehen (Winick/Kress 2004); wenn der in Texas entwickelte Test³⁷ genauer untersucht wird, kann ebendas bezweifelt werden, weil der Test einerseits viele Ausnahmen zulässt und die Schulen andererseits einen hohen Aufwand betreiben, damit ihre Schülerinnen und Schüler den Test bestehen, ohne dass sicher wäre, dass sich ihre Lernleistungen insgesamt verbesserten (Haney 2000).

Die *Education Week* legt einen jährlich erscheinenden Bericht *Quality Counts* vor, der die Bildungsarbeit in allen amerikanischen Bundesstaaten analysiert. Im Januar 2006 veröffentlichte die Zeitschrift ihre zehnte Untersuchung, die sich detailliert damit auseinandersetzt, ob die Bundesstaaten Strategien und Methoden hin zu einem standardbasierten Unterricht eingeführt haben und ob es ihnen gelungen ist, die schulischen Leistungen der Schülerinnen und Schüler zu verbessern (Education Week 2006a). Die Analyse kommt sowohl zu ermutigenden als auch zu ernüchternden Ergebnissen. In den 4. und 8. Jahrgangsstufen stehen bundesstaatliche Bemühungen zur Entwicklung von Standards, Tests und Accountability-Systemen in einem positiven Zusammenhang mit den sich verbessernden Schülerleistungen im Untersuchungszeitraum von 1996 bis 2005. Gesteigert haben sich insbesondere die Leistungen im Fach Mathematik und die Leistungen von Schülerinnen und Schülern der einkommensschwachen Minderheiten. Allerdings ließ sich kein positiver Zusammenhang zwischen der Verbesserung der Schülerleistungen und der Qualität des Unterrichts bzw. des Lehrpersonals feststellen. Zudem sind die Leistungssteigerungen in anderen Fä-

³⁷ Texas Assessment of Academic Skills (TAAS).

chern kaum signifikant, auch weil diese Fächer Gefahr laufen, mit den Tests abgewertet zu werden.

Ein Review von Studien (Lauer/Snow/Martin-Glenn et al. 2005), die seit 1995 zur Frage der Auswirkung von Standards und standardbasierten Curricula auf Unterricht und Schülerlernen durchgeführt worden sind, ergibt tendenziell positive Effekte bezüglich der Schülerleistungen. Gemischt sind die Ergebnisse hinsichtlich des Einflusses auf die Unterrichtspraxis und -methodik. Bedingung für einen positiven Zusammenhang ist hier die Verfügbarkeit systemischer Unterstützung, die mit den Curricula im Sinne des Alignment korrespondiert. Interessant ist, dass Lehrpersonen oftmals angaben, ihren Unterricht stark an den Standards auszurichten, wenngleich dies häufig mit Unterrichtsbeobachtungen nicht bestätigt wurde. Der stärkste Einfluss auf den Unterricht und das Lernen geht allerdings, wie die Studien zeigen, von den Tests aus. „The overall conclusion is that standards-based policies influence teaching and student learning in K-12 classrooms, but the nature of these influences depends on how standards-based policies are perceived and implemented by teachers“ (S. viii).

Die standardbasierten Testmethoden haben, wie gezeigt, zu ausgiebigen Diskussionen und scharfen Kontroversen geführt. Die Befürworter gehen davon aus, dass testbasierte staatliche Assessments die Leistungserwartungen erhöhen und sicherstellen, dass alle Schülerinnen und Schüler gemäß den gleichen Standards unterrichtet werden. Sie leisten damit einen Beitrag zur Chancengleichheit (Gandal/McGiffert 2003). Die Kritiker argumentieren, dass das Lernen der Schülerinnen und Schüler auf die Inhalte verengt wird, die testrelevant sind, und dass die getesteten Inhalte nur einen kleinen Ausschnitt von dem darstellen, was die Lernenden wissen sollten. Hinzu kommt, dass die Testergebnisse in den Accountability-Systemen meist mit „high stakes“ verknüpft werden. Die Kernfrage ist dann, wozu die Tests und ihre Resultate verwendet werden sollen.

„High stakes“ bezeichnen keine besondere Form oder Art des Testens, sondern die mit den Resultaten verbundenen Konsequenzen. Der amerikanische Ausdruck „stake“ lässt sich einerseits mit dem deutschen Wort „Grenzpfehl“ übersetzen, andererseits mit der Idee des Spiel- oder Wetteinsatzes. High-Stakes-Testing hat ernsthafte Konsequenzen für die Person, die getestet wird, etwa wenn es sich um einen finalen Test handelt, der nicht wiederholt werden kann und der über Promotionen oder Selektionen entscheidet. Alle Tests sind „high stakes“, die dazu verwendet werden, Urteile über Zugänge, Graduierungen oder Promotionen zu fundieren oder gar festzulegen. Sie dienen nicht primär der Förderung, sondern der Zu- und Abweisung innerhalb eines selektiven Systems.

Selektionsentscheide gibt es in jedem Bildungssystem, wenngleich nicht basiert auf nationalen Tests, die gleiche Standards voraussetzen, sondern auf Urteilen von Lehrkräften. Kontrovers ist, ob eine High-Stakes-Testpolitik zu besseren Testergebnissen und damit zusammenhängend zu besseren Lernleistungen führt. Diese Frage wird hitzig diskutiert, ohne eine abschließende Antwort gefunden zu haben. Amrein/Berliner (2002) vergleichen die Resultate der NAEP-Tests mit denen der Tests verschiedener Bundesstaaten und ziehen ein negatives Fazit:

„At the present time, there is no compelling evidence from a set of states with high-stakes testing policies that those policies result in transfer to the broader domains of knowledge and skills for which high stakes scores *must* be indicators. Because of this the high-stakes tests being used today do not, as a general rule, appear valid as indicators of genuine learning, of the types of learning that approach the American ideal of what an educated person knows and can do. Moreover, as predicted by the Heisenberg University Principle, data from high-stakes testing programs too often appear distorted and corrupted“ (Amrein/Berliner 2002).

Rosenshine (2003) untersucht einen vergleichbaren Datensatz und fügt aber eine Kontrollgruppe hinzu, also Bundesstaaten, die den gleichen Test *ohne* High-Stakes-Konsequenzen durchführen ließen. Diese Staaten fallen gegenüber denen *mit* High-Stakes-Konsequenzen deutlich zurück, und zwar in allen drei NAEP-Tests über eine Periode von vier Jahren. Green/Winters/Forster (2004) zeigen, vor allem am Beispiel des Bundesstaates Florida, dass Tests verlässlich Auskunft geben über die Fortschritte im Leistungsverhalten der Schülerinnen und Schüler, ohne dass die mit dem High-Stakes-Anspruch verbundenen Konsequenzen darauf einen negativen Einfluss nehmen würden. Cunningham (2004) beschreibt das Scheitern eines Accountability-Systems im Bundesstaat Kentucky. Der Einfluss reformpädagogischer Ideale habe eine klare Ausrichtung an Exzellenz-Standards verhindert und dazu geführt, dass halbherzig, methodisch schwach und ohne Konsequenzen getestet wurde. Die Lehrkräfte aber werden nur dann effektive Unterrichtsstrategien anwenden, wenn sie sich ernsthaften Konsequenzen gegenübersehen (ebd., S. 298).

Häufiger werden die nicht intendierten Folgen selektiver Tests thematisiert und entsprechende Befürchtungen artikuliert (Jones/Jones/Hargrove 2003; Thomas 2005 und diverse andere). Dazu zählen die Verengung des curricularen Angebots, das die Schülerinnen und Schüler ernst nehmen, auf die testrelevanten Fächer, der faktische Ausschluss bestimmter Gruppen von Schülerinnen

und Schülern, etwa solchen mit besonderen Lernbedürfnissen, oder auch die Reduktion der Lernkultur auf „rote memorization“. Auf die Bedeutung nicht intendierter Folgen von externen Testprogrammen ist schon früh verwiesen worden (etwa Smith/Rottenberg 1991), ohne dass bis heute klar wäre, ob sie vermeidbar sind, und wenn nicht, wie damit umgegangen werden soll.

Ein Schlüssel zu dieser Kritik ist bezeichnet mit dem Ausdruck „*Teaching to the Test*“. Schulen, so die Kritik, würden viel Zeit für die Vorbereitung auf die Tests vergeuden und einen hohen Aufwand betreiben, ein mehr oder weniger primitives Lernen zu vermitteln.

„It is no surprise that attaching high stakes to the test results in an accountability system leads to a narrowing of the instructional focus of teachers and principals. There is considerable evidence that teachers place greater evidence on material that is covered on a high-stakes test than they do on other material“ (Linn 2003, S. 4; vgl. unter anderem Mehrens/Kaminski 1989; Shepard 1990; Taylor/Shepard/Kinner et al. 2001; Stecher/Hamilton 2002).

Eine Konzentration des Unterrichts auf die inhaltliche Ausrichtung der Tests hat allerdings nicht nur negative Aspekte. Vorteilhaft ist die curriculare Verdichtung dann, wenn sich damit qualitativ hochstehende Content Standards verbinden, an denen sich die Tests ausrichten und über die ein Konsens besteht, ohne damit die gesamte Breite eines Fachs oder einer Lerndomäne abzudecken. Dagegen hat eine *ausschließliche* Ausrichtung des Unterrichts an den Testitems und den damit abgedeckten Inhalten negative Auswirkungen. Mit der Fokussierung des Unterrichtes lediglich auf die Inhalte, die getestet werden, geht ein reduziertes Verständnis der zu unterrichtenden Fächer insgesamt einher (Linn 2003). Die Autoren von McREL fassen die Vor- und Nachteile des *Teaching to the Test* wie folgt zusammen:

Tabelle 1: Vor- und Nachteile des *Teaching to the Test*

For	Against
Teaching to test and teaching the test are two different things. If a teacher is teaching to the test and doing good teaching that enhances learning, what's wrong with that?	In response to high-stakes tests, many schools substitute test-prep materials for the regular curriculum – mostly in minority students' classroom. These test-prep materials have little, if any, value beyond practicing for the tests. Scores go up, but the quality of education goes down.

If the test measures what students are supposed to learn, then teaching to the test is the best thing to do.	Accountability is based on multiple-choice, standardized testing that requires memorization and regurgitation. These tests do not cover large chunks of state standards, and things that aren't tested probably aren't taught.
Teaching to the test is fine, as long as students are learning good, valuable skills and important content knowledge.	Many states use "off-the-shelf" tests to hold schools accountable without considering whether the test measures the state's most important academic standards.
Teachers should be teaching to the test if the test requires students to use higher-order thinking skills and problem solving.	Teaching to the test throttles teacher creativity and encourages cheating.

Quelle: Marzano/Mayeski/Dean 2000

Unabhängig von der Frage, wie stark letzten Endes das *Teaching to the Test* den Unterricht bestimmen wird – testbasierte Assessment- und Accountability-Systeme werden in der amerikanischen Entwicklung vor dem Hintergrund des NCLB-Gesetzes weiterhin eine einflussreiche Rolle spielen. Die angemessene Anwendung des Testings ist dabei ein Schlüsselproblem. Ohne Nutzung der Tests sind Rückschlüsse auf Schülerleistungen und die Leistungen von Schulen gegenstandslos. Dementsprechend hat die *American Educational Research Association* (AERA) (American Educational Research Association/American Psychological Association/National Council on Measurement in Education 1999), unterstützt durch das *National Research Council* (NRC), verschiedene Methoden und Verfahren entwickelt, die Assessments berücksichtigen sollten (Heubert/ Hauser 1999). Dazu gehört etwa das Kriterium, dass Entscheidungen, die Konsequenzen für die Schülerin oder den Schüler haben, nicht auf einem Testergebnis allein beruhen sollten.

Das entspricht der Position des *National Council of Teachers of Mathematics* vom Januar 2006. „Tests, after all, are snapshots that capture one event in one context rather than a wide array of events in multiple contexts. The results of large-scale tests must be balanced against a broader sampling of student performance“ (NCTM 2006, S. 1). Genau diese Ausweitung wird von der Gegenseite bestritten, die davon ausgeht, dass einzig und allein Large-Scale-Tests verlässliche Aussagen über die Leistungen der Schülerinnen und Schüler machen können. Dagegen schreibt das NCTM aus der Sicht des Schulfaches:

„Valid, reliable large-scale assessments are useful and important tools for examining students’ progress and making a variety of comparisons. However, because they may not measure the full range of important mathematics, they must be combined with a more complete sampling of student performance. This sampling might include classwork, tests, quizzes, observations, projects, and interviews. Such a collection of both formal and informal assessments can provide teachers and others with a more complete picture of student performance. By contrast, placing too much emphasis on a single test or on testing can undermine the quality of education and jeopardize equality of opportunity“ (ebd.).

Die Bedeutsamkeit der Frage der angewandten Assessments und Testmethoden im Hinblick auf Unterrichtspraxis und Schülerlernen lässt sich anhand der großen Zahl entsprechender Untersuchungen in den USA gut nachweisen (vgl. Lauer/Snow/Martin-Glenn et al. 2005). Daraus ergibt sich die Forderung von selbst, dass standardbasierte Assessments immer im Licht der Frage beurteilt und ausgewählt werden, ob sie potenziell dazu beitragen, dass Lehrpersonen wünschenswerte curriculare und pädagogische Entscheide fällen (ebd.).

3.1.4 Erfahrungen mit der Implementation von Standards

Seit der Publikation von *A Nation at Risk* haben Beobachter drei Wellen der Schulreform registriert. „First-wave reforms called largely for systemic changes, such as increasing standards and regulations, and resulted in increased teachers’ salaries, increased core requirements, and a lengthened school day and year“. Diese Maßnahmen wurden in der Folge wegen ihres Top-down-Charakters kritisiert, und es wurde moniert, dass sie lediglich zu einer Intensivierung³⁸ des vorhandenen Systems führten, jedoch keine zusätzliche Kapazität produzieren würden. Darauf wurde in einer zweiten Phase reagiert: „The second wave focused primarily on broadening and deepening the relationship between schools and families, addressing the needs of special groups of students, attracting and retaining effective teachers, upgrading teacher education, and restructuring teachers’ role to make them more professional“ (Desimone 2002, S. 433).

Untersucht man die Ebene der Schuldistrikte, dann haben sich diese Reformen weder auf Schulorganisation noch auf den Unterricht der Lehrpersonen maßgeblich ausgewirkt (Cuban 1984). Als Antwort darauf wurde mit einer dritten Welle der Ansatz der *Comprehensive Schoolwide Reform* (CSR) intensiver verfolgt.

³⁸ Zur These der Intensifikation im Bildungswesen vgl. z.B. Covington (1996).

Im Gegensatz zu den früheren Strategien, welche gewöhnlich auf eine bestimmte Schülerpopulation innerhalb der Schule und/oder spezielle Fächer, Programme oder Methoden gerichtet waren, fokussiert CSR die Reform der einzelnen Schule als Ganze. Mit und neben dem CSR gehört das Konzept der *Standards-based Systemic School Reform* in die dritte Phase der Entwicklung. Systemic school reform unterscheidet sich gemäß Lusi (1997)

„from the reform attempts of the 1980s in at least two important ways. First, systemic school reform strives to reform the education system as a system; it works for coherence across the system’s component policies, something that the piecemeal reform of the past did not achieve. Second, systemic school reform explicitly strives to support school-site efforts at redesigning teaching and learning with the goal that all students will learn ambitious content knowledge and higher-order skills. It is insufficient to promulgate mandates such as increased graduation requirements from the ‚top‘ of the education system (state). The ‚bottom‘ of the system (schools and districts) must be supported and activated to transform teaching and learning“ (ebd., S. 6).

Als Konzept, das in der dritten Reformphase und vor dem Hintergrund eines anhaltenden achievement gap zunehmend zur Diskussion steht, ist auf *capacity*³⁹ zu verweisen. Es steht auch für die Feststellung, dass eine anhaltende Verbesserung von Schülerleistungen weiterer Maßnahmen bedarf, als lediglich der Einführung von Standards, Tests und Accountability; vor diesem Hintergrund ist die Rede von *capacity-building*.

„Capacity“ wird sowohl von Vertreterinnen und Vertretern systemischer Reform wie von school-by-school change (etwa mittels Comprehensive School Reform) thematisiert. Im ersteren Fall zielt capacity-building neben der Adaption von Standards, der Entwicklung von Curricula und Assessments sowie einer Veränderung der Steuerung auf das Alignment von Reformpolitiken und -strategien verschiedener Ebenen. Letztere meinen in erster Linie die Bildung von Lerngemeinschaften (learning communities) und Möglichkeiten für Lehre-

³⁹ Newmann/King/Rigdon (1997) geben folgende Definition: „capacity: the degree to which the human, technical, and social resources of an organization are organized into an effective *collective* enterprise“ (S. 47). Ingram/Colby (1998): „In a school with a high organizational capacity, all resources are directed toward a clear, shared purpose for student learning. Teacher knowledge, leadership and a school’s technical and financial resources are effectively integrated to accomplish the school’s vision for student achievement“ (S. 5).

rinnen und Lehrer, ihr Wissen und praktisches Können zu entwickeln und auszutauschen (Corcoran/Goertz 1995).

Neben Kompetenzen von Lehrkräften beinhaltet capacity auf der Ebene Unterricht für Cohen/Ball (1996) anforderungsreiches Unterrichtsmaterial sowie Schülerinnen und Schüler und Lehrpersonen, die motiviert sind zu lernen. Die Autoren folgen in ihrer Konzeption Argumenten aus der Kognitionspsychologie, die besagen, dass das Verhalten von Akteuren stark beeinflusst ist von der umgebenden Situation; Ressourcen und Personen im Klassenzimmer sind deshalb bei der Umsetzung von Reformzielen bedeutende Elemente von capacity.

Goertz/Floden/O'Day (1995) beziehen capacity auf die das Klassenzimmer umgebenden Kontexte, Strukturen und Organisationen (Schule, Distrikt, Staat) und attestieren diesen eine herausragende Bedeutung für die erfolgreiche Implementation standardbasierter Reformen. Massell (1998) hat, ausgehend von diesen beiden einschlägigen Konzepten, sieben Bereiche von capacity abgeleitet, die für die Verbesserung von Unterricht und Lernen von Bedeutung sind:

Classroom-level capacities:

- teachers' knowledge, skills, and dispositions,
- students' motivation and readiness to learn,
- curriculum material for students and teachers,

Organisational-level capacities (school, district, state):

- number and kinds of people supporting the classroom,
- number and quality of social relationship,
- material (non-human resources),
- organization and allocation of school and district resources.

Wie oben ausgeführt ist sowohl der Ansatz der Systemic Reform wie derjenige der Comprehensive Schoolwide Reform Ausdruck einer dritten Reformphase innerhalb der US-amerikanischen Anstrengung zur Hebung der Schülerstandards seit den 80er-Jahren; beide haben, bei unterschiedlicher Gewichtung, den Anspruch, Top-down-Reformen mit Bottom-up-Reformen zu vereinigen. Die Implementation von Comprehensive School Reform Designs in einzelnen Schulen hat innerhalb des Reformansatzes des Comprehensive School Improvement eine systemische Ausrichtung. Es handelt sich um eine breite Strategie mit dem Ziel der Hebung der Schülerleistungen über eine kohärente Kopplung der Implementation von Modellen der Comprehensive School Reform (CSR) mit der Politik und Praxis auf der Distrikt- und Staatsebene und umfasst

weitere Elemente wie school leadership und governance, accountability, Evaluation, Ressourcenallokation und Einbindung der community.

Die Verbreitung von Comprehensive School Reform wurde durch eine Reihe nationaler Entwicklungen befördert: die Bewegung in Richtung systemischer und standardbasierter Reform, eine neue Gesetzgebung auf Bundesebene, die es erlaubt, Title-I-Gelder für die Implementation von schoolwide programs in sogenannten high-poverty schools zu verwenden, und eine CSR-Gesetzgebung, die hunderte von Millionen Dollar zur Verfügung stellt, um Schulen die Adaption von extern entwickelten Reformmodellen zu ermöglichen. Zusätzlich von Einfluss waren Ergebnisse der effective schools research (Teddlie/Reynolds 2000).

Bezüglich der Implementation und des scaling-up von sogenannten whole-school designs im Rahmen der Comprehensive School Reform (CSR) liegen u.a. mit der Initiative New American Schools (NAS) der unter Präsident George H. Bush gegründeten privaten New American Schools Development Corporation und deren Evaluation mittels RAND⁴⁰-Untersuchungen Ergebnisse vor (vgl. dazu ausführlich Berner/Stolz 2006). Im Folgenden sollen lediglich Erfahrungen und Forschungsergebnisse zur Implementation von Standards im Rahmen des Systemic-Reform-Ansatzes wiedergegeben werden. Als eine herausragende Initiative ist in diesem Zusammenhang auf die Systemic State Initiative (SSI) zu verweisen, deren Implementation in verschiedenen Staaten gut untersucht ist.

Gemäß Knapp (1997) hat sich die Idee der Systemic Reform in den späten 80er-Jahren artikuliert, so etwa in Publikationen von Smith/O'Day (1991, 1993) oder der National Governors' Association (1989). Chatterji (2002) stellt fest:

„Specifically, as the federal government became increasingly involved in setting the education reform agenda on a national level, the term standards-based reforms was used to refer to the ways in which individual states responded to the push for higher standards and school accountability“ (S. 346).

Im Zentrum verschiedener systemischer Initiativen steht die signifikante Verbesserung der Qualität von Curricula und Unterricht. Dieses Ziel wird hauptsächlich verfolgt mit der Formulierung von Standards und Curriculum Frame-

⁴⁰ RAND ist eine non-profit Forschungsorganisation, die im Auftrag von Politik und Entscheidungsträgern im öffentlichen und privaten Sektor wissenschaftliche Analysen bereitstellt.

works, einer kohärenten Reformpolitik (Alignment) und einer Kombination von Zentralisierung mit Dezentralisierung. Staatliche Zielvorgaben sind zentral zu verstehen, sie gelten für alle Schulen in einem Zielbereich. Dezentrale Maßnahmen setzen eine Erhöhung der Flexibilität voraus sowie eine Verlagerung der Verantwortung an lokale Einheiten, also Distrikte und einzelne Schulen. Folgende Elemente stehen gemäß Smith/O'Day (1993) im Zentrum systemischer Reform:

- (1) anspruchsvolle Curricula, fordernder Unterricht, hohe Leistungserwartungen an *alle* Schülerinnen und Schüler;
- (2) eine klare Abstimmung (Alignment) von Rahmencurricula (Curriculum Frameworks), Standards, Unterrichtsmaterial, Lehreraus- und -weiterbildung und Leistungsbeurteilung;
- (3) Rechenschaftslegung und Verantwortung gemäß Leistungsdaten (performance-based accountability);
- (4) Professionalisierung der Lehrkräfte;
- (5) ein abgestimmtes Konzept von Standards (input- und output-standards: performance standards, resource standards und practice standards).

Gemäß dem Report der American Federation of Teachers (AFT) von 1995, *Making Standards Matter*, befassten sich schon Mitte der neunziger Jahre nahezu alle amerikanischen Bundesstaaten in irgendeiner Weise mit der Durchführung standardbasierter Reformen. Bereits 1991 hatte die National Science Foundation (NSF) das Programm Statewide Systemic Initiative (SSI) ins Leben gerufen, welches die 25 partizipierenden Staaten und Puerto Rico bei der Durchführung systemischer Reformen unterstützte. SSI in verschiedenen Staaten ist mit zahlreichen Untersuchungen gut dokumentiert und evaluiert.⁴¹ Die Strategien differierten in ihren Schwerpunktsetzungen je nach Staat, weisen jedoch gemeinsame Elemente auf. Auf staatlicher Ebene bestand überall eine Hauptaufgabe darin, Konsens betreffend Zielsetzung von SSI unter den verschiedenen beteiligten Gruppen herzustellen. Strategien kombinierten folgende Elemente: Herstellung eines Alignment zwischen SSI und staatlicher Politik, Entwicklung und Dissemination von neuen Inhalten und Materialien, Unterstützung für die Lehrerweiterbildung, Aufbau von Infrastrukturen, Finanzierung von lokalen systemischen Anstrengungen, Versuche, die Lehrerbildung zu reformieren,

⁴¹ Die Evaluationen wurden von SRI International durchgeführt und können abgerufen werden unter: <http://www.sri.com/policy/cep/mst/ssi.html> (Stand 14.01.06).

Kampagnen zur Mobilisierung der professionellen und öffentlichen Unterstützung.

Massell (1998) hat ausgehend von einem Capacity-Konzept systemische Strategien der Implementation (capacity-building) in acht amerikanischen Bundesstaaten untersucht, nämlich Kalifornien, Colorado, Florida, Kentucky, Maryland, Michigan, Minnesota und Texas. Verschiedene dieser Bundesstaaten nahmen am SSI teil. Die Staaten zeigten folgende gemeinsame Umsetzungen:

- *Aufbau einer Infrastruktur für Weiterbildung und technische Unterstützung der Schulen und Lehrkräfte.* Als Institutionen einbezogen wurden regionale Weiterbildungs- und Servicezentren, staatliche Berufs- und Fachorganisationen. Dabei wurden professionelle Netzwerke für Lehrerinnen und Lehrer, Schulen und Distrikte etabliert. Netzwerke zielten entweder auf die Vermittlung von Wissen und Können oder sie wurden verwendet zur Dissemination spezifischer Produkte für den Unterricht. In diesem Zusammenhang wurden auch Universitäten ermutigt, den Schulen gezielt Unterstützung anzubieten. In größeren Distrikten setzte man auf eigene Weiterbildungsangebote. Angebote von Seiten des Staates waren gewöhnlich auf die bedürftigsten Schulen gerichtet.
- *Setzung von Standards für Weiterbildungsangebote.* Die Personalentwicklung in den Schulen orientiert sich neu an Vorgaben, die vom Bedarf ausgehen und der Verbesserung der Qualität dienen. Daneben wurden Akkreditierungsstandards für Programme der Lehrerausbildung und für Lizenzierungen revidiert oder neu entwickelt.
- *Entwicklung von Curriculummaterial.* Damit sollte die Lücke hinsichtlich Anleitung des Unterrichts, die im Zuge der Reformpolitik bestehend aus zentraler Standardsetzung und dezentraler Verantwortung entstanden ist, geschlossen werden. Staaten entwickelten spezifischere Curriculum Frameworks oder anderes zusätzliches Material und Beispiele, wie die neuen Standards im Unterricht anzuwenden sind. Staaten etablierten auch Datenbanken mit Materialien und Unterrichtswerkzeugen, außerdem ermutigten sie Partnerschaften mit nationalen Curriculumprojekten. Traditionell besitzen weder die Lehrpersonen noch das Distriktpersonal vor Ort die notwendige Expertise, um selber Curricula und Programme zu entwickeln. Daraus resultierte eine mehr oder weniger starke Abhängigkeit von kommerziellen Lehrmitteln und Testanbietern. Die Staaten waren also gefordert, eine aktivere Rolle zu übernehmen, um die Schulen und Distrikte bei der Entwicklung standardkompatibler Curriculummaterialien zu unterstützen. Das Ausmaß der staatlichen Aktivitäten differierte zwischen stärker lokal und zentral kontrollierten Staa-

ten. Generell überließen die Staaten die primäre Verantwortung für die Curricula und Unterrichtsmaterialien den Distrikten.

- *School Improvement Planning*. Die Schulen sollten ihre Bedürfnisse selber im Licht der staatlichen Reformziele identifizieren und die finanziellen Mittel, Zeit, persönliche und berufliche Weiterbildung selbstverantwortlich allozieren.

Knapp (1997) hat systemische Reformen auf staatlicher und Distriktebene, insbesondere solche, die im Rahmen von SSI von der National Science Foundation unterstützt wurden, in einer Metaanalyse hinsichtlich ihrer Auswirkungen auf den naturwissenschaftlichen und Mathematikunterricht evaluiert. Kalifornien ist dabei ein instruktives Beispiel. Der Bundesstaat hat an der Statewide Systemic Initiative teilgenommen,⁴² verfügt jedoch zusätzlich über längere Erfahrungen, die auf die Implementation neuer Curriculum Frameworks (Standards)⁴³ in den achtziger Jahren zurückgehen.

Bezüglich der kurzfristigen Dynamiken und Effekte der California Mathematics Framework Reforms stützt sich Knapp (1997) auf qualitative Untersuchungen, die auf der Ebene der Distrikte und Schulen (Marsh/Odden 1991) angesiedelt sind, respektive Untersuchungen zu einzelnen Lehrpersonen (Guthrie 1990) untersuchen. Die beiden Studien zeigen, über welche verschiedenen Wege Reformen das Schulzimmer erreichten oder nicht erreichten. Marsh/Odden (1991) hatten die lokale Politik und ihr System im Blick und konnten substantielle Erfolge der Reform feststellen; demgegenüber sind die Studien in Guthrie (1990)⁴⁴ pessimistischer. Zusammen geben sie folgendes Bild: Die Curriculum-Reformpolitik war in den ersten Jahren erfolgreich, insofern verschiedene Akteure, etwa über regionale Netzwerke, in Distrikten und in Schulen aktiviert worden sind. Dies gilt besonders für „reife“ Distrikte, d.h. solche, die genügend Kapazität und Interesse zeigten, die Reform an eine große

⁴² Eine Evaluation zu SSI in Kalifornien liegt vor mit Shields/Marsh/Marder et al. 1998.

⁴³ In Kalifornien stehen „Frameworks“ für „Standards“. Smith/O’Day (1993, S. 300) verwenden „curriculum frameworks“ für state-level Dokumente und „content standards“ für national-level Dokumente, wie etwa die NCTM-Standards.

⁴⁴ Fünf Fallbeispiele von Lehrerinnen und Lehrern, publiziert in: Educational Evaluation and Policy Analysis, 12 (1990), No. 3: Ball, D.L.: Reflections and deflections of policy: The case of Carol Turner; Cohen, D.K.: A revolution in one classroom: The case of Mrs. Oublier; Peterson, P.L.: Doing more in the same amount of time: Cathy Swift; Wiemers, N.J.: Transformation and accommodation: The case of Joe Scott; Wilson, S.M.: A conflict of interests: The case of Mark Black.

Zahl von Lehrkräften zu vermitteln. Anstrengungen, die Politik in die Praxis umzusetzen, waren erfolgreich, insofern leicht und schnell zu ändernde Aspekte adressiert wurden. Grenzen ergaben sich auf der Verwaltungs- und Schulebene dort, wo es um das tiefere Verständnis für den damit bezweckten neuen (konstruktivistischen) Unterricht, die Änderung von Überzeugungen (beliefs) bezüglich mathematischen Fachwissens und die Bildung von Kapazitäten, dieses in neuen Unterricht zu transformieren, ging.

Die relativ schnelle Adoption der Politik in den von Marsh/Odden (1991) untersuchten Distrikten und Schulen kann auf die dort bereits vorhandenen Kapazitäten, die Bildung von Koalitionen, die der Reform positiv gegenüberstanden, die Existenz stützender staatlicher Initiativen wie ein Mentoring-Programm und Subventionen für Schulreformprogramme sowie die kreative Anwerbung und Verwendung von Ressourcen zurückgeführt werden. Mit ihrem Ansatz können sie aber wenig aussagen über die tatsächliche Umsetzung der Reformen im Unterricht. Hierzu müssen die Studien zu einzelnen Lehrkräften in Guthrie (1990) herangezogen werden. Sie zeigen bezogen auf die intendierten Effekte und die tatsächlichen Konsequenzen eine komplexere Geschichte. In den geschilderten Fällen nahmen die Lehrkräfte einzelne Aspekte der Reform auf, während sie in vielem beim bereits Bekannten blieben. Dieses Verhalten kann auf ein begrenztes Verständnis dessen zurückgeführt werden, was die neuen Frameworks an verändertem Unterricht verlangen. Dieser Mangel resultierte einerseits aus einem zu geringen Maß an Unterstützung durch die Distrikte. Andererseits hatten auch die Curriculumspezialisten ein eher oberflächliches Verständnis für die Reform. Damit lässt sich erklären, warum trotz Abstimmung zwischen dem Curriculum und den Unterrichtsmaterialien die Wirkung der Reform begrenzt blieb.

Hinsichtlich der längerfristigen Entwicklung liegen durch verschiedene Untersuchungen innerhalb der *Educational Policy and Practice Study* (EPPS) Ergebnisse vor. Die EPPS-Forscher haben die einzelnen Lehrkräfte über weitere Jahre hinweg beobachtet. Deren Lernprozesse setzten sich fort, jedoch nicht notwendig in der Weise, die von der Reformpolitik vorgesehen war. Verschiedene Lehrkräfte derselben Schule verstanden die an sie gestellten Erwartungen unterschiedlich (vgl. auch Grant/Peterson/Shojgreen-Downer 1996). Das Lernen der Lehrkräfte war vor allem geleitet durch ihr mathematisches Fachverständnis und -wissen. Hinderlich war die Existenz eines Testprogramms, das nicht mit den Frameworks übereinstimmte. Ihr Lernen und der Transfer des Gelernten in den Unterricht hingen ab von den unmittelbaren Arbeitsbedingungen, den Peers, Lern- und Weiterbildungsgelegenheiten, Zugang zu Anleitungen, etwa über Unterrichtsdemonstrationen, und Support. Eine zentrale Rolle kam dem

pädagogischen Personal des Distrikts zu, das zwischen den verschiedenen Ebenen vermittelte und dabei die Reform in Teilen neu interpretierte.

Mit O'Day (1995) und Page (1995) liegen Untersuchungen zur systemischen Reform in Kalifornien und der Implementation von Frameworks vor, die die unterschiedlichen Möglichkeiten von High-Capacity und Lower-Capacity-Schulen und Distrikten zeigen. In High-Capacity-Distrikten übten die Frameworks gemäß den Ergebnissen von O'Day (1995) einen sichtbaren Einfluss auf Curricula und Unterricht aus. Lehrkräfte nahmen extensiv an verschiedenen Weiterbildungskursen teil, die Bezug nahmen auf die Frameworks. Sowohl die Frameworks wie auch die daran gekoppelten Performance-Based Assessments fanden hier Akzeptanz bei den Lehrkräften. Lerngemeinschaften wurden initiiert, und die Schulen unterstützten die Reform mit Umstrukturierungen in ihrer Organisation. Lehrpersonen erhielten Gelegenheit, an der Bewertung der Schülerinnen und Schüler über Assessments mitzuwirken, bei der Auswahl von Lehrbüchern mitzubestimmen etc.

Pessimistischer sehen die Ergebnisse von Page (1995) bezüglich der Implementierung eines stärker integrierten Naturwissenschaftsunterrichts in gewöhnlichen oder Lower-Capacity High Schools aus. Diese Distrikte investierten wenige Ressourcen in die Reform bzw. reservierten diese für Tätigkeiten auf der administrativen Ebene. Dieser Mangel wirkte sich auf die Qualität der curricularen Materialien, das Unterrichtsmaterial und die Weiterbildung aus. Spannungen zwischen Distriktbehörde einerseits und oppositionellen Lehrpersonen und Eltern andererseits störten die Zusammenarbeit.

Kalifornien ist als Beispiel einer systemischen Reform einerseits instruktiv, andererseits müssen die Ergebnisse im speziellen Kontext dieses Staates gelesen werden. Dazu gehören die Größe Kaliforniens, die Geschwindigkeit demographischer Veränderungen, die ungünstige Entwicklung der Staatsfinanzen oder auch wechselnde politische Mehrheitsverhältnisse. Knapp (1997) verweist deshalb auf SSI in Montana, Delaware und Connecticut sowie auf systemische Reformen in Michigan, South Carolina, Vermont und Kentucky.

In Montana, Delaware und Connecticut überzog auf der Distriktebene in den ersten beiden Jahren die Investition in die Etablierung einer gemeinsamen Vision der Reform. Besondere Anstrengungen wurden unternommen, über Weiterbildungsmaßnahmen „teacher leaders“ zu profilieren. Weiter setzten die Distrikte auf die Beteiligung von Lehrpersonen in die Curriculumentwicklung oder -auswahl und investierten in logistischen Support für Unterrichtsmaterialien. Erste Evaluationen dieser Prozesse nach zwei Jahren konnten noch kaum Wirkungen auf der Ebene des Unterrichts feststellen (Zucker/Shields 1995). Das verweist auf lange Zeiträume einer wirksamen Implementation.

Von Michigan, South Carolina, Vermont und Kentucky ist aus Studien im Rahmen von Forschungsprogrammen des *Center for Policy Research in Education* (CPRE) und der *Educational Policy and Practice Study* (EPPS) (Price/Ball/Luks 1994; Grant 1995; ergänzt unter anderem durch Spillane/Thompson/Lubienski et al. 1995) am meisten bekannt über die Implementation von systemischen Reformen in den Bereichen Mathematik und Naturwissenschaften. Insgesamt bestätigen diese Untersuchungen die bereits für Kalifornien festgestellten Mechanismen und Wirkungen. Frühe Effekte auf der Ebene Klassenzimmer zeigen eine partielle Implementation, etwa durch die Verwendung neuer Lehrmittel, bei gleichzeitiger Beibehaltung traditioneller Unterrichtsmuster und methodischer Grundanschauungen. Wie in Kalifornien haben die Reformen Aktivitäten und Initiativen bei den Akteuren auf der Distriktebene ausgelöst, diese wiederum haben bei den Lehrkräften neue Vorstellungen von Unterricht stimuliert, auch wenn es sich hierbei um individuelle Versionen der Reformidee handelte. Gleichzeitig zeigt etwa die Gegenüberstellung von Kalifornien mit Vermont und Michigan (Grant 1995) auch Unterschiede, die ihren Ausgangspunkt in unterschiedlichen staatlichen Strategien haben und etwa in Beziehung stehen zu Traditionen bezüglich der Verteilung von Autonomie und zentralstaatlicher Kontrolle.

Massell/Kirst/Hoppe (1997) haben Gemeinsamkeiten und Differenzen systemischer Reformen Mitte der neunziger Jahre in neun Bundesstaaten untersucht, nämlich in Kalifornien, Connecticut, Florida, Georgia, Kentucky, Minnesota, New Jersey, South Carolina und Texas. Die Reformen waren teilweise unterstützt durch SSI. Bezogen auf Fragen der Implementation wird in der Studie festgehalten:

- Aus politischen und pädagogischen Gründen wurden staatliche Standards und Curriculum Frameworks eher allgemein formuliert, um den Distrikten und Schulen die detaillierte Ausführung zu überlassen. Distriktadministratoren und Lehrkräfte verlangten oftmals mehr Anleitung und Support, als die Staaten offerierten. Die Unterstützungssysteme waren nicht zureichend.
- Befürchtungen, dass diese Reformen die lokale Autonomie beschneiden würden, traten eher nicht ein. Die verstärkte staatliche Rolle im Schulwesen stimulierte, zumindest in High-Capacity-Distrikten, die Initiativen auf lokaler Ebene. Einige Distrikte übernahmen eine Führungsrolle.
- Staatliche Initiativen förderten oder waren zumindest nicht hinderlich für die schul- oder distrikteigenen Aktivitäten bezüglich Curriculum und Unterricht. Obwohl sich der Einfluss der staatlichen Reformen auf der Distrikt-

ebene in vielen Fällen feststellen ließ, waren die staatlichen Reformen aber nicht immer die wichtigste Quelle eigener Umsetzungen. Gleichzeitig und parallel wirkte sich die Adaption nationaler Regierungs- und Nicht-Regierungsprogramme (z.B. NAS; SSI der National Science Foundation; American Association for the Advancement of Science Project 2061; NCTM-Standardsdokumente etc.) auf Gehalt und Fortgang der Reformen aus.

- Das lokale Personal betrachtete die staatlichen Standards als lediglich eine unter anderen Ressourcen, die verwendet wurden, um die eigenen, detaillierteren curricularen Dokumente und Leitlinien zu generieren. Besonders initiative Distrikte bemühten sich, zusätzliche finanzielle Mittel und technische Unterstützung von Stiftungen, nationalen Organisationen und anderen Gruppierungen zu erhalten. Diese zusätzlichen Projekte können, wenn sie in ihren Zielsetzungen nicht übereinstimmen, die Gefahr der Fragmentierung mit sich bringen.
- Zu Beginn der Standards-Bewegung fokussierte die Politik die Entwicklung von Dokumenten zur Anleitung des Unterrichts, also curriculare Leitfäden, Frameworks und Assessments. Die Notwendigkeit, zusätzliche Kapazität aufzubauen, trat erst im Verlaufe des Prozesses in den Vordergrund. Weiterbildung etwa gewann vor allem dann an Bedeutung, wenn die Frameworks, wie im Fall von Kalifornien, stark konstruktivistisch orientiert waren. Auch der Übergang von den traditionellen Multiple-Choice Tests zu Performance-Based Assessments bedeutete eine neue Herausforderung für Lehrkräfte und verlangte zusätzliche Investitionen.
- In Reaktion darauf wurden neue Programme in den Bereichen Lehrerausbildung und Lehrerweiterbildung lanciert sowie die Lehrerlizenzierung neu geregelt und beides an die Schülerstandards gekoppelt; auch Netzwerke zwischen Schulen und Lehrkräften wurden aufgebaut. Mancherorts wurden Lehrkräfte in die Entwicklung von Assessments und Standards eingebunden. Gewöhnlich fehlte den lokalen Administratoren und den Lehrkräften aber die Expertise, eigene Programme zu entwickeln.

Die Ergebnisse von Studien der neunziger Jahre machen deutlich, dass die Implementation von standardbasierten systemischen Reformen, wird deren Erfolg an der tatsächlichen Wirkung auf den Unterricht und das Lernen der Schülerinnen und Schüler gemessen, ein langwieriger Prozess ist. Diesen Befund bestätigen zahlreiche neuere Untersuchungen (wie Michlin/Seppanen/Sheldon

2001; Wong/Anagnostopoulos/Rutledge et al. 2001; Corcoran/Christman 2002; Marlette/Goldston 2003; O'Shea 2003).⁴⁵ Eine Studie zu New Jersey zeigt etwa, dass viele Lehrpersonen auch nach intensiver Weiterbildung, die sich dezidiert an den neuen Standards orientierte, Schwierigkeiten hatten, eine Lektion mit Aktivitäten zu planen, die auf einen Content Standard hinführen (O'Shea 2003). O'Shea schließt, dass die Ausrichtung der Weiterbildung auf curriculare Standards nicht ausreiche. Die Lehrkräfte müssen konkrete Lektionen planen und durchführen, deren Fokus nicht einfach Standards, sondern Lernergebnisse sind, die die Standards einlösen.

„Teachers need to learn how to plan a standards-based lesson that focuses on the achievement of outcomes related to the standards rather than simply conducting an activity aligned with standards. Furthermore, teachers need to learn how to evaluate student work samples in relation to the performance outcomes included in the standards-based lessons“ (O'Shea 2003, S. 16).

Die bisherigen Untersuchungen zur standardbasierten Schulreform verweisen auf eine eher bescheidene Wirkung auf Unterricht und Schülerleistungen. Zumindest gilt das kurzfristig. Schmoker (2004) führt diese eher enttäuschenden Resultate auf die Komplexität der Ansätze zurück, die zu Überforderung, Überfrachtung und Frustration bei den Lehrpersonen, den Schulleiterinnen und Schulleitern sowie anderen Akteuren geführt habe. Die Involvierung verschiedener Interessengruppen, wie von systemischen Ansätzen gefordert, habe sich bisher nicht als effektiv erwiesen (Hatch 2001). Oft sind die Konsequenzen einer Systemreform nicht klar, vielfach verlangen sie einen Mehraufwand, der sich nicht unbedingt auszahlt, und die Effekte zeigen sich wenn, dann zeitlich gestaffelt und kaum in Übereinstimmung mit den Zielen. Die Komplexität des Vorhabens führt dazu, dass zu viele wichtige Variablen im Spiel sind, von denen kaum eine begründet ausgeschlossen werden kann.

Michael Schmoker bezeichnet diese Reformen insgesamt als gescheitert. Er ist der Vertreter einer einflussreichen Reformlinie,⁴⁶ die einen dezidierten Bottom-up-Ansatz vertritt, also auf Teamentwicklung, professionelles Lernen und Ca-

⁴⁵ Die aufgeführten Untersuchungen sind alle über ERIC abrufbar.

⁴⁶ Dazu gehören bekannte Autorinnen und Autoren wie Linda Darling-Hammond, Michael G. Fullan, Milbrey McLaughlin, Allan Odden, Judith Warren Little, Tom Corcoran, Susan Fuhrman, Richard F. Elmore, Stephanie Hirsh oder Anne Lieberman.

capacity-building in den Schulen setzt. An die Stelle langfristiger und komplexer Pläne zur Reform einer Schule oder des Schulsystems seien kurzfristige Ziele mit direktem Fokus auf eine Verbesserung des Unterrichts zu setzen. Lehrpersonen müssten Erfahrungen sammeln können, in denen sich die Innovationen als sinnvoll erweisen. Dazu sollen in und zwischen den Schulen Lerngemeinschaften oder Netzwerke gebildet werden. Hier entwerfen, planen und evaluieren Lehrpersonen gemeinsam Unterrichtsmaterial und wenden es mit konkreten Ergebnissen auf ihre Situation hin an. Die Erwartung ist: *Results now* (Schmoker 2006).

Durch die bloße Übernahme von extern entwickelten Programmen, so die Kritik, werde das Schulpersonal zu passiven und abhängigen Empfängern fremder Scripts degradiert, statt als aktive Mitglieder eines Forschungs- und Entwicklungsteams zu agieren. Die von Schmoker und anderen stark gemachte Perspektive mündet ein in zahlreiche Arbeiten, die vor dem Hintergrund der Implementation standardbasierter Reform auf die zentrale Bedeutung der Aus- und Weiterbildung von Lehrpersonen sowie von Schulleiterinnen und Schulleitern verweisen. Diese Fragen spielen auch innerhalb der Systemreformansätze eine bedeutende Rolle. Die Gegensätze dürften weniger stark sein, als es den Anschein hat. Die konkreten Strategien lassen sich unterschiedlich gewichten, je nachdem, ob die Basis oder die darüberliegenden Ebenen thematisiert werden. Aber quer dazu liegen sämtlichen Ansätzen bestimmte Theorien zugrunde, über die Implementationsprozesse auf den verschiedenen Ebenen des Systems und mit Bezug auf beteiligte Individuen und Gruppen konzeptualisiert werden.

3.1.5 Theoretische Ansätze der Implementation

Die Analyse der amerikanischen Forschungsliteratur zu Innovationen im Bildungsbereich lässt einen bestimmten Schluss zu. In den vergangenen zehn Jahren ist ein zunehmendes Bewusstsein darüber entstanden, dass Implementationsprozesse entgegen vielen Planungspapieren nicht linear verlaufen, ihre anfängliche Komplexität eher steigern als reduzieren und unvorhergesehene Schwierigkeiten erzeugen. Damit einher gehen eine verstärkte Berücksichtigung von Reaktionsweisen, Lern- und Umsetzungsprozessen an der „Basis“ sowie eine zureichende Unterscheidung der verschiedenen Ebenen. Die Vielzahl der Variablen und interagierenden Faktoren verlangt nach theoretischen Modellen. Vorherrschend ist in vielen Studien ein Konsens, dass beide Perspektiven – die systemische Außensicht sowie eine (konstruktivistische) kognitionspsychologi-

sche Sichtweise – von Bedeutung sind und sich in der Forschung und Evaluation zu ergänzen haben (Knapp 2002; Spillane 2004).

Im Folgenden werden drei theoretische Modelle dargestellt, die in der amerikanischen Diskussion seit dem Ende des Zweiten Weltkriegs von Bedeutung waren und in der einen oder anderen Variante immer noch sind (vgl. Knapp 1997). Die Modelle unterscheiden sich unter anderem in ihrer Perspektivierung von Mikro- und Makroebenen und geben damit Bottom-up- und Top-down-Prozessen unterschiedliches Gewicht.

Innovations- und Change-Theorien:

Dieses Theoriekorpus hat sich bereits in den sechziger Jahren ausgehend vom privaten Sektor im Bereich der Bildung etabliert. Reformen werden in dieser Perspektive von der oberen Systemebene (Staat, Distrikt) initiiert. Strategien des „planned change“ führen auch gegen Widerstände zur Übernahme der Innovation an der Systembasis (Schule, Lehrpersonen). Die Schwäche dieses Ansatzes liegt vor allem darin, dass er von einem abgrenzbaren Produkt ausgeht, das an den Endnutzer zu bringen ist. Bei Schulreformen handelt es sich aber um komplexe Veränderungen, die verschiedene Prozesse und Praktiken betreffen, in welche wiederum verschiedene (intermediäre) Akteure und Agenten einbezogen sind. Sie setzen kein Produkt um, sondern stellen eines her, das sich im Laufe des Prozesses und je nach Anwendung verschiedener Praktiken verändert.

Theorien der Policy-Implementation:

Diese Theorien stammen aus der Policy-Forschung, die davon ausgeht, dass Wandel immer eine bestimmte Politik hat. Implementiert werden nie einzelne Elemente, sondern Bündel von Konzepten und Strategien, die mit einem politischen Willen verbunden sind. Die entsprechenden Theorien werden im Bildungsbereich seit Mitte der siebziger Jahre rezipiert. Wie die Innovationstheorie beschäftigt sich auch dieser Ansatz mit Prozessen des gezielten Wandels. Der Fokus liegt jedoch auf den Dynamiken, über die politische Maßnahmen und Verordnungen auf ihrem Weg durch verschiedene Ebenen und Instanzen des Systems gedeutet und (um-)interpretiert werden, bevor sie schließlich zur Umsetzung gelangen. Besondere Aufmerksamkeit erhält die Interaktion von Politiken in verschiedenen Kontexten auf verschiedenen Ebenen, also etwa von nationalen Regulationen mit Gesetzgebungen und Maßnahmen auf staatlicher und lokaler Ebene (z.B. Sunderman 1999).

Die Rolle der Distrikte als Kontexte der Umsetzung staatlicher Reforminitiativen wurde ab Mitte der neunziger Jahre Gegenstand vermehrter Aufmerksamkeit (vgl. Elmore 1993; Spillane 1996; Spillane/Thompson 1997; Spillane 2002b; Sipple/Killeen 2004; Spillane 2004). Ausgehend von einer wechselseitigen Beobachtung und Adaption, bei der es zu einer Reinterpretation der Politik durch die lokale Umgebung kommt, finden lokale Voraussetzungen und Umsetzungskapazitäten Berücksichtigung. Wie diese bedeutenden kontextuellen Faktoren und Bedingungen jedoch zustande kommen oder was genau sie beeinflusst, darüber vermögen diese Theorien wenig auszusagen.

Theorien des professionellen und organisationalen Lernens

Das zentrale Konzept dieser Theorien ist „Lernen“. Das Konzept wird sowohl auf Personen als auch auf Organisationen bezogen. Im Mittelpunkt des Interesses stehen die Möglichkeiten und Voraussetzungen für anhaltendes und nachhaltiges Lernen der Lehrpersonen, der Schulleiterinnen und Schulleiter sowie der anderen beteiligten Akteure. Der Blick richtet sich auch auf die zur Veränderung notwendigen Ressourcen und Unterstützungsangebote. Auch die Organisation Schule „lernt“, allerdings ist organisationales Lernen nicht einfach die Summe individueller Lernprozesse. Organisationen wie die Schule

„are seen as learning by encoding inferences from history into routines that guide behavior. The generic term ‚routine‘ includes the forms, rules, procedures, conventions, strategies, and technologies around which organizations are constructed and through which they operate. It also includes the structure of beliefs, frameworks, paradigms, codes, cultures, and knowledge that buttress, elaborate, and contradict the formal routines. Routines are independent of the individual actors who execute them and are capable of surviving considerable turnover in individual actors“ (Levitt/March 1996, S. 517).

Im Folgenden sollen vor dem Hintergrund dieser groben Klassifikation drei Arbeiten, die Prozesse der Implementation ausgehend von verschiedenen Theorieansätzen unter Bezug empirischer Erfahrungen zu konzeptualisieren versuchen, dargestellt werden.

a) Policy Attributes Theory

Mittels einer Sekundäranalyse von Studien zur Umsetzung von Projekten der Comprehensive School Reform hat Desimone (2002) Faktoren untersucht, die

die Implementation von Reform Designs befördern oder behindern. Die Analyse erfolgt gemäß der *policy attributes theory* (Porter/Floden/Freeman/Schmitt/Schwille 1988; Porter 1994), die näher fasst, wann und wie sich Innovationen positiv entwickeln.⁴⁷ Erfolgreiche Implementation lässt sich ausgehend von dieser Theorie bestimmen über die Attribute *specificity*, *consistency*, *authority*, *power* und *stability*.

Spezifität (*specificity*): Je spezifischer eine Innovationsstrategie ist (Material, Information, professionelle Entwicklung, Leitung, Instruktion), desto wahrscheinlicher wird sie vom Lehrkörper implementiert. Studien zur Comprehensive School Reform haben entgegen der Kritik gezeigt, dass die Implementation von Programmen schneller und einfacher verläuft, wenn es sich um extern entwickelte handelt, im Gegensatz zu lokalen oder schuleigenen. Bei der Spezifizierung der Programme kommt der Schulleitung eine zentrale Rolle zu, die jedoch nur dann trägt, wenn bei den anderen Akteuren hinreichend Klarheit darüber besteht, welche Rolle ihnen im Reformprozess zukommt.

Verschiedene Untersuchungen haben auf die herausragende Bedeutung von Weiterbildung hingewiesen, wobei anschauliches Material, konkrete und beispielbezogene Instruktion von den amerikanischen Lehrerinnen und Lehrern bevorzugt wird. Präferiert werden schulbezogene und den speziellen Bedürfnissen der Lehrkräfte entsprechende Weiterbildungsprogramme. Generalisieren lässt sich der Befund, dass dies in besonderem Maße für schwächere Lehrer und Lehrerinnen gilt, während exzellente Lehrpersonen stärker auf Autonomie ansprechen.

Aber: Je spezifischer ein Reformprogramm ist, desto weniger zählen Autonomie und Kreativität. Hier gilt es, das Anliegen der Reform gegenüber den Befürchtungen und tatsächlichen Gefahren einer Deprofessionalisierung transparent zu machen. Geschätzt werden von den Lehrkräften Evaluationen ihres Unterrichts und persönliches Feedback, mit dem sie direkten Aufschluss über die erreichten und noch zu erwartenden Fortschritte erhalten. Dabei ist die Qualität der Kommunikation und des Supports eine ausschlaggebende Größe.

Konsistenz (*consistency*): Damit ist die möglichst hohe Übereinstimmung von Schulreformstrategien mit Strategien der Politik gemeint, die von übergeordneten Ebenen ausgehen. Die Konsistenz bezieht sich auf die schuleigene Orga-

⁴⁷ Die *policy attributes theory* wurde unter anderem von Clune (1998) und Berends/Chun/Schuyler et al. (2002) zur Evaluation von systemischen Reformen (SSI der National Science Foundation) und Comprehensive School Reforms (NAS) verwendet. Eine weitere Studie ist die von Desimone/Smith/Hayes et al. (2005).

nisation, den Unterricht und die Lehrpläne. Als problematisch bezeichnet werden Ausgangslagen, in denen die Stossrichtung der Schulreform nicht mit den staatlichen Standards und Assessments übereinstimmt.

Autorität (*authority*): Nachdem verschiedenste Studien zur Schulentwicklung gezeigt haben, dass die Lehrkräfte die Schlüsselkomponente des Erfolgs seien (Desimone 2002, S. 446), liegt es auf der Hand, dass Autorität der einflussreichste Faktor ist. Zugleich ist dieser Faktor der am schwierigsten zu erreichende, da er auf Überzeugungen beruht, aber auch gewisse personale Fähigkeiten und Eigenschaften auf beiden Seiten voraussetzt sowie komplexe Interaktionen zwischen den Lehrkräften und den Trägern der Reform. Neben der möglichst weitgehenden Beteiligung der Akteure an den Entscheidungen kann die Überzeugungsarbeit dadurch erleichtert werden, dass den Lehrpersonen die Möglichkeit gegeben wird, auf verschiedenen Ebenen Netzwerke zu bilden. Dadurch wird der professionelle Austausch angeregt, Reformen können so zum täglichen Gespräch werden. Die Kooperation in Netzwerken trägt zum positiven Schulklima sowie zur Entwicklung der Schulkultur bei und ist wirksamer als politische Direktiven. Schließlich wird dadurch auch die Expertise der Lehrerschaft gefördert, wobei entsprechende Zeitgefäße vorhanden sein müssen. Implementierungsprozesse sind insbesondere in der Anfangsphase sehr zeit- und aufwändig.

Im Zusammenhang mit der Frage der Autorität ist wiederum auf die Rolle der Schulleiterinnen und Schulleiter hinzuweisen (vgl. die Studien von Anderson/Shirley 1995; Cooper/Slavin/Madden 1998; Haynes 1998). In einigen Untersuchungen sind sie der stärkste Prädiktor von Veränderung (Berends/Chune/Schuyler et al. 2002 und weitere RAND-Studien; vgl. Desimone 2002, S. 449). Mit diesen Befunden steht wiederum die Frage der dezentralen Autonomie zur Diskussion, die wiederum positiv korreliert mit der Implementation von Schulreformprogrammen (Bodilly/Berends 1999).

Durchsetzungsmacht vs. Autorität (*power vs. authority*): „Power operates through the force of the rewards and sanctions associated with the reform efforts while authority operates through the persuasive mechanisms of law, norms, and individual and expert backing and support“ (Desimone 2002, S. 453). „Power“ kann kurzfristig erreicht werden durch verschiedene Arten der Belohnung, etwa indem staatliche Reformprogramme monetär unterstützt werden. Aber die damit verbundene Motivation bricht oft mit den Incentives ab. Eine ausgewogene Balance zwischen „power“ und „authority“ ist wichtig etwa im Blick auf die Frage, wer bei der Wahl des Reformprogramms mitbestimmen kann und wie groß die Unterstützung der Wahl durch Informationen ist.

Stabilität (*stability*): Stabilität ist derjenige Faktor, der am schwierigsten zu beeinflussen sein dürfte. Darunter sind personelle Kontinuitäten, insbesondere Schulleitung, Lehrkörper, Administration, die Verlässlichkeit des politischen Umfelds sowie der anhaltende Verlauf der Reformen zu fassen. Implementierungen im Schulbereich sind langfristige Unternehmungen, es können Jahre vergehen, bis sich Effekte zeigen, wobei auch Zeit für Evaluationen eingeplant werden muss. Auf sichtbar große Effekte zu Beginn des Prozesses folgt gewöhnlich eine Abnahme in den folgenden Jahren und daraufhin eine mehr oder weniger kontinuierliche Effektzunahme (Borman/Hewes/Overman et al. 2003). Befragungen von amerikanischen Lehrerinnen und Lehrer zeigen, dass es Jahre dauern kann, bis das Verständnis für die Reformen und ihre Ziele bei den Lehrkräften durchgedrungen ist (Bodilly 1996).

b) Policy Implementation and Cognition

Erfahrungen mit der Implementation von Reformen und der Replikation von Schuldesigns (vgl. Coburn 2002) haben nochmals bestätigt, dass es sich dabei um einen komplexen Prozess handelt, bei dem den lokalen Ebenen (Distrikt, Schule, Klassenzimmer) und besonders den Lehrpersonen eine zentrale Rolle zukommt. Aufgrund dieser Wahrnehmung haben sich Forschung und Theoriebildung zunehmend auf diese Größen eingestellt. Entsprechende Ansätze lassen sich grob den Theorien des professionalen und organisationalen Lernens zuordnen. Ein Beispiel ist die Arbeit von Spillane/Reiser/Reimer (2002).

Ausgehend von einem umfangreichen Korpus an theoretischer und empirischer Literatur betrachten die Autoren Probleme der Policy Implementation unter einem kognitions- und sozialpsychologischen Blickwinkel. Eine Kernfrage richtet sich dabei auf Voraussetzungen des konstruktiven Sense-Making bei den „Agenten“ der Implementation. Sense-Making im Sinne von Karl Weick (1995) wird dabei nicht lediglich als individueller kognitiver Prozess untersucht, sondern situiert in den lokalen sozialen Kontexten. Von Bedeutung für die Interpretation von politischen Signalen und das Sense-Making sind Faktoren wie die Geschichte und Traditionen einer Profession (Yanow 1996), professionelle Diskurse sowie formelle und informelle Netzwerke (Coburn 2001).

Neben dieser Bottom-up-Sicht wird für das Sense-Making auch die Bedeutung von politischen Signalen im Top-down-Prozess in Betracht gezogen. Damit sollen insbesondere behavioristische Ansätze abgewiesen werden, die misslungene Implementation vorschnell auf „Sabotage“, Defizite und Ignoranz gegenüber neuen Anforderungen zurückführen. Dem wird die Komplexität von Sinnstif-

tungsprozessen angesichts der Ambiguität von Informationen, die Interpretationsspielräume freisetzen, gegenübergestellt. „Thus implementing agents’ sense-making provides numerous opportunities, aside from any willfull or intentional efforts to revise policy to fit with local agendas, for the transformation of policymakers’ ideas about changing local practice“ (Spillane/Reiser/Reimer 2002).

Die Konstruktion von Sinn basiert immer auf Vorwissen und wird von sozialen Kontexten beeinflusst. Besonders schwierig wird es, wenn von verschiedenen Seiten und Ebenen unterschiedliche Botschaften mit widersprüchlichen Zielsetzungen ausgehen. In bildungspolitischen Prozessen ist das fast immer der Fall, wie sich nicht zuletzt an den Diskussionen zu den Bildungsstandards zeigen lässt.

Folgende Faktoren müssen bei Prozessen der Implementation als potentiell hindernd berücksichtigt werden:

Vorgängiges Wissen und eingespielte Praktiken. „What is novel is always seen in terms of past understandings. In large part, ‚people generate what they interpret‘ – they create the environment and select the cues and signals that they interpret“ (Weick 1995, S. 34; Spillane 2002b, S. 394; vgl. auch Hill 2001). Studien zur Implementation standardbasierter Unterrichtsreformen haben gezeigt, dass dasselbe Reformkonzept von verschiedenen Lehrpersonen individuell ganz unterschiedlich verstanden werden kann. Neue Informationen werden immer im Licht dessen, was bereits verstanden wurde, interpretiert – wie aus der Gestaltpsychologie oder der philosophischen Hermeneutik bekannt ist und es auch die Entwicklungspsychologie Piagets kennzeichnet.

Übertragen auf das Feld der Schule lässt sich festhalten: Neue Ideen von Unterricht und Lernen standen und stehen immer in der Gefahr, als geringe Abweichung von dem betrachtet zu werden, was man schon immer gewusst und getan hat. Auch die mit den Standards verknüpften Konzepte von Qualitätssicherung und Schulreform werden von vielen Lehrpersonen und auch von anderen „Implementationsagenten“ assimiliert, ohne die Kernidee hinter dem Neuen zu verstehen. Das Neue erscheint fast wie das Alte, so dass besonderer Wandel gar nicht nahe liegt. Fullan (2001) beschreibt entsprechende Wahrnehmungen von Lehrpersonen als „false clarity“. Die Klarheit entsteht aber nur aus der Bestätigung des Bekannten.

Im Blick auf Prozesse des Sense-Making müssen neben kognitiven Aspekten auch emotionale und motivationale einbezogen werden. Anforderungen, bisherige Praktiken und Wissensbestände zu reformieren oder gar zu revidieren, bedeuten eine Infragestellung der positiven Selbstwahrnehmung, die an be-

stimmte Werte und Überzeugungen geknüpft ist. Wenn Reformanliegen als konsistent mit vorhergehenden Glaubenssätzen (beliefs) und Wissensbeständen interpretiert werden, dann kann das auch als eine Art Schutzmechanismus gedeutet werden. Das korreliert mit der Wahrnehmung: Auf Reformen, die nur relativ wenig Veränderung bringen, wird von den „Implementationsagenten“ mehr Aufmerksamkeit und Energie gerichtet, als auf weitergehende (Firestone/Fitz/Broadfoot 1999).

Sozialer Kontext: Grundsätzlich gilt hier: „Implementing agents encounter policy in a complex web of organizational structure, professional affiliations, social networks, and traditions. Both macro and micro aspects of the situation are important for implementing agents’ sense-making“ (Spillane/Reiser/Reimer 2002, S. 404). Der in der Policy-Forschung häufig gewählte Fokus auf die Institution ist im Blick auf die Schule nicht in gleicher Weise ergiebig. Schulen tendieren dazu, ihr Kerngeschäft, den Unterricht, von formalen Strukturen zu entkoppeln. Entsprechend gut dokumentiert ist der geringe Einfluss von Politik auf Lehrkräfte.

Bedeutsamer für Fragen der Implementation auf Handlungsebene und speziell bezogen auf Sense-Making ist daher der Mikrobereich des unmittelbaren Arbeitskontextes. Verschiedene Studien verweisen auf die Rolle sozialer Interaktion und Kommunikation in Kollegien oder Netzwerken, also kollektiver Formen des Sense-Making (z.B. Spillane/Zeuli 1999; Coburn 2001). Auch die Zugehörigkeit zu einer Profession ist von Bedeutung, weil sich dadurch eine berufliche Identität entwickelt, die über die Stufe, auf der man unterrichtet, das Fach, das man unterrichtet, hinausgeht.

Design und Repräsentation von Policy: Textdokumente sind das häufigste Medium der Kommunikation von Reformen. Sie bilden damit einen zentralen Ausgangspunkt des Sense-Makings, weil sie gelesen und als repräsentativ erachtet werden. Ausgehend von der politischen Ebene werden Reformen oftmals in Form von kurzen Aussagen kommuniziert, die sich auf Ziele und Zwecke beziehen, ohne die zugrundeliegenden Ideen und Konzepte auszuführen. Vor dem Hintergrund der Tendenz, Neuerungen in althergebrachte Schemata zu assimilieren, sind diese Verkürzungen problematisch. Auch ein Vorgehen, bei dem Neuerungen als lediglich geringfügig „verkauft“ werden, um die Akzeptanz zu erhöhen, ist sehr zweischneidig. Die Spannung besteht zwischen der notwendigen Erzeugung von Dissonanz und Unzufriedenheit mit dem bislang Bewährten und der Gefahr, Abwehr zu erzeugen.

Störend für den Sense-Making-Prozess ist zudem, wenn eine Reform über verschiedene Dokumenten etwa der Gesetzgebung, Standards, Testinstrumenten oder in Informations- und Weiterbildungsunterlagen unterschiedliche Definitionen erhält.

c) Die Outside-in- und inside-out-Perspektive: Versuch einer Synthese

Unter dem Titel *Rethinking Scale: Moving Beyond Numbers to Deep and Lasting Change* schlägt Cynthia Coburn (2003) vier Dimensionen vor, anhand derer das Scaling-up von Comprehensive School-Reform-Programmen evaluiert werden soll. Sie geht davon aus, dass die Forschung zu Problemen des Scaling-up noch in den Anfängen steckt. Ihr eigener Ansatz wendet sich unter anderem gegen die Evaluationsdesigns der RAND-Studien, mittels derer die Implementation von NAS in Teilen verfolgt worden ist.

Traditionelle Definitionen von „scale“ erfassen die quantitative Breitenwirkung von Reformen, also die Zahl der Schulen, die ein Reformdesign übernommen und appliziert haben (Bodilly 1996; Blum 1997; Fullan 2000; McDermott 2000). Theorien, die sich neben dieser Breitenwirkung auch mit der Tiefenwirkung und der Nachhaltigkeit befassen, fehlen weitgehend (Elmore 1996; Stokes/Sato/McLaughlin et al. 1997; Gamson 1998; McLaughlin/Mitra 2001). Gemäss Coburn müssen Definitionen von „scale“ so gefasst werden, dass sie

„include attention to the nature of change in classroom instruction; issues of sustainability; spread of norms, principles, and beliefs; and a shift in ownership such that a reform can become self-generative“ (Coburn 2003, S. 3).

Die Kritik Coburns stellt einen Versuch dar, Implementation jenseits von Zahlen und oberflächlichen Manifestationen in ihren tiefgreifenden und nachhaltigen Wirkungen zu erfassen. Das Maß der Implementation bzw. des Scaling-up eines Comprehensive-School-Reform-Modells müsse demnach die Dimensionen

- depth,
- sustainability,
- spread und
- shift in reform ownership

einbeziehen (so auch Glennan/Bodilly/Galegher et al. 2004). Dies impliziert, dass der gemessene Grad an Implementation entsprechend verschieden ausfallen kann.

Knapp (1995, 1997) hat mit der Gegenüberstellung zweier Untersuchungen – die eine mit einer „Inside-out“-Perspektive (Guthrie 1990) und die andere mit

einer „Outside-in“-Perspektive (Marsh/Odden 1991) – vorgeführt, wie die Ergebnisse zum selben Phänomen (Implementation der kalifornischen Mathematik-Frameworks von 1985) unterschiedlich ausfallen können. Er hat daraufhin versucht, einen neuen Forschungszugang zu konzipieren, der beide Perspektiven vereinigt (Knapp 2002).⁴⁸

Knapp (2002) fasst unter die *Outside-in-Perspektive* Theorien der Policy Implementation (Knapp/Bamburg/Ferguson et al. 1998) oder Forschung zum Alignment systemischer Reformen (Fuhrman/O’Day 1996). Was hierbei weitgehend unberücksichtigt bleibe, sei ein spezifisches Verständnis von der Wirkung dieser Politiken auf das individuelle und kollektive Denken und Handeln von Lehrkräften. Diese Fragen bilden den Ausgangspunkt der *Inside-out-Perspektive* (vgl. etwa Cohen/Ball/Raudenbusch 2000). Sie basiert auf der Forschung zum Unterricht, zum Lehrerlernen⁴⁹ und zur professionellen Entwicklung von Lehrkräften. Die entsprechenden Theorien nehmen Bezug auf kognitions- und sozialpsychologische Ansätze, zum Teil unter Einschluss einer soziokulturellen Perspektive. Im Zentrum steht nicht die politische Umgebung, sondern stehen die praktizierenden Lehrkräfte und die Entwicklung ihrer professionellen Identität und Kompetenz. Anwendung findet dabei häufig ein konstruktivistisches Verständnis von Lernen. Eine Linie innerhalb des Ansatzes bilden Studien zu den Auswirkungen multipler Kontexte auf den Unterricht von Lehrpersonen. Eine andere Linie erforscht die Rolle von Netzwerken und Gemeinschaften (communities) (Grossman 1996; Lieberman/McLaughlin 1996; Nelson/Hammerman 1996). Eine dritte Linie beschäftigt sich mit der Frage, wie Lehrpersonen Unterrichtsentscheide treffen. Neuere Forschungen in dieser Tradition untersuchen die Realisierung und Reinterpretation von Reformanforderungen durch Lehrpersonen (Withford/Jones 2000).

⁴⁸ Positiv im Hinblick auf eine Vermittlung beider Perspektiven hebt Knapp (2002) hervor: Bezug nehmend auf die Bedeutung professioneller communities innerhalb von Schulen und Distrikten: Coburn (2001), McLaughlin/Talbert (2001), Gallucci (2003); bezüglich der Bedeutung der Organisation und Allokation von Unterrichtsressourcen: Miles/Darling-Hammond (1998); bezüglich der Bedeutung von instructional leadership: Peterson (1999), Glickman (2002); Lehrerdemographie und Arbeitsmarkt: Ingersoll (2000); Design und Implementation von Weiterbildung: Cohen/Hill (2001); bezüglich Einfluss der Distriktpolitik auf den Unterricht eines Faches: Grossman/Thompson/Valencia (2001).

⁴⁹ Einen Überblick zur Forschung zum Lehrerlernen bieten Bransford/Brown/Cocking (1999).

Knapp (2002) plädiert im Hinblick auf zukünftige Forschungsdesigns für eine Erweiterung der Inside-out-Perspektive. Ein solches Design sollte folgende Bereiche und Kontexte der Lehrerarbeit mit einschließen:

- professional environment – which defines good practice, shapes preparation and support for professional work,
- community environment – which is a source of preferences and resources, not to mention the student population itself,
- school and district organisational environment – which builds structures and organisational cultures surrounding professional work (ebd., S. 19).

Outside-in-Untersuchungen können ihrerseits verschiedene Dimensionen des Unterrichts einbeziehen, die sensibel sind für politische Einflüsse:

- how learners are grouped and who is assigned to teach them,
- what is taught and how,
- how students engage with learning tasks and with each other,
- how instruction is designed and planned,
- what is learned and how learning is demonstrated (ebd.).

Eine weitere Theorievariante sieht die Implementation von Schulreformen als ko-konstruktiven Prozess (Datnow/Hubbard/Mehan 1998). Auch hier werden die linearen Modelle der sechziger Jahre (wie Carlson 1965) kritisiert. Der Verlauf von Innovationen hat weder nur eine Richtung noch nur eine Perspektive, verschiedene Richtungen und unterschiedliche Perspektiven unterscheiden sich nach ihrer Gestaltungs- und Durchsetzungsmacht, sind aber immer auf Ko-Akteure angewiesen. Auch Strukturen sind handlungsabhängig, können also nicht einfach durch Vorschriften oder Erlasse gesetzt werden. „Social actions constitute social structures and are simultaneously constrained or enabled by them“ (Datnow/Hubbard/Mehan 1998, S. 3).

Reformen werden generell wie eine Verhaltensmatrix verstanden, „as a web of interrelated conditions and consequences, where the consequences of actions in one context may become the conditions of the next“ (Hall/McGinty 1997, S. 461). Damit ist eine Heuristik gewonnen, die zeigt, dass Implementierungen nicht einfach wie ein Prozess des Weiterreichens konzipiert werden können. Die Idee der flexiblen Matrix zeigt, „that educational implementation is generated in face-to-face-interactions among real people confronting real problems in concrete social contexts, such as classrooms, school board meetings, courts of law, and state legislation“. Was im einen Kontext geschieht, kann on-

sequenzen in anderen haben, die oft weit entfernt sind. Beteiligt sind immer Akteure, die nie nur eine Wahl haben, auf Reformen zu reagieren.

„In the real world, educators may act in a variety of ways in response to reforms. Some may initiate reform efforts. Some may push or sustain reform efforts. Others may resist or actively subvert these efforts. Regardless of the course of action taken, the agency of educators is part of a complex dynamic, shaping and shaped by the structural and cultural features of school and society“ (Datnow/Hubbard/Mehan 1998, S. 2).

Die Studie basiert auf Untersuchungen von Reformprogrammen in Florida und Kentucky. Sie kommt zu folgenden Schlussfolgerungen:

- (1) Reform efforts in schools do not succeed simply on technical considerations, nor do they proceed in a linear fashion, fixed in time and space. Analyzing the policy implementation process as a conditional matrix affords a better opportunity to understand the complexities of a successful prototype's implementation than does a unidirectional (especially top-down) interpretation.
- (2) The consequences of actions taken in one context become the conditions for actions taken in other contexts.
- (3) The implementation process is viewed differently from different perspectives. Although local arrangements may vary significantly from what designers intended, the changes may be necessary for the reform.
- (4) The culture of the school mediates educators' actions and structural constraints.
- (5) The actions of local educators reflect institutional distribution and application of power (ebd., S. 13).

3.1.6 Veränderungen und Problemfelder der Lehrerbildung

Der Fokus der Forschung und Entwicklung zu Fragen der Implementation standardbasierter Reformen mit dem Ziel einer Verbesserung der Schülerleistungen hat sich bereits in den neunziger Jahren zunehmend auf die Lehrpersonen gerichtet (z.B. Darling-Hammond 1990, 2000; Sanders/Horn 1995; Sanders/Rivers 1996; Wright/Horn/Sanders 1997; Darling-Hammond/Loewenberg Ball 1998; Sanders/Horn 1998; Lewis/Parsad/Carey et al. 1999; Sanders/Topping 1999; Mayer/Mullens/Moore 2001). Der Grund ist bereits mehrfach genannt worden. Die Schulreform im Rahmen der Standardsbewegung war in

eine Sackgasse gelangt, weil die meisten Schulen und Lehrkräfte sich nicht als fähig erwiesen, den geforderten standardsorientierten Unterricht umzusetzen. Oft lag das nicht an mangelndem Willen oder fehlendem Engagement, sondern daran, dass das System, in dem sie arbeiteten, die Lehrkräfte nicht ausreichend unterstützte (Darling-Hammond 1998).

Resonanz fand in diesem Zusammenhang insbesondere der Report der National Commission on Teaching and America's Future *What Matters Most: Teaching for America's Future* (1996). Der viel zitierte Bericht stellte als die zentrale Botschaft in den Vordergrund, dass das Wissen und Können der Lehrkräfte der bedeutendste Faktor im Hinblick auf das Lernen der Schülerinnen und Schüler sei und dass Schulreformen nur dann erfolgreich sein könnten, wenn sie auf die Schaffung guter Unterrichtsbedingungen für die Lehrkräfte zielten. Unter dieser Prämisse sind erhebliche Anstrengungen unternommen worden, den Faktor „Lehrkraft“ zu entwickeln. Ansatzpunkt sind die Grundbildung der Lehrerinnen und Lehrer sowie die Programme zu ihrer Weiterbildung.

Mit der Reauthorisierung des *Higher Education Act* 1998 wurde unter Title II auf Bundesebene ein Reportingsystem eingerichtet, das von den Staaten und Institutionen der höheren Bildung (higher education) Berichte über die Qualität ihrer Lehrerbildungsprogramme einfordert. Seit 2002 publiziert das U.S. Department of Education jährlich einen Report, der Auskunft gibt über den Stand der Bemühungen der Staaten, hoch qualifizierte Lehrer und Lehrerinnen auszubilden (U.S. Department of Education 2002, 2003). Seit der Reauthorisierung der *ESEA* 1994 (*LASA*) und neuerdings mit dem *NCLB*-Gesetz macht sich eine verstärkte Einflussnahme des Bundes im Hinblick auf Standards der Lehrerbildung bemerkbar. Mit dem *NCLB*-Gesetz werden erstmals Kriterien auf Bundesebene festgelegt, die Personen in den USA erfüllen sollten, um als Lehrerin oder Lehrer von den Staaten zugelassen zu werden.

Der 2002 erstmals erschienene Report des U.S. Department of Education bezüglich Lehrqualität (2002) wurde von wissenschaftlicher Seite wegen der darin gezogenen Schlüsse kritisiert. Der Report führte Forschungsergebnisse an, welche aber etwa von Darling-Hammond/Youngs (2002) wiederum mit anderen Studien widerlegt oder relativiert wurden. Der Report und die Kritik daran spiegeln eine Diskussion in den USA, in der – aufgrund widersprechender Befunde – angezweifelt wird, dass die vorhandenen Studien zur Frage der Effizienz der Lehrerbildung und -lizenzierung aufgrund ihrer methodologischen Qualität überhaupt Schlüsse, statt lediglich Meinungen, zulassen (Darling-Hammond 2002 vs. Allen 2003; Whitehurst 2002; National Education Association 2004).

Entwicklungen, die in den Vereinigten Staaten Anlass waren, die Neukonzeption der Lehrerbildung zu diskutieren, sind ein veränderter Begriff von Accountability, wie er Mitte der sechziger Jahre aufgekommen ist (Cuban 2004), die Bildungsreformbewegung seit den achtziger Jahren (Sirotnik 2004), das Andauern der Leistungsdifferenzen zwischen verschiedenen Schülergruppen (achievement gap), die zunehmende Rolle des Bundes⁵⁰ sowie ein verstärktes Forschungsaufkommen zu Fragen der Lehrerbildung und auch der Einfluss von Marktmodellen auf die Bildungspolitik (z.B. Thomas B. Fordham Foundation 1999; Hess 2001; The Teaching Commission 2004).

In einem Artikel vom Oktober 2005 hat Cochran-Smith die aktuellen Trends und Reformen im Lehrerbildungsdiskurs in den USA, wie sie sich seit den späten 90er-Jahren manifestieren, analysiert und folgendermaßen zusammengefasst:

- *Lehrerbildung wird gegenwärtig diskutiert als ein Problem der Bildungspolitik.* Die Diskussion ist durch einen zentralen Befund stimuliert worden, der lautet: Die Lehrerbildung in ihrer bisherigen Form leistet keinen Beitrag zur Erhöhung von Schülerleistungen. Die Reform der Lehrerbildung gelangt damit in die obersten Ränge der politischen Agenda, begleitet von einer vereinfachten Wahrnehmung, wie Politik auf die Outcomes des Bildungssystems zu wirken vermag (zum Beispiel U.S. Department of Education 2002, 2003, 2004).
- *Die kritischen Befunde werden auf Forschungsergebnisse und Evidenz gegründet.* Mit den Title-II-Bestimmungen des 1998 reauthorisierten *Higher Education Act* verlangt der Bund für die Vergabe von Zuschüssen an die Bundesstaaten Evidenz der Qualität von Lehrerbildungsprogrammen; Akkreditierungsinstitutionen wie NCATE verlangen unter anderem, dass die Ausbildungsprogramme ein datenbasiertes Assessment-System einschließen. Die Zuverlässigkeit der Forschungsergebnisse und vor allem die daraus abgeleiteten Schlüsse, Empfehlungen und politischen Initiativen stehen zumindest in Frage (vgl. Darling-Hammond/Youngs 2002; Allen 2003; Walsh/Hale 2004).
- *Die Lehrerbildung ist gesteuert durch Outcomes.* Das fordern analog zur Entwicklung der Bildungsstandards Autoren wie der Bildungsökonom Hanushek (2002) mit Erfolg. Staatliche und regionale Reformprogramme koppeln die

⁵⁰ Title II des 1998 reauthorisierten *Higher Education Act* enthält eine Anzahl Forderungen nach Berichterstattung und Accountability bezüglich der Qualität der Lehrerbildung in den Vereinigten Staaten.

Accountability der Lehrerbildung zunehmend an Schülerleistungen bzw. -leistungszuwächse. Elmore (2002) kritisiert demgegenüber, dass das *NCLB*-Gesetz die schlechteste Version unter den gegenwärtigen Accountability-Modellen befördert habe: Performance-Based Accountability beschränke sich allein auf das Testen.

Alle Argumente und alle Kontroversen, die auf Schulebene festzustellen waren, finden sich in der Diskussion um die Reform der Lehrerbildung wieder. Mit diesen Tendenzen verbinden sich Spannungen und Konflikte in mehreren Bereichen: zwischen Forderungen nach Selektivität und Diversifikation der Ausbildung, zwischen stärker betontem Inhaltswissen bezogen auf das zu unterrichtende Fach (subject matter knowledge) und zurückgedrängter Ausbildung in Pädagogik und Psychologie. Es entsteht ein Wettbewerb zwischen Universitäten und anderen Ausbildungsinstitutionen, die alternative und oft kostengünstigere Wege zur Lehrerlizenzierung und -zertifizierung anbieten. Schließlich besteht auch hier ein Widerspruch zwischen Regulation und Deregulation.

Im Folgenden werden fünf größere Themenkomplexe dargestellt, die sich mit Fragen der Aus- und Weiterbildung der Lehrpersonen befassen. Speziell behandelt wird deren Rolle im Kontext der standardbasierten Reformen. In der amerikanischen Diskussion wird immer wieder betont, dass eine erfolgreiche Implementation von Standards eine intensive Verbesserung der Personalentwicklung und so der Lehrerbildung verlangt. Es geht dabei grundsätzlich um die Frage, welches Wissen die Lehrpersonen benötigen und wie sie lernen müssen, wenn sie standardbasiert unterrichten sollen. Die Frage bezieht sich auf Aus- und Weiterbildung gleichermaßen und erhält in der Durchsetzung der Reformen einen hohen Stellenwert.

Allgemein geht es darum, wie Lehrerbildungsprogramme im Hinblick auf die Implementation von Standards umgestaltet werden müssen, damit sie zur Verbesserung der Leistungen sowohl der Schüler als auch der Lehrkräfte beitragen können. Programme der Lehrerbildung müssen zeigen, wie erfolgreiche Lernformen und Lernarrangements ausgebaut werden können, aber sie müssen auch die Möglichkeiten des Lernens von Lehrkräften in einem Accountability- und Assessment-System verständlich machen. Besondere Beachtung findet schließlich die Induktionsphase in der berufsbiographischen Entwicklung angesichts zunehmender Innovationsanforderungen.

Subject Matter vs. Pedagogical Knowledge

Einen wichtigen Ausgangspunkt der Diskussionen um eine Lehrerbildungsreform in den Vereinigten Staaten bilden verschiedene Studien aus den neunziger Jahren (etwa Ferguson 1991; Greenwald/Hedges/Laine 1996; Sanders/Rivers 1996; Darling-Hammond/Loewenberg Ball 1998). Diese Studien legen dar, dass die Kompetenz von Lehrkräften gemessen an absolvierter Ausbildung, Abschlüssen, Zertifikaten und Berufserfahrung der wichtigste interne Faktor für das Lernen und die Leistungen der Schülerinnen und Schüler ist. Aus diesem Befund ergab sich die Kernfrage der Reform: Wie entsteht professionelle Kompetenz und was beeinflusst ihre Qualität? In der Diskussion dieser Frage brachen alte Gegensätze aus der Geschichte der Lehrerbildung erneut auf. Strittig nämlich ist, ob das disziplinäre Wissen bezüglich der unterrichteten Fächer (subject matter knowledge) oder das pädagogisch-psychologische sowie allgemeindidaktische Wissen (pedagogical knowledge)⁵¹ den Schwerpunkt der Ausbildung ausmachen soll.

Entsprechend scheiden sich die Geister an der Frage, in welchen Bereichen für eine Reform der Aus- und Weiterbildung im Zusammenhang mit standardbasierten Reformen zu investieren wäre. *A Nation at Risk* führte die Bildungsmisere in den Vereinigten Staaten unter anderem darauf zurück, dass durchschnittlich 41% der Ausbildung von Lehrkräften auf Kurse in Pädagogik und Didaktik entfallen würden, was der Bericht als übermäßig viel im Vergleich mit der Ausbildung in den Unterrichtsfächern ansah. Tatsächlich hatte in den siebziger und achtziger Jahren eine Verlagerung des Gewichts von subject matter zu pedagogical knowledge stattgefunden, neuerdings lässt sich jedoch eine Rückverlagerung der Ausbildungsanteile zugunsten subject matter knowledge feststellen (Sosniak 1999).

Verschiedene Studien haben im Blick auf Mathematik und Naturwissenschaften gezeigt, dass ein starker Zusammenhang besteht zwischen der Ausbil-

⁵¹ In der Literatur wird gewöhnlich unterschieden zwischen: knowledge of the subject (knowing what to teach), skills in teaching (knowing how to teach) und pedagogical content knowledge (knowing which method to use with particular topics, with particular kinds of students, and in particular kinds of settings). Der Begriff „pedagogical content knowledge“ stammt von Shulman (1986), der davon pedagogical und content knowledge unterscheidet. Darling-Hammond (1999, S. 224ff.) unterscheidet innerhalb von pedagogical content knowledge: knowledge of development, differences, learner knowledge, understanding of motivation, knowledge about learning, assessing, teaching strategies, curriculum resources and technologies, collaboration, ability to analyze and reflect.

derung in den Fachdisziplinen und dem Lernzuwachs der Schülerinnen und Schüler (zum Beispiel Monk 1994; Goldhaber/Brewer 2000). Gerade für die Durchführung von curricularen und standardbasierten Reformen sei ein vertieftes inhaltliches Wissen bei den Lehrkräften eine bedeutende Voraussetzung, die häufig fehle (Corcoran/Goertz 1995).

Weiterbildung: Forderung nach Inhalt- und Praxisbezug

Wenn sich die Implementation von Standards und Reformen in der Umsetzung im Unterricht zu erweisen hat (Spillane 1999) und wenn die Kompetenzen der Lehrpersonen der zentrale Faktor für verbessertes Lernen der Schülerinnen und Schüler sind, dann muss der professionellen Entwicklung der Lehrkräfte generell verstärkt Beachtung zukommen. Diese Forderung ergibt sich aus einer großen Zahl von Studien, die Effekte der Aus- und Weiterbildung untersucht haben. Änderungen der Unterrichtspraxis und so neue Problemlösungen stellen hohe Anforderungen an das Lernen der Lehrkräfte und sind nicht leicht zu erreichen, vor allem, weil die Ressourcen weitgehend festgelegt sind (Putnam/Borko 1997; Ball/Cohen 1999). Standardbezogene Reformen verlangen bessere fachliche und methodische Kenntnisse, die im Blick auf das amtierende Personal nur mit einer effektiven Weiterbildung aufgebaut werden können (Garet/Porter/Desimone 2001; Desimone/Porter/Garet et al. 2002; Fishman/Marx/Best et al. 2003; zur Forschung zur Lehrerweiterbildung in den USA vgl. Borko 2004).

Die herkömmlichen Angebote der Weiterbildung gelten, wie bereits mehrfach erwähnt, als ungenügend. Gut untersucht ist die Rolle der Distrikte in Bezug auf Weiterbildung und Unterrichtsreformen (z.B. Spillane 1996; Elmore/Burney 1999; Marsh 2000; Desimone/Garet/Birman et al. 2002). Die Studien zeigen, dass Distrikte und Lehrerbildungsinstitutionen (Colleges, Universitäten) meist separat und kaum koordiniert operierten (Desimone/Garet/Birman et al. 2003). Die von den Distrikten angebotenen Weiterbildungen werden traditionell charakterisiert durch Kurzworkshops ohne spezifischen Fokus und deshalb ohne Einfluss auf die Praxis (Darling-Hammond 1998). Gleichzeitig ist gut belegt, dass für die Durchführung von Reformen die Berücksichtigung lokaler Prioritäten für das Design und die Implementation von Weiterbildungsprogrammen zentral ist (z.B. Elmore/Burney 1999).

Sparks/Hirsh (1997) haben bereits Mitte der neunziger Jahre ein Umdenken festgestellt, das eine Verschiebung des Fokus in der Weiterbildung vom Distrikt auf die einzelne Schule und ihre Bedürfnisse verlangen würde. Drei Ideen hätten hauptsächlich dazu beigetragen, dass Weiterbildung neu konzipiert wird: die

Outcome-Steuerung, das Denken der Reformen vom Gesamtsystem aus und der Einfluss konstruktivistischer Lerntheorien. Damit rückt zunehmend die Entwicklung der Schule als Organisation in den Blick. Die Autoren sehen bereits eine Ablösung der fragmentierten Weiterbildungskurse durch Anlässe, die sich an einem kohärenten Programm orientieren. Dabei spielt eine entscheidende Rolle, wie die Leistungen der Schülerinnen und Schüler verbessert werden können. Formen des job-embedded Lernens lösen die vom Unterricht getrennten Kurse ab, die die Lehrperson in eine passive Situation versetzen. Thematisch nimmt die Ausrichtung auf die Entwicklung fachlicher und fachdidaktischer Kompetenzen zu.

Wilson/Berne (1999) haben verschiedene Aspekte von Weiterbildung analysiert und folgende Anforderungen als zentral identifiziert: Die Weiterbildung sollte auf Inhalte zielen, mit denen die Lehrpersonen selber im Unterricht konfrontiert sind; genügend Beachtung sollte den Lernprozessen von Schülerinnen und Schülern geschenkt werden; wertvoll sind vor allem realitätsnahe Versuche mit neuen didaktischen Formen. Hawley/Valli (1999) entwickelten acht Prinzipien effektiver Weiterbildung, deren Essenz auf der Erkenntnis beruht, dass Lehrpersonen dann am meisten profitieren, wenn ihre eigenen Fragestellungen und Interessen den Ausgangspunkt des Lernens bilden. Dazu tragen besonders Arrangements des kollaborativen Problemlösens bei.

Insgesamt lassen sich zwei Tendenzen aus der Forschung ablesen: Effektive Lehrerweiterbildung hat sich an der Praxis und ihren Problemstellungen zu orientieren. Der Fokus eines erfolgreichen Angebotes ist auf die zu unterrichtenden Inhalte zu fokussieren.⁵² In diese Richtung gehen auch die von Darling-Hammond/Loewenberg/Ball (1998) untersuchten Ansätze, die als positive Beispiele gelten können:

- *Integration von Theorie und Praxis*. Dabei kann es sich um Ausbildungsprogramme handeln, die ein zusätzliches Praxisjahr in einer Professional Deve-

⁵² Im Zusammenhang mit einer inhaltsorientierten Weiterbildung sind verschiedene kalifornische Initiativen hervorzuheben, etwa die Governor's Professional Development Reading Institutes von 1999, die 2000 auf weitere Fächer ausgeweitet wurden und Bezug nehmen auf die Schülerstandards sowie die im jeweiligen Fach staatlich approbierten kommerziellen Lehrmittel (z.B. Houghton Mifflin Reading oder McGraw/Hill Reading). Bekannt und gut untersucht sind zudem die California Subject Matter Projects (CSMP) (St. John/Dickey/Heenan et al. 1995; Pennell/Firestone 1996; Medina/St. John 1997; Firestone/Pennell 1997; Youngs 2001). Vgl. auch Kent/Lingman (2000) zur umfassenden California Professional Development Initiative (CPDRI) des California Department of Education.

lopment School beinhalten (vgl. Leading Edge Professional Development Schools Network (Griffin 2002; Whitford/Metcalf-Turner 1998).

- *Entwicklung einer professionellen Auseinandersetzung mit Problemen aus der Praxis.* Ein Beispiel ist das National Writing Project,⁵³ eine Weiterbildungsinitiative in Bezug auf Schreiben und den Schreibunterricht (vgl. Lampert 1999).
- *Content-based professional development:* Es handelt sich um Formen der Weiterbildung, die von konkretem Unterrichtsmaterial ausgehen. Ein Beispiel sind die Student Curriculum Workshops im Zusammenhang mit der Mathematik-Reform in Kalifornien (Cohen/Hill 1997).
- *Lernen aus der Analyse der eigenen Praxis.* Eine Möglichkeit stellt das INTASC-Portfolio-Assessment dar, welches neben verwendetem Unterrichtsmaterial und Schülerarbeiten auch Videoaufnahmen der Kandidatinnen und Kandidaten mit einbezieht.

Diese Beispiele rücken ab von der Praxis des Kurswesens, in dem großteils nur die Anbieter profitieren – meist Kurzanlässe ohne spezifischen inhaltlichen Fokus, die keinen großen Wirkungsgrad entfalten. Der Ertrag für Schule und Unterricht ist gering, der Transfer ist höchst zufällig und lässt sich vielfach auch gar nicht nachweisen. Einer solchen Praxis der Weiterbildung, die immer noch verbreitet ist, stellt Darling-Hammond (1998) effektivere Ansätze gegenüber, „that are curriculum-based, sustained over time, linked to concrete problems of practice, and built into teachers’ ongoing work with their colleagues“ (S. 9). Solche Programme eröffnen Lernmöglichkeiten, „that are

- connected to teachers’ work with their students,
- linked to subject matter and to concrete tasks of teaching,
- organized around problem-solving,
- informed by research, and
- sustained over time by ongoing conversations and coaching“ (ebd.).

Verschiedene Studien verweisen darauf, dass diese Maßnahmen eher Erfolg haben. Schülerinnen und Schüler in Schulen und Distrikten, deren Lehrkräfte an entsprechenden Weiterbildungen teilnehmen, zeigten höhere akademische Leistungen (ebd.). Die Bedingung dafür ist, dass die Lehrkräfte lernen können, wie sie ihre bisherige Unterrichtspraxis verändern können. Neue Curricula, Stan-

⁵³ <http://www.writingproject.org/> (Stand 13.01.06).

dards und Assessments bilden den Ausgangspunkt für Schulreformen, aber sie sind bedeutungslos, wenn die Lehrkräfte sie nicht anwenden können.

Linda Darling-Hammond hat ihre Empfehlungen für die Implementation standardbasierter Reformen wie folgt zusammengefasst:

„1) infuse new knowledge about learning and development in teacher preparation that includes extended clinical work in restructured schools, 2) support strong professional communities and networks that foster both practitioner learning and reform, and 3) restructure schools and schools of education as learning organizations that operate through inquiry and that examine practice through the lense of professional standards and student learning“ (ebd., S. 12).

Dieses Design von Fort- und Weiterbildung gründet auch auf der Erfahrung, dass der Einfluss von Forschungswissen auf die Unterrichtspraxis bisher eher gering ist. Ursache dafür ist die fehlende Verknüpfung, die eine institutionelle Trennung voraussetzt. Die besondere Betonung wissenschaftlichen Wissens hat dazu geführt, dass die Lehrerbildung vornehmlich an den Universitäten stattfindet. Bisherige Ansätze gingen davon aus, dass das dort generierte und vermittelte akademische Wissen die „Basis“ bilde, auf der das professionelle Lehrerwissen aufgebaut werden soll und kann.

Eine alternative Sichtweise betont, dass diese Wissensform grundsätzlich verschieden ist von derjenigen, auf welche Praktikerinnen und Praktiker gewöhnlich zurückgreifen (Hiebert/Gallimore/Stigler 2002). Der Transfer wissenschaftlichen Wissens in Wissen, welches die Lehrperson in der schulischen Praxis anwenden kann, setzt die Möglichkeit der Anknüpfbarkeit an das mitgebrachte Wissen und berufliche Erfahrungen voraus, welche nicht ohne weiteres gegeben ist (Korthagen/Kessels 1999). Lässt man die institutionelle Trennung bestehen – Universität hier, Schule dort – so entwertet der Praxisdruck die Verwendung wissenschaftlicher Kenntnisse. Diese müssen sich im professionellen Setting der Lehrkräfte als nützlich erweisen und dementsprechend vermittelt werden.

Netzwerke und Coaching

Neben alternativen Ansätzen in der Weiterbildung stehen zunehmend direkte Formen der Personalentwicklung im Mittelpunkt des Interesses. Netzwerke dienen dabei dem Austausch gleicher oder ähnlicher Erfahrungen, die sich unmittelbar im eigenen Praxisfeld anwenden lassen. Daneben stehen Rückmeldungen und Interventionen durch Coaches, die fach- und inhaltsbezogen ar-

beiten. Im Vordergrund des Reform-Designs steht dann ein differenziertes System an professioneller Fort- und Weiterbildung, dessen Zentrum Arrangements von Peer-Coaching in Netzwerken sowie das Prinzip des „content-focused coaching“ bilden (vgl. West/Staub 2003; Staub 2004).

Die Effekte dieses Ansatzes lassen sich etwa am Beispiel des Community School District #2, New York City, aufzeigen. Das Beispiel ist vergleichsweise gut untersucht und wird wegen des systematischen Ansatzes und der Verbesserung der Schülerleistungen häufig zitiert. Es gilt als erfolgreiches Modell der Schulentwicklung, wenngleich nicht alle Kritiker zustimmen.⁵⁴ Das Beispiel zeigt auch und erneut, welche bedeutende Rolle der lokalen Ebene des Distrikts in der Schulreform zukommen kann, sofern hier Personen agieren, die über brauchbare Konzepte verfügen und persönliches Engagement zeigen. Die zentrale Figur war Anthony Alvarado, der zwischen 1987 bis 1998 Superintendent des zweiten New Yorker Schuldistrikts war. Er hat die Reformen maßgeblich initiiert, deren Grundmerkmale Organisationsentwicklung, professionelles Lernen und Accountability waren.

Ein weiteres zentrales Element bildete die Ausbildung von Schulleiterinnen und Schulleitern über das *Aspiring Leadership Program*, welches in Zusammenarbeit mit dem Baruch College an der New York University entwickelt worden war. Eine Kooperation besteht ebenfalls mit der University of Pittsburgh und dem dort angesiedelten Learning Research and Development Center (unter der Leitung von Lauren Resnick). Das Institut hat Performance Standards und Assessments sowie ein darauf bezogenes Set von *Principles of Learning* entwickelt und unterstützte deren Implementation im Distrikt.

Den Fokus aller Strategien zur Qualitätssicherung bildet der Unterricht und mit ihm die Verbesserung der Schülerleistungen. Zu den Gelegenheiten der professionellen Entwicklung der Lehrerinnen und Lehrer gehören (Elmore/Burney 1997):

- *Professional Development Laboratory*: Eine ausgezeichnete und erfahrene Lehrperson fungiert als Resident Teacher. Visiting Teachers besuchen diese über

⁵⁴ Kritisch dazu Weiner (2003): In seinem Artikel spricht er von einer Cheerleader-Rolle des Designs und weist auf, dass die Schulen des Distrikts eine verhältnismäßig hohe Stratifikation bezüglich Rasse und sozialem Status aufweisen; zudem genieße der Distrikt eine überdurchschnittlich gute materielle Ausgangssituation. Insgesamt ist er skeptisch, dass professionelle Weiterbildung *der* bedeutende Faktor für die erfolgreiche Implementation von standards und school improvement sei, wie vielfach behauptet und teilweise mit Studien belegt wird.

drei Wochen im Klassenzimmer und lernen über Unterrichtsbeobachtungen. Währenddessen übernehmen qualifizierte Adjunct Teachers die Klassen der Visiting Teachers.

- *Instructional Consulting Services*: Externe Experten eines Unterrichtsfaches oder Berater des Distrikts arbeiten mit einzelnen Lehrpersonen oder Gruppen an konkreten Problemen zu spezifischen fachlichen Inhalten, in diesem Fall besonders in Bezug auf Literacy und Mathematik. Das Ziel ist die konkrete Verbesserung des Unterrichts.
- *Intervisitation and Peer Networks*: Lehrpersonen innerhalb einer Schule besuchen sich gegenseitig im Klassenzimmer, und Gruppen von Lehrkräften besuchen andere Schulen. Daneben existieren Schulleiterkonferenzen, die dem gegenseitigen Austausch über Themen des Unterrichts dienen, zum Beispiel die Entwicklung und Implementation von Standards zur Evaluation von Schülerarbeiten.
- *Off-Site Training* während des Schuljahres und im Sommer, zum Beispiel zu Fragen der Entwicklung und Implementation von Standards und Curriculum Frameworks. Die Themen dieser externen Weiterbildung sind auf eine begrenzte Auswahl von Unterrichtsinhalten gerichtet (content-focused) und werden über eine längere Zeitspanne bearbeitet. Bei der Themenwahl wird versucht, Prioritäten der Schulen und des Distrikts auszubalancieren.
- *Oversight and Principal Site Visits*: Häufige Schulvisiten durch den Superintendent und sein Team. Beurteilungsgrundlage der Besuche sind die von den Schulleitenden jährlich verfassten Zielsetzungen und die zur Erreichung der Ziele geplanten Aktivitäten.

Im Distrikt existiert zudem ein *Distinguished Teachers Program*. In diesem Programm werden hervorragende Lehrerinnen und Lehrer anderen Lehrpersonen mit besonderem Potential in einer der Focused Literacy Schools zur Seite gestellt. Focused Literacy Schools sind Schulen mit tiefen Schülerleistungen, die besondere Lese- und Schreibförderung benötigen. Die Zusammenarbeit dauert ein Jahr und zielt auf die gemeinsame Entwicklung eines verbesserten Unterrichts, der die Situation vor Ort verändert. Distinguished Teachers müssen zuerst ein Selektionsverfahren durchlaufen und dabei unter anderem die erfolgreiche Implementation von Standards im eigenen Unterricht aufweisen können. Sie werden für ihre Arbeit als Coach und Mentor nennenswert entschädigt.

Professionalisierung und Deprofessionalisierung

Der *National Education Summit* der Gouverneure unter Beteiligung der Wirtschaftselite 1996 in Palisades, NY, hatte mit seiner Stossrichtung in Richtung Effizienz und Accountability zahlreiche Kritik seitens der amerikanischen Erziehungswissenschaft hervorgerufen.⁵⁵ Die Kritiker sahen in den Beschlüssen des Gipfels die erneute Bestätigung einer Politik, die seit *A Nation at Risk* einseitig auf Standards, Tests, Accountability oder ICT-Technologie setze und die bürokratische Kontrolle über Schulen und Lehrpersonen verstärke (Smylie 1996). Diese Strategie hätte aber, wie die Forschung zeige, bisher kaum Effekte in der Unterrichtspraxis der Lehrpersonen erzeugt, weil es sich um externe Reformverordnungen handle, in die die eigentlichen Akteure, die Lehrerinnen und Lehrer, nicht einbezogen würden. Sie würden kontrolliert, aber nicht zum Lernen angehalten.

Eine alternative und wirkungsvollere Strategie sehen Autoren wie Mark Smylie in einer Fokussierung des Lernens von Lehrkräften im Sinne eines permanenten Prozesses und damit der Entwicklung von „Humankapital“: „While the primary change mechanism of regulatory policy is control, the crucial mechanism for building human capital is learning“ (ebd., S. 10). Gelegenheiten des Lehrerlernens müssen problemorientiert sein, gegründet in Forschung, Experiment und Reflexion, und sie müssen vertraute „beliefs“ und Praktiken mit einbeziehen. Der schulische Arbeitsplatz sei dabei eine reiche Quelle für Lehrerlernen (Smylie 1995). Ausserhalb der Schule bilden kollaborative Netzwerke, Lehrerzentren und Angebote von Fachorganisationen (subject matter associations) ebenfalls effektive Lernmöglichkeiten (Lieberman 1995).

Jennifer O'Day (2002) greift ebenfalls auf die Gegenüberstellung von professionellem Lernen und bürokratischer Kontrolle zurück und beschreibt die jeweils damit in Verbindung gebrachten Accountability-Mechanismen. Sie vergleicht das als bürokratisch attribuierte und Outcome-basierte Accountability-System in Chicago mit dem auf Professional Accountability setzenden System in Baltimore und plädiert für eine Kombination beider Mechanismen zur Qualitätssicherung. Gleichzeitig sieht sie mit der *NCLB*-Gesetzgebung die Gefahr einer Vernachlässigung professioneller Verantwortungslegung zugunsten vermehrter externer Kontrolle und Sanktionen, wie sie im Kern bürokratisch-administrativer Logik stehen.

⁵⁵ Ein Beispiel sind die Beiträge in der November-Nummer 1996 der Zeitschrift *Educational Researcher* von Weinstein; Sabers/Sabers; Natriello; Machr/Machr; Covington.

Zuvor ging bereits Spencer (1996) davon aus, dass die bisherigen Forderungen nach Standards und Accountability, aber auch die Forderungen der *National Commission on Teaching and America's Future* (1996) nach Ausdehnung der Lehrerbildung nicht zu einer Professionalisierung der Lehrkräfte führen werden. Diese Tendenzen forcierten vor dem Hintergrund der politischen Entwicklung in Richtung Dezentralisierung und Schulautonomie vielmehr eine Intensivierung der bisherigen Arbeitsanforderungen (intensification of tasks). Auf die Lehrkräfte komme mehr vom Gleichen zu, ohne dass wirkliche Innovationen initiiert würden.

Die Frage, ob die standardbasierten Reformen tatsächlich die Professionalisierung der Lehrkräfte unterstützen oder nicht vielmehr gerade ihre Deprofessionalisierung befördern, wird auch an der Frage der Rolle von Assessments⁵⁶ diskutiert: Werden Assessments lediglich zum Zweck der externen Kontrolle verwendet oder eröffnen sie – als Bestandteil des Unterrichts – Gelegenheiten des professionellen Lernens? Können Assessments im ersten Fall, der ein Misstrauen gegenüber den diagnostischen Kompetenzen der Lehrpersonen evoziert und Tendenzen der Deprofessionalisierung wachruft, der Implementation eines besseren Unterrichts dienen?

Ausgehend von einem Reformkonzept, das Top-down-Support bietet für Bottom-up-Reformen, verneint Darling-Hammond (1994) die letzte Frage. Den Hintergrund ihrer Ausführungen bietet die Einführung neu entwickelter performanz-basierter, „authentischer“ Assessments, die den Anspruch erheben, das Postulat der Chancengleichheit (equity) besser einzulösen als die traditionellen Multiple-Choice-Leistungstests. Beurteilt wird so nicht nur die Leistung zu einem bestimmten Testzeitpunkt, sondern eine Palette verschiedener Leistungen.

Bestimmte Reformansätze gehen davon aus, dass Leistungsüberprüfungen das Verhalten von Lehrpersonen stark beeinflussen würden und somit ein wirksames Mittel zur Herbeiführung von Schulreformen seien. Testbasierte Entscheidungsprozesse etwa bezüglich der Selektion von Schülerinnen und Schülern haben jedoch in erster Linie zu einem Unterricht geführt, der sich auf lower-order cognitive skills richtet. Die Selektionsmaßnahmen wiederum hätten die Ungleichverteilung der Lerngelegenheiten für Lernende befördert. Nicht in der Präzision von Selektionsentscheidungen, nicht in der „Objektivität“ von As-

⁵⁶ Glaser (1990) folgend unterscheidet Darling-Hammond zwischen Tests und Assessments: Testing zielt auf Selektion von Schülerinnen und Schülern und damit auf die Voraussage von Lernerfolgen vor der Platzierung in einer Klasse oder Schulform. Assessments eruieren Outcomes und messen somit die Lernergebnisse.

assessments zur Legitimation von Strafen und Belohnungen, sondern in deren Informationsgehalt für Lehrpersonen und deren Einbettung in Prozesse der Schul- und Unterrichtsreform liegen die zentralen Kriterien. In diesem Kontext können Assessments der Lehrperson Informationen über das Lernen der Schülerinnen und Schüler und die Effektivität des eigenen Unterrichts und damit Gelegenheiten des Lernens bieten. Lehrerinnen und Lehrer müssen dann aber selber in die Durchführung, Auswertung und Ergebnisinterpretation von Assessments einbezogen werden (vgl. Murnane/Levy 1996). Die Frage hat sich am Ende immer darauf zu richten, ob Assessments Curriculum und Unterricht positiv beeinflussen.

Für eine Verbindung von Unterricht und Leistungsüberprüfung plädiert auch Shepard (2000). Der Umgang mit entsprechenden Instrumenten könne nicht in separaten Kursen vermittelt werden, sondern müsse Bestandteil sein der unterrichtsmethodischen Ausbildung. Shepard geht bei ihrer Argumentation von (sozial-)konstruktivistischen Prämissen aus, die sie behavioristischen Annahmen diametral gegenüberstellt. Von Lehrpersonen durchgeführte Assessments, Feedbacks von Kolleginnen und Kollegen sowie Ergebnisse von Selbstevaluationen der Schülerinnen und Schüler sind in dieser Sichtweise zentrale Elemente eines sozialen Prozesses, der die Entwicklung von intellektuellen Kompetenzen, die Konstruktion von Wissen und die Entwicklung der Identität der Schülerinnen und Schüler vermittelt. Dieser kognitivistische Ansatz sieht Performance Assessments mit offenen Aufgabenstellungen vor, die das kritische Denken und die Lösung komplexer Probleme sowie die Fähigkeit zur Anwendung von Wissen in realen Kontexten bei den Schülerinnen und Schülern befördern sollen.

Die damit präferierte Funktion von Assessments zielt auf den Aufbau von Kapazität und die professionelle Entwicklung der Lehrkräfte (McLaughlin/Shepard 1995) und wendet sich gegen Accountability-Systeme, die auf Strafe und Belohnung setzen und zur Deprofessionalisierung beitragen würden.

Entwicklung von Lehrpersonen:

affektive Momente und die Bedeutung der Induktionsphase

Ausgehend von einem kognitions- und sozialpsychologischen Ansatz wird die professionelle Entwicklung von Lehrpersonen im Zusammenhang mit der Frage der Implementation von „change“ auch unter Einbezug affektiver Momente diskutiert (vgl. etwa Berg v.d. 2002). Im Zentrum stehen dabei die unterschiedlichen Wirkungen von Anforderungen des Wandels auf die individuelle Lehrkraft im berufsbiographischen Kontext. Es wird davon ausgegangen, dass die

Wahrnehmung und Sinngebung von Lehrpersonen (teachers' meaning) angesichts von Reformanforderungen durch höchst resistente Faktoren wie die persönlichen Überzeugungen, die Einstellungen und Haltungen, die professionelle und persönliche Identität sowie schwer steuerbare Emotionen vermittelt werden.

Die bildungspolitischen Entwicklungen und Reformen mit ihren erhöhten Anforderungen an die Lehrkräfte werden unter dem Stichwort „*intensification*“ behandelt (vgl. neuerdings Valli/Buese 2007). Zu diesem Phänomen der Intensivierung von Belastungsfaktoren trägt auch die Wahrnehmung bei, dass die Arbeit der Lehrerinnen und Lehrer in der Öffentlichkeit wenig beachtet und ihre Kompetenzen eher gering geschätzt werden. Intensification geht also einher mit einer Infragestellung der Professionalität der Lehrerschaft und wirkt sich nicht nur auf deren Selbstwahrnehmung negativ aus, sondern auch auf die Haltung gegenüber Innovationen. Mangelnde Möglichkeiten der Sinngebung sind hindernd für die Karrieregestaltung über verschiedene berufliche Phasen. Die Forschung zeigt dabei, dass die Erfahrungen von Lehrkräften in der Berufseinstiegsphase die professionelle Identität stark bestimmen und den Verlauf der Karriere determinieren können. Die Gestaltung der Induktionsphase in der Ausbildung gewinnt damit an Bedeutung.

Die Bedeutung der Induktionsphase wird auch im Zusammenhang mit dem Konstrukt „*teacher efficacy*“ hervorgehoben. Die Wirksamkeit der Lehrkräfte und die damit verknüpften Überzeugungen stehen sowohl in Bezug auf Schülerleistungen, die Implementation von Innovationen als auch auf Fragen der Lehrerbildung zur Diskussion. Bereits eine RAND-Untersuchung von 1977 (Berman/McLaughlin/Bass et al. 1977; vgl. auch Bandura 1977) hatte gezeigt, dass *teacher efficacy* einen positiven Effekt auf Schülerleistungen hat, aber auch auf den Erfolg der Implementation von Innovationen, etwa neue Unterrichtsmethoden. Zugleich machen neuere Studien deutlich (Valli/Buese 2007), dass sich Effekte einer *intensification* als Folge zunehmender Einflussnahme und Kontrolle (Accountability) von Seiten lokaler, staatlicher und Bundespolitiken gegenüber der Arbeit der Lehrkräfte negativ auf verschiedene *Efficacy*-Faktoren auswirken können und auch auswirken.

Teacher Efficacy wird definiert als „teachers' belief or conviction that they can influence how well students learn, even those who may be difficult or unmotivated“ (Guskey/Passaro 1994, S. 4). Eine andere Definition bezieht sich auf „the extent to which the teacher believes he or she has the capacity to affect student performance“ (Berman/McLaughlin/Bass et al. 1977, S. 137). Diese Überzeugungen beeinflussen den Energieaufwand und die Ausdauer angesichts von Hindernissen sowie die Flexibilität im Umgang mit Misserfolgen, die eine

Person aufbringen kann. Schließlich hängt davon auch ab, wie viel Stress ertragen wird und wie hoch die Toleranz gegenüber negativen Gefühlen ist. Goddard/Hoy/Woolfolk Hoy (2004) sehen in der weiteren Erforschung des Konstrukts der *efficacy* einen wichtigen Beitrag zur Einlösung der Forderungen des NCLB-Gesetzes. Die Optimierung der Selbstwirksamkeitsüberzeugungen sei ein wichtiger Faktor für den Wandel in Schulen und speziell für die Verbesserung der Leistungen sowohl der Lehrkräfte als auch der Schülerinnen und Schüler.

Teacher efficacy im Sinne der Einschätzung der eigenen Möglichkeiten, die Schülerleistungen positiv beeinflussen zu können, nimmt gemäß einer Untersuchung im Verlauf des ersten Praxisjahres ab. Die Lehrkräfte müssen erst lernen, nicht einfach von den Idealen ihrer Ausbildung auszugehen, sondern sich selbst im Umgang mit der Klasse zu erproben. Mit zunehmender Erfahrung entwickeln Lehrpersonen eine stabile Wahrnehmung der eigenen Wirksamkeit. Der Aufbau dieser Überzeugungen kann von der Ausbildung beeinflusst werden. Das erklärt, warum die Berufseingangsphase in allen Konzepten zur Lehrerbildungsreform ein so starkes Gewicht erhält. Die Investition in eine starke *teacher efficacy* zu Beginn einer Karriere lohnt sich, weil die ersten guten Problemlösungen nachhaltig wirksam sind. Wenn die Lehrkräfte standardbasiert unterrichten sollen, dann müssen sie mit dieser Fähigkeit in den Beruf hineinwachsen.

Im Anschluss an Bandura (1977) ist postuliert worden, dass das Vertrauen in das eigene Können als Lehrerin oder Lehrer durch die Ausbildung und den Berufsanfang geformt werden müsse. Frühe Gelegenheiten für Studierende, Erfahrungen im Unterrichten und mit der Führung einer Klasse in diversen und zunehmend komplexen Situationen zu machen und dabei Rückmeldungen zu erhalten, werden als bedeutsam für die Bildung und Aufrechterhaltung einer positiven Grundüberzeugung angesehen. Über die Entwicklung von *efficacy* bei Novizinnen und Novizen existieren zwar erst wenige Studien, jedoch scheint diese Größe zur Meisterung der Berufseinstiegsphase und zum Verbleib im Beruf beizutragen. In den späteren Berufsphasen wird *efficacy* im Zusammenhang mit der Implementation von Reformen untersucht und ein positiver Zusammenhang festgestellt. Die Einräumung von Mitbestimmungsmöglichkeiten, das Gefühl der Kontrolle über Situationen, ein unterstützendes Verhalten der Schulleitung sowie konstruktive Rückmeldungen von dieser Seite tragen insgesamt zur *efficacy* bei und wirken sich positiv auf den Implementationserfolg aus.

3.2 Niederlande

Seit dem Ende der französischen Besatzung und der Gründung des Königreichs der Niederlande (1815) war die Schulentwicklung in den Niederlanden gekennzeichnet durch politisch-gesellschaftliche Auseinandersetzungen. Am Ende des Jahrhunderts stand ein Kulturkampf um die Ausrichtung des Bildungssystems. Bedingt durch die revolutionären Entwicklungen im 19. Jahrhundert wurde im November 1848 eine liberale Verfassung ausgerufen, die auch Einfluss auf die Konstitution des Bildungswesens hatte. Die Verfassung sah die Freiheit des Unterrichtens vor, Gründung und Unterhalt der Schulen waren nicht mehr an eine staatliche Erlaubnis gebunden, der Staat sollte nur noch die Eignung der Lehrkräfte kontrollieren. Die Zulassung konfessioneller Schulen geht auf eine Modifikation des ersten holländischen Schulgesetzes aus dem Jahre 1806 zurück. Seit der liberalen Verfassung wurde die Gründung von freien Schulen ermöglicht, ein Aspekt, der bis heute von hoher Relevanz ist. Der Anteil privater Schulen im niederländischen Schulsystem ist sehr hoch und beträgt etwa 70% des Gesamtbestandes.

Die zweite Hälfte des 19. Jahrhunderts war geprägt durch eine Emanzipationsbewegung der Katholiken und Protestanten, die einen eigenen Schulunterricht basierend auf ihrer Religion und die finanzielle Gleichbehandlung von der Regierung forderten (Schwarz 1997, S. 25; Liket 1997, S. 248f.). Mit der tiefgreifenden Änderung des Grundgesetzes, das seit 1917 die Gleichwertigkeit von öffentlichen und privaten Bildungseinrichtungen festschrieb, wurde der sogenannte „Schulstreit“ (schoolstrijd) beigelegt, der seit der Einführung der allgemeinen Schulpflicht im Jahre 1900 eskaliert war und kulturkampfbähnliche Züge angenommen hatte. Bis 1917 wurden nur die öffentlichen Schulen (openbare scholen) mit staatlichen Mitteln unterstützt. Die konfessionellen oder sonstigen Privatschulen (bijzondere scholen) konnten den im 19. Jahrhundert kontinuierlich gestiegenen staatlichen Anforderungen an Ausstattung und Schulqualität nicht standhalten, übernahmen einen Teil der Unterrichtsversorgung und wollten gleich behandelt werden.

Der Streit wurde 1917 mit der Vollfinanzierung aller Privatschulen gelöst, die seitdem gilt. In der heutigen Verfassung gehört die freie Erteilung von Unterricht zu den Grundrechten, ebenso, dass der Unterricht „ganz oder teilweise“ aus öffentlichen Mitteln zu finanzieren ist (Verfassung des Königreichs der Niederlande §23).⁵⁷ Der Staat finanziert den Kauf und den Unterhalt der

⁵⁷ <http://deatch.de/niederlande/verfassung.niederlande.htm> (Stand 17.07.06).

Gebäude, die Lehrmittel, die Gehälter der Lehrkräfte und alle sonstigen Ausgaben vollumfänglich. An dieser Regelung hat auch das Basisschulgesetz von 1985 nichts geändert. Mindestens 200 Eltern, in den großen Städten mindestens 300, sind nötig, wenn eine Schule gegründet werden soll. Unabhängig von der didaktischen und/oder weltanschaulichen Ausrichtung muss der Staat für den Unterhalt dieser wie aller Schulen aufkommen. Der Finanzierungsmodus ist in der Literatur gelegentlich als Musterbeispiel für ein Voucher-System dargestellt worden, weil die Steuermittel an die Schulträger gehen, die wiederum frei gewählte Schulen unterhalten (James 1984).

Die auch als „Bildungspazifikation“ (pacificatie) bezeichnete Modifizierung des Grundgesetzes von 1917 sorgte für Religionsfrieden, hatte aber auch eine enorme Diversifikation zur Folge. Nur rund 30% der Schülerinnen und Schüler besuchen heute eine öffentliche Schule, um 1900 waren es noch etwa 70%. Es gibt im Privatschulsektor mehr als 6000 verschiedene Schulträger, die zu drei Vierteln in Vereinen oder Stiftungen organisiert sind. Die örtlichen Träger stellen die Lehrkräfte ein, was auch für die staatlichen Schulen gilt, nur dass es hier die Gemeinderäte sind. Die Zahl der privaten Träger ist im Bereich des *Basisonderwijs* (Kindergarten und Grundschule: 4. bis 12. Lebensjahr) sogar noch höher und liegt bei etwa 90%. Von den 8100 Schulen sind 7200 privat geführte. Die Mehrzahl der Schulen bezieht sich auf ein christliches Glaubensbekenntnis. Im Jahre 2006 gab es in den Niederlanden 6318 christliche Schulen, davon 4955 im Bereich Kindergarten und Grundschule sowie 1054 im Sekundar- und 309 im Tertiärbereich.

Die Schulen erhalten nahezu unbeschränkte Freiheiten nicht nur zu ihrer Gründung, sondern auch im Blick auf die Gestaltung und Organisation der Curricula, die Durchführung des Unterrichts sowie bezogen auf die Einstellung des Lehrpersonals. Eltern; Schülerinnen und Schüler können die Schule frei wählen, ohne dass der Staat eingreifen könnte. Es gibt auch keine Beschränkungen für die Einzelschule hinsichtlich der Auswahl der Schulbücher und anderer Lehr- und Lernmittel. Dementsprechend verfügte die Einzelschule innerhalb des niederländischen Bildungssystems im Vergleich zu anderen europäischen Ländern bereits zu Beginn des 20. Jahrhunderts über ein hohes Maß an Autonomie.

Allerdings sind inzwischen die Beteiligung aller Bildungseinrichtungen und so auch aller Schulen, der privaten wie der öffentlichen, an regelmäßigen Schulinspektionen durch das Bildungsministerium, die Teilnahme an landesweiten Abschlussprüfungen der Sekundarstufe sowie die Berücksichtigung arbeitsrechtlicher Bestimmungen verpflichtend. Die Verstärkung der Aufsicht erfolgt in einem bislang weitgehend autonomen System, das seit Beginn des 20. Jahr-

hunderts immer auf die einzelne Schule und deren Einbindung in örtliche Gemeinschaften gesetzt hat.

3.2.1 Der Wandel der Bildungspolitik

Charakteristisch für das niederländische Schulwesen heute ist die Kopplung zwischen einer zentralen Steuerung der Qualitätssicherung und einem dezentralen Management der Schulen. Das Bildungsministerium auf nationaler Ebene ist überwiegend für die grundlegende Organisation, die Finanzierung, die Verwaltung der öffentlichen Schulen, die Schulaufsicht sowie für die nationalen Prüfungen verantwortlich. Während auf der Ebene der Provinzen vornehmlich kontrollierende und rechtsprechende Aufgaben wahrgenommen werden, verwalten die Gemeinden die öffentlichen Schulen. Die privaten Schulen unterstehen, wie gesagt, eigenen Trägergruppen, die mit den Gemeinden in der einen oder anderen Form kooperieren.

Seit Mitte der achtziger Jahre weisen die Entwicklungen innerhalb des niederländischen Schulsystems auf der einen Seite deutliche Tendenzen zu einer noch weitergehenden Dezentralisierung und Deregulierung auf. Das Bildungsministerium übertrug den Schulen zunehmend mehr Autonomie hinsichtlich der Verwaltung und Organisation sowie für die Gestaltung des Unterrichts. Auf der anderen Seite sind aber auch neue Formen der Qualitätssicherung gefunden worden, die über die einzelne Schule und die örtliche Aufsicht hinausgreifen. Die Bildungspolitik ist dadurch aber nicht grundlegend anders geworden. Die liberalen Grundsätze gelten weiterhin, was die Niederlande von den drei anderen Ländern unterscheidet. Es gab weder einen radikalen Politikwechsel wie in England noch einen weitgehenden Umbau der Organisation wie in Schweden. Gleichwohl sind auch in den Niederlanden nach 1985 deutliche Veränderungen erfolgt.

1985 wurde das Primarschulgesetz verabschiedet. Das Gesetz ist 1997 revidiert worden und gilt heute in dieser Fassung. Die wesentliche Änderung 1985 war die Zusammenfügung von Kindergärten, Vorschuleinrichtungen und Primarschulen zu einer Stufe. Zudem wurde die Schulpflicht auf das fünfte Lebensjahr vorverlegt und ein freiwilliger Schulbesuch vom vierten Lebensjahr an ermöglicht. Das niederländische Schulsystem umfasst zwei Phasen und vier Stufen. Die erste Phase ist der bereits erwähnte achtjährige *Basisonderwijs*, der aus Kindergarten und Grundschule besteht. Daran schließt sich die Sekundarstufe an, die wiederum geteilt ist. Sie umfasst drei Säulen und beginnt jeweils mit einer drei- bzw. zweijährigen Grundbildungsphase. Die drei Säulen sind unterschiedlich lang, der Gymnasialzweig dauert mit der Grundbildungsphase

sechs Jahre, der allgemeinbildende Sekundarbereich fünf Jahre und der berufsvorbereitende Sekundarbereich vier Jahre.

Nach Liket lassen sich die heutigen Steuerungsmechanismen der niederländischen Regierung für das Bildungswesen durch folgende Grundsätze skizzieren:

- Formulierung globaler, aber eindeutiger Ziele und Grundsätze auf nationaler Ebene,
- erweiterte Schulautonomie in vielen Bereichen,
- pauschales Finanzierungssystem der Einzelschule,
- deutliche Motivation zum Management der eigenen Bildungseinrichtung,
- Unterstützung von Funktionsdifferenzierungen in der Schule,
- Entwicklung eines neuen Qualitätssicherungssystems,
- Etablierung einer schulinternen Evaluation und eines externen Evaluationssystems,
- Reform der Schulleiter- und Lehrerausbildung (Liket 1997, S. 249f.; Liket 1998, S. 62).

Der niederländische Staat verlagerte somit zunehmend die zentralen Entscheidungskompetenzen und Reglementierungen im Schulsystem auf die Ebene der Einzelschule. Infolgedessen beschränkten sich die gesetzlichen Vorgaben etwa im Blick auf den Unterricht nur noch auf die Formulierung allgemeiner Ziele (Liket 1997, S. 250). Durch die Pauschalfinanzierung erhielt jede Einzelschule ein Budget, über dessen Verwendung sie frei verfügen kann (Ven v.d. 2002, S. 333). Ferner bieten die eigenverantwortliche Gestaltung der Schulorganisation, des Curriculums und der Personalplanung ansprechende Profilierungsmöglichkeiten für die Einzelschule im Wettbewerb um die Schülerinnen und Schüler. Die Schulwahl ist, wie gesagt, freigestellt.

Dennoch oder gerade aus diesem Grunde wird ein Gleichgewicht zwischen der staatlichen und der Verantwortlichkeit der Einzelschule angestrebt. Dieses Gleichgewicht soll mit dem Prinzip der Rechenschaftslegung erreicht werden. Zudem ist es die Aufgabe des Staates, ein qualitativ hochwertiges Schulsystem zu gewährleisten, das allen Lernenden gleichen Zugang zu einer gleichwertigen Ausbildung ermöglicht. In der Verfolgung dieser Aufgabe behält sich der Staat eine kontrollierende Funktion in Form staatlicher Prüfungen und Evaluationsverpflichtungen vor. Eines der wichtigsten Elemente für die schulische Qualitätsentwicklung ist das Evaluationssystem, bestehend aus internen und externen Evaluationsmethoden.

Für die Ausweitung der Eigenverantwortlichkeit auf der Ebene der Einzelschule lassen sich vorwiegend drei Begründungen anführen (Liket 1997, S. 253f.; vgl. auch Liket 1995):

1) Erziehungswissenschaftliche Beweggründe:

- In der Forschung wurde zunehmend die Ansicht erhärtet, dass Top-down-Strategien kaum anregende und unterstützende Auswirkungen auf die Schulentwicklung und -innovation haben und Reformen oder Umstrukturierungen nur dann greifen, wenn die Schulen und Lehrer davon überzeugt sind.
- Die zuvor bestehenden detaillierten Reglementierungen und Vorschriften auf nationaler Ebene gaben dem zunehmenden Interesse an vielfältigeren Unterrichtsmethoden und deren Entwicklung und Anwendung nur wenig Raum.
- Erfahrungen zeigten, dass die 45- bis 50-minütigen Unterrichtseinheiten nicht für die Vielfalt didaktischer Konzepte und deren Umsetzung im Unterricht geeignet waren.
- Das starre Arbeitszeitmodell beeinträchtigt die Kompetenzen und Fähigkeiten der Lehrkräfte, für die andere Arbeitszeitregelungen notwendig seien.
- Die zunehmenden Ansprüche an die Schulqualität und -entwicklung machten die Entwicklung eines umfassenden Evaluationssystems erforderlich.

2) Soziale Beweggründe

- Die multikulturelle Zusammensetzung der Lernenden stellt erhebliche Anforderungen an die Lehrerinnen und Lehrer, die ihren pädagogischen Auftrag nur dann erfüllen können, wenn sie größere Spielräume und Freiheiten für ihre tägliche Arbeit zur Verfügung haben.
- Durch die Internationalisierung und Vernetzung der Gesellschaft ist es unumgänglich geworden, die Reformen und Entwicklungen der Bildungssysteme anderer Länder mitzuverfolgen und entsprechende Rückschlüsse für das eigene System zu ziehen.
- Auch die Einzelschulen bewegen sich, bedingt vor allem durch das Internet, im internationalen Raum. Demnach sind Rahmenbedingungen zu gestalten, die den Anschluss an die internationale Diskussion und an den internationalen Austausch ermöglichen.

3) Finanzielle Beweggründe

- Durch die Ausweitung der dezentralen Steuerungsmechanismen erhofft sich die Regierung, die Kosten im Voraus besser einschätzen und planen zu können.
- Durch eine Pauschalfinanzierung der Einzelschule können die Finanzmittel besser eingesetzt werden.

Liket beschreibt die neuen staatlichen Steuerungsmechanismen als ein Zusammenspiel von großer, an Vielfalt im Bildungsangebot ausgerichteter Unterrichtsfreiheit und einem konsequenten System interner und externer Qualitätsprüfungen, das gleichermaßen um Zielerreichung als auch um die Verbesserung der Unterrichtsqualität bemüht ist (Liket 1997, S. 252). Die externen Prüfungen sind nicht verbunden mit Sanktionen, ebenso wenig existiert ein Testsystem mit „high stakes“. Die Rückmeldungen sind professionsnah und orientiert an der Entwicklung der einzelnen Schulen, die gefördert und nicht bestraft werden sollen. Ihre Besonderheit muss schon aus Gründen der Anlage des Systems respektiert werden.

Ein zentrales und inzwischen viel genutztes Mittel der neuen Politik der Qualitätssicherung sind die Selbstevaluationen der Schulen. Anders als in anderen Ländern liegt der Fokus nicht auf der Entwicklung der Organisation, sondern auf der Entwicklung der Leistungen. Ein entsprechendes System ist von CITO, dem zentralen Institut für Leistungsmessung, in Kooperation mit anderen Einrichtungen zur Qualitätssicherung, vorgelegt worden. Die Leistungen der Schülerinnen und Schüler werden als der wichtigste Faktor der Bildungsqualität angesehen. Entsprechend steht dieser Faktor im Mittelpunkt der Evaluation, die Testdaten benutzt und das Leistungsmonitoring der Lehrkräfte verbessern will (Moelands/Ouborg 1998).

3.2.2 Verständnis und Bedeutung von Standards

Die niederländischen Schulen können eigenständig über die Organisation, die Inhalte und Methoden des Unterrichts entscheiden. Es bestehen keine bindenden staatlichen Lehrpläne wie in den deutschen Bundesländern und auch kein nationales Curriculum wie in England. Eine unlängst durchgeführte Vergleichsstudie der Curriculum-Politik mit anderen, vorwiegend europäischen Ländern kam zum Schluss: „[...] then the final conclusion is inevitable that the Netherlands are the uncrowned king of curriculum autonomy“ (Kuiper/Akker v.d./Hooghoff et al. 2005, S. 75). Auch die Auswahl und Nutzung der Schulbücher sowie anderer Lehrmittel und Unterrichtsmaterialien liegt im Verantwortungs-

bereich jeder einzelnen Schule; sie werden von kommerziellen Anbietern auf dem freien Markt produziert. Die Schule und ihre Trägerschaft sind zuständig für die Anstellung der Lehrkräfte ebenso wie für die Anstellung der Schulleitung. Beides unterliegt einem Ausschreibungs- und Auswahlverfahren. Die Lehrkräfte werden der Schule nicht einfach zugeteilt. Angesichts solcher Regelungen ist die staatliche Einflussnahme auf die Kontext- und Prozessqualitäten des Unterrichts traditionell und konstitutionell sehr gering (Akker v.d./Letschert 2002, S. 47; Ackeren v.d. 2003a, S. 105f.).

Dennoch sind die Leistungsziele und vor allem die zentralen Abschlussprüfungen, aber auch die Schulbücher starke Steuerungsinstrumente (Scheerens/Glas/Letschert 2005, S. 5).⁵⁸ Ferner werden die Schulen durch ein 1998 eingeführtes Qualitätsgesetz (*kwaliteitswet*) zur Entwicklung eines Schulplans (*schoolplan*) und eines Schulführers (*schoolgids*) verpflichtet, die als Grundlage für ein Qualitätskonzept dienen sollen, dass die Schule auch in der Öffentlichkeit repräsentiert (Ackeren v. 2003a, S. 106). Das niederländische Institute for Curriculum Development (SLO) vertritt ein Konzept, das verschiedene Ebenen unterscheidet, aber auch deutlich auf die einzelne Schule ausgerichtet ist.

„The curriculum concept concerns three levels: **micro, meso and macro**. The activities of the SLO are currently aimed at a national curriculum or a set of core objectives or attainment targets for a certain type of education. A curriculum at meso level for example, is a plan for a school or a group of schools. This concerns for example a group of schools with a common board and supporting a certain identity. A curriculum at micro level is a text-book for one course year or a lesson series for a certain subject.“⁵⁹

Bis Anfang der neunziger Jahre gab es keine nationalen Regelungen über fachspezifische Unterrichtsziele. Um jedoch die Schul- und Unterrichtsqualität, eine gleichwertige Ausbildung und eine Vergleichbarkeit sicherzustellen, wurde ein komplexes System von Kernzielen (*kerndoelen*) entwickelt. Die erste Ausarbeitung der Kernziele durch das Bildungsministerium von 1988 umfasste für die Primarstufe 464 sehr detailliert formulierte Ziele. Diese Fassung wurde vom niederländischen Parlament abgelehnt, weil der Schritt von bis dahin kaum

⁵⁸ Im Rahmen dieser Expertise wurde im Juni 2005 ein Interview mit drei niederländischen Bildungsexperten Jaap Scheerens, Cees Glas und Jos Letschert von der Universität Twente durchgeführt.

⁵⁹ <http://www.slo.nl/international/map7/> (15.08.06).

vorhandenen Zielen zu diesen sehr spezifischen Vorgaben als zu einschneidend empfunden wurde.

Daraufhin wurde 1993 ein neues Set von Zielen entwickelt, das aus 122 Kernzielen bestand (Scheerens/Glas/Letschert 2005, S. 6). Eine durch das niederländische Bildungsministerium in beratender Funktion einberufene „Kommission Kernziele“ (*commissie kerndoelen*),⁶⁰ kam 2002 zum Schluss, dass die Anzahl der Kernziele auf unter 70 Vorgaben zu reduzieren sei (Ackeren v. 2004, S. 252). Im August 2005 wurden dann, den Empfehlungen der Kommission folgend, 58 sehr allgemein definierte Kernziele eingeführt; ähnlich sind die Tendenzen bezüglich der Kernziele der unter Sekundarstufe, während auf der oberen Sekundarstufe, je näher die landesweiten Abschlussprüfungen rücken, die curricularen Vorgaben zunehmen. Auch an der Entwicklung und Ausgestaltung der Kernziele wird deutlich, dass das niederländische Bildungssystem weiteren Dezentralisierungs- und Deregulierungsprozessen unterliegt. Die zentrale Regulation ist immer noch schwach, wodurch der Einzelschule ein hohes Maß an Accountability abverlangt wird (Scheerens/Glas/Letschert 2005, S. 2). Und offenbar besteht in der Lehrerschaft kein Verlangen nach stärkerer zentraler Steuerung; stattdessen findet man aber zunehmend den Wunsch nach Unterstützung und nach der Verbreitung gelungener Beispiele für die Konstruktion der spezifischen Schulcurricula in Bezug auf die Kernziele (Kuiper/Akker, v.d./Hooghoff 2005). „The expectation is that such procedurally and substantively specified examples – possibly including descriptions of good performance for certain age levels [...] – will be more inspiring for schools and teachers than the current set of attainment targets“ (ebd., S. 76). Es kann also durchaus die Gefahr bestehen, dass mit in Zahl und Konkretion abnehmenden curricularen Vorgaben eine Art Vakuum entsteht. Eine Stärkung zentraler Regulierung erscheint jedoch im niederländischen Kontext wenig gewünscht und vorgesehen, mehr verspricht man sich hingegen von der Investition in den Aufbau von „Kapazität“ vor Ort über die professionelle Weiterbildung (ebd.).

Die Kernziele legen fest, was die Schülerinnen und Schüler am Ende der Grundschulzeit erlernt haben sollten. Wie die Einzelschule diese Vorgaben erreichen will, liegt in ihrem Verantwortungsbereich. Die Ziele werden alle fünf Jahre überprüft und angepasst.

Die Kernziele beziehen sich auf folgende Lernbereiche der Primarschule:

⁶⁰ Ein entsprechender Bericht ist unter dem Titel „Verantwoording delen. Herziening van de Kerndoelen basisonderwijs met het oog op beleidsruimte voor scholen“ 2002 erschienen (Advies van de commissie kerndoelen basisonderwijs 2002).

- Niederländisch,
- Arithmetik und Mathematik,
- Englisch (mindestens in den letzten beiden Jahren),
- Sinneskoordination und Körperbewegung,
- Künstlerischer Ausdruck,
- Sozialverhalten und „life skills“ wie Verkehrserziehung und
- verschiedene Wissensfächer (Geographie, Geschichte, Naturwissenschaften mit Biologie, Gesellschaftswissenschaften und Politik).⁶¹

Daneben wird auch die Ausbildung überfachlicher Kompetenzen angestrebt. Sie umfassen Lernhaltungen, die Fähigkeit, nach Plan zu arbeiten, verschiedene Lernstrategien, das Selbstbild der Schülerinnen und Schüler, ihr Sozialverhalten sowie die neuen Medien. Die Schulen können neben den gesetzlich vorgeschriebenen auch weitere Fächer oder Lernbereiche anbieten. Religionsunterricht gehört zum Angebot, kann aber abgewählt und durch andere Lernbereiche kompensiert werden. In der Festlegung des Stundenplans, also in der Zeitverteilung, sind die Schulen frei. Nur die Gesamtzeit des Unterrichts ist vorgeschrieben.

Erfahrungen mit den Kernzielen vor 2002 zeigten auf, dass sie insgesamt zu allgemein gehalten waren, nur unzureichend auf die Unterrichtspraxis übertragen werden konnten und keine Angaben zu den Kompetenzniveaus enthielten. Niveauangaben fehlten generell. Die Interpretationsspielräume bei der Anwendung in den Schulen waren groß, und derart allgemeine Ziele können mit sehr vielen Erfahrungen in Verbindung gebracht werden. Zur Vermeidung von Beliebigkeit und zur besseren Handhabung der Ziele wurden für die Fächer Mathematik und Niederländisch Lehrlinien und Zwischenziele (*leerlijnen* und *tussendoelen*) entwickelt (Ackeren v. 2004, S. 252).

In der bildungspolitischen Diskussion in den Niederlanden hat der Begriff „Bildungsstandards“ keine lange Tradition. Allerdings befinden sich nunmehr Lehr- und Lernstandards (*leerstandards*) in der Entwicklung, die als Qualitätsmaß fungieren sollen, um auf Lernrückstände der Schülerinnen und Schüler reagieren zu können. „Hierbei geht es um operationalisierte und genormte Mindestkompetenzen, die am Ende bestimmter Unterrichtsphasen beherrscht werden müssen. Sie geben an, was Schulen ihren Schülerinnen und Schülern mindes-

⁶¹ Die beste nicht-holländische Übersicht über das niederländische Schulcurriculum bietet das International Review of Curriculum der englischen QCA an. Die Daten sind dieser Seite entnommen: <http://www.inca.org.uk/1330.html> (Stand 15.08.06).

tens beibringen müssen und sollen auch Grundlage für die Schulbuchentwicklung sein“ (Ackeren v. 2003a, S. 106).

In einem Bericht des niederländischen Bildungsrates⁶² in Bezug auf die Empfehlungen der *commissie kerndoelen* werden zwei Niveaus für die zu entwickelnden *leerstandards* vorgeschlagen. Diese Standards geben an, in welchem Maße die Schülerinnen und Schüler die Kernziele beherrschen müssen. Grundsätzlich wird ein Niveau der Minimumstandards vorgesehen, das festlegt, welche Unterrichtsinhalte die Lernenden mindestens erlernt haben und das von 90% der Schülerinnen und Schüler erreicht werden sollte. Das zweite Niveau liegt höher und wird als befriedigend bezeichnet (*voldoende standaarden*). Dieses Niveau sollte von 70-75% der Lernenden erreicht werden (Ackeren v. 2003a, S. 106; Ackeren v. 2004, S. 253).

Hinsichtlich der Schulcurricula sprach die *commissie kerndoelen* (2002) ebenfalls Empfehlungen aus. Insbesondere regte sie an, den Schulen mehr Verbindlichkeit für die Gestaltung der schulinternen Curricula abzuverlangen. Dementsprechend schlägt die Kommission eine Differenzierung in ein Kerncurriculum, das verpflichtend für alle Schulen gilt, und in ein schulspezifisches Curriculum, das von der Einzelschule individuell gestaltet werden kann, vor. Vorgesehen ist, dass das Kerncurriculum sich auf 70% der Unterrichtsinhalte bzw. -zeit beziehen soll, während die restlichen 30% dem schulspezifischen Lehrplan vorbehalten werden sollen. Das wäre ein sehr weiter Eingriff in die bisherigen Freiheiten.

3.2.3 Überprüfung von Standards: Tests und Prüfungen

Das zentral gesteuerte niederländische Test- und Prüfungssystem kann als ein funktionales Äquivalent zur umfänglichen Eigenständigkeit der Einzelschule aufgefasst werden. Es handelt sich dabei aber nicht um eine ausdifferenzierte und komplexe Prüfungs- und Testkonzeption, wie diese in anderen europäischen Ländern vorliegt. Dennoch gehören Tests und zentrale Prüfungen zum Alltag in den niederländischen Schulen, was in der öffentlichen Berichterstattung etwa in Deutschland oft übersehen wird. Externe Prüfungen sind die andere Seite der großen Freiheiten, die das System immer noch bietet.

Eine zentrale Position im niederländischen Test- und Prüfungssystem nimmt das unabhängig von den staatlichen Institutionen bestehende zentrale

⁶² Onderwijsraad (2002): De Kern van het doel. Reactie op het advies van de commissie Wijnen over de kerndoelen basisonderwijs. Den Haag: Onderwijsraad.

Institut für Testentwicklung (CITO – *Centraal Instituut voor Toetsontwikkeling*) ein. CITO, das 1968 gegründet wurde, arbeitete seitdem im Auftrag des Bildungsministeriums und ist im Wesentlichen für die Entwicklung von Prüfungen, Tests und ganzen Examenssystemen zur Beurteilung der Leistungen der Lernenden sowie von Unterrichtsprozessen und von Bildungssystemen verantwortlich. 1999 wurde das Institut privatisiert, auch um sich kundenorientierter ausrichten zu können. Es nimmt aber weiterhin Aufträge des Bildungsministeriums oder auch von anderen Bildungseinrichtungen und Behörden entgegen. CITO erstellt unter anderem jährlich Leistungstests für das Ende der Grundschulzeit sowie die in Teilen zentral formulierten *Abschlussprüfungen* zum Ende der Sekundarstufe. Diese wurden bereits 1973/74 eingeführt, sind obligatorisch und finden zu unterschiedlichen Zeitpunkten statt, je nach Schultyp:

- in der Vorbereitenden Mittlere Berufsbildung (VMBO) mit 16 Jahren,
- in der Höheren Allgemeinbildung (HAVO)⁶³ mit 17 Jahren und
- in der Vorbereitenden Wissenschaftlichen Bildung (VWO)⁶⁴ mit 18 Jahren.

Die Abschlussprüfungen setzen sich zu einem Teil aus schulinternen Prüfungen (*schoolexamen*), die sich an den zentral vorgegebenen Zielen orientieren, zusammen und zum anderen Teil von CITO (*Commissie examens voortgezet Onderwijs*) erstellten und zentral ausgewerteten Prüfungsaufgaben (*centraal examen*). Der schulinterne Prüfungsanteil kann einerseits als eine Art Mittelweg zwischen den staatlichen Vorgaben und der Schulautonomie aufgefasst werden, zum anderen bietet er auch die Möglichkeit einer schulspezifischen inhaltlichen Akzentsetzung. Das Institut befasst sich des Weiteren im Bereich der Schulevaluation mit der Organisation von Leistungsstudien und der Weiterbildung der Lehrenden.

Seit 1994 müssen mit dem Ende der Grundschulzeit den aufnehmenden Sekundarschulen Ergebnisse eines *Abschlusstestes*, des *eindtoets basisonderwijs* vorliegen (Akker v.d./ Letschert 2002, S. 48). Es liegt jedoch im Entscheidungsspielraum der jeweiligen durchführenden Grundschule, welchen Test sie anwenden möchte. Der genannte CITO-Test wird von ca. 80% der Schulen gewählt und existiert für die Fächer Niederländisch, Mathematik, Sachkunde und Informationsverarbeitung; dementsprechend wird immer wiederkehrend darüber diskutiert, diesen für alle Grundschulen verbindlich einzuführen. Ebenso wie der landesweiten Abschlussprüfung kann den CITO-Tests eine bedeutende Steue-

⁶³ Dieser Schulabschluss ist mit dem deutschen Realschulabschluss vergleichbar.

⁶⁴ Dieser Schulabschluss ist mit dem deutschen Abitur vergleichbar.

rungswirkung zugesprochen werden. Die Ergebnisse des Abschlusstests erhalten die Schulen in Form von schülerspezifischen Berichten nach drei Wochen. Folgende Funktionen des Tests lassen sich unterscheiden:

- Vorwiegend sollen durch den Test der individuelle Wissensstand und die individuellen Kompetenzen der Schülerinnen und Schüler aufgezeigt werden, um eine objektive Beurteilungsgrundlage für die Planung der weiteren Schullaufbahn zur Verfügung zu haben.
- Ferner sollen durch die Testergebnisse, die sowohl nationale als auch eine regionale Bezugsbasis haben, schulinterne Evaluationen angeregt und ermöglicht werden.
- Zudem sollen die Testresultate auch Eingang in ein Systemmonitoring finden.

Die jährlich neu aufgelegten Tests des CITO werden vom Institut selber in Zusammenarbeit mit Grundschullehrkräften ausgearbeitet und zudem durch entsprechende Fachexperten sowie durch Pre-Tests kontrolliert.

Neben den obligatorischen Prüfungen gibt es noch weitere Tests im niederländischen Schulsystem, deren Einsatz jedoch fakultativ ist. Die *diagnostischen Tests* werden ebenfalls von CITO entwickelt und dienen vor allem den Lehrkräften, um die Qualität des eigenen Unterrichts besser beurteilen zu können. Da sie landesweit Anwendung finden, können die Ergebnisse zudem als Vergleichsreferenz herangezogen werden. Daneben gibt es mit dem *leerlingvolgsysteem* (LVS) ein weiteres diagnostisches Instrument von CITO, das es ermöglichen soll, den Unterricht besser auf die individuellen Bedürfnisse der Lernenden abzustimmen. Insgesamt besteht das CITO-Testsystem aus aufeinander abgestimmten Testpaketen, die kontinuierlich eingesetzt werden können. Zudem erhalten die Lehrkräfte noch weitere Unterlagen und Materialien, die zum Beispiel Unterstützung bei der Analyse der Testergebnisse und deren Ursache bieten.

3.2.4 Qualitätssicherung und Evaluation

Eine zentrale Rolle besetzt hinsichtlich der Qualitätssicherung und der Entwicklung von Schulqualität im niederländischen Bildungssystem die staatliche Schulaufsicht (*Inspectie van het Onderwijs*). Das Schulinspektorat ist eine 1801 gegründete Aufsichtsbehörde, die seither kontrollierende, evaluierende sowie beratende Aufgaben im Schulwesen wahrnimmt. Geleitet wird das Inspektorat

durch einen Generalinspektor, der die Behörde mit vier Hauptinspektoren verwaltet; zusammen bilden sie den Rat der Inspektoren.

Die Schulaufsicht untergliedert sich in vier Abteilungen, Primarschulbildung, Sekundarschulbildung, Berufsausbildung und Erwachsenenbildung sowie Hochschulbildung, für die jeweils ein Hauptinspektor zuständig ist (Felber 2005, S. 10; Ackeren v. 2003a, S. 174). Die Aufgaben und Verpflichtungen der *Inspectie van het Onderwijs* werden in den Grundsätzen der Aufsichtsbehörde, in verschiedenen Bildungsgesetzen und in einem unlängst eingeführten Schulaufsichtsgesetz festgehalten. Das Gesetz über die Schulaufsicht bildet die Grundlage für eine unabhängige, fundierte und öffentliche Bewertung der Qualität der Schulen und des Schulsystems. Darin werden die Ausführungen zu den Aufgaben und Verpflichtungen der Schulaufsicht, zu den Qualitätsmerkmalen und dem Qualitätsrahmen, zum Prinzip der Verhältnismäßigkeit und die Ausführung des Unterstützungssystems gesetzlich festgehalten (Felber 2005, S. 11). Die Arbeit der Inspektoren ist bei den Lehrkräften akzeptiert und der Umgang mit den externen Evaluationen gehört zur professionellen Kompetenz. Die Schulaufsicht umfasst vier grundlegende Aufgabenbereiche:

- *Kontrolle* – Überprüfung der Einhaltung der gesetzlichen Bestimmungen durch die einzelnen Schulen;
- *Evaluation* – Darstellung der gegenwärtigen Situation im Bildungswesen, insbesondere hinsichtlich der Qualitätsentwicklung in den Schulen;
- *Förderung/Stimulierung* – Unterstützung der Schul- und Qualitätsentwicklung durch Beratung von Schulen und relevanten Behörden;
- *Berichterstattung und Beratung* – Berichterstattung an den Bildungsminister sowie auf Anfrage oder eigene Initiative beratende Funktion, zudem öffentliche Berichterstattung zur Schulqualität (Felber 2005, S. 14-17).

Das wichtigste Instrument des Inspektorats für den umfassenden Aufgabenbereich ist der Schulbesuch. Im Gesetz zur Schulaufsicht wird festgehalten, welche Qualitätsaspekte die Schulinspektion umfassen soll:

In Bezug auf die Lernleistungen:

- Lernergebnisse,
- Entwicklungsfortschritt der Schülerinnen und Schüler.

In Bezug auf die Struktur der Unterrichts- und Lernprozesse:

- Unterrichtsstoff,
- Lehr- und Lernzeit,

- Unterrichtsklima,
- Schulklima,
- didaktische Kompetenz der Lehrkräfte,
- Schülerbetreuung,
- Inhalt, Niveau und Implementationsgrad von Prüfungen, Tests, Übungen und Examen (Inspectie van het Onderwijs 2005, S. 5f.).

Das Verfahren zur Schuluntersuchung hat in den letzten Jahren verschiedene Umstrukturierungen durchlaufen, da es nicht mehr als realistisch galt, alle Schulen innerhalb eines Schuljahres besuchen und evaluieren zu können. Um jedoch zu vermeiden, dass Schulen möglicherweise über mehrere Jahre hinweg nicht beurteilt werden, was im englischen Inspektionssystem lange als kritisch galt, hat man Ende der neunziger Jahre ein mehrgliedriges Inspektionsverfahren entwickelt, das aus einer integralen Schulaufsicht (*integraal schooltoezicht – IST*), einer regulären Schulaufsicht (*regulier schooltoezicht – RST*) und einer proportionalen Schulaufsicht (*porportioneel schooltoezicht – PST*) bestand.

Die reguläre Schulaufsicht, die als schmalere Variante der integralen Inspektion galt, konzentrierte sich überwiegend auf die Evaluation der Schülerleistungen, der Unterrichtsorganisation, der Lernmaterialien und der Schülerbetreuung. Verlieft die RST negativ, folgte die integrale Inspektion, die auf der Grundlage eines umfassenden Indikatorensets die gesamten Lehr-Lern-Prozesse, die erreichten Ergebnisse und die Schulbedingungen bewertete. Die PST zielte auf die Förderung der Selbstevaluation der Schulen ab, indem das Verhältnis zwischen der externen Evaluation durch die Schulaufsicht und der internen Evaluation proportionaler gestaltet wurde. Je besser das System der Selbstevaluation innerhalb der Schule entwickelt war, desto weniger bedurfte es der externen Inspektion (vgl. dazu Felber 2005, S. 15 f., Ackeren v. 2003a, S. 176ff.). Dieses Inspektionsverfahren wurde im letzten Jahr wieder durch die *Inspectie van het Onderwijs* (2005) reformiert. Es differenziert sich nun in fünf Inspektionstypen:

- die jährliche Inspektion,
- die regelmäßige Qualitätsinspektion,
- die weitere/ergänzende Inspektion,
- die qualitätsverbessernde Inspektion und
- die außerplanmäßige Inspektion.

Der herkömmliche Zyklus für die Evaluation der Einzelschule sieht eine Inspektion alle vier Jahre vor. Für die durchschnittlichen und besseren Bildungseinrichtungen gestaltet sich der Inspektionszyklus wie folgt (vgl. *Inspectie van het Onderwijs* 2005, S. 11-15):

- (1) Jede Schule wird alle vier Jahre durch die regelmäßige Qualitätsinspektion beurteilt.
- (2) In den zwei darauffolgenden Jahren werden die Leistungen der betreffenden Schule jährlich durch eine Dokumentenanalyse im Inspektorat evaluiert. Die Ergebnisse dieser Analyse können gegebenenfalls zu einer Anpassung der Qualitätskarte (s.u.) einer Schule führen.
- (3) Im Jahr vor der regelmäßigen Qualitätsinspektion wird die jährliche Inspektion in Form eines Schulbesuchs durchgeführt. Gegenstand der Evaluation sind sowohl die Leistungen als auch die Qualitätssicherung der Schule. Auf der Grundlage dieser Beurteilung wird zusammen mit der Einzelschule festgelegt worauf sich die umfassende Inspektion, im nächsten Jahr, neben dem Indikatorenset thematisch konzentrieren soll.

(1) Jährliche Inspektion

Die jährliche Inspektion (annual inspection) unterteilt sich in die Dokumentenanalyse und -auswertung und in den Schulbesuch. Informationsgrundlage für die Beurteilung der Schulen durch die Auswertung von Dokumenten ist ein Fragebogen, der alljährlich durch die Schulen zu beantworten ist. Der Schulbesuch im Vorfeld der regelmäßigen Qualitätsinspektion dient einerseits der Vorbereitung der Evaluation, andererseits wird den Schulen in Gesprächen die Möglichkeit gegeben, mit dem Inspektorat die Schulentwicklung und die Entwicklung des Umfeldes zu diskutieren.

(2) Regelmäßige Qualitätsinspektion

Von essentieller Bedeutung für die Arbeit der Aufsichtsbehörde ist gegenwärtig die regelmäßige Qualitätsinspektion (periodical quality inspection – PQI), die auf der Grundlage eines Basissatzes von Indikatoren durchgeführt wird (*Inspectie van het Onderwijs* 2005, S. 6).⁶⁵

⁶⁵ Die folgende Auflistung des Indikatorensets ist nicht vollständig, sie soll vielmehr beispielhaft die Gestaltung der Indikatoren aufzeigen.

The basic set of indicators

Curriculum

- The subject matter taught for the Dutch language and for arithmetic and mathematics covers the core objectives.
- The subject matter for the Dutch language and for arithmetic and mathematics is taught to a sufficient number of pupils up to and including the level of year 8.
- The subject matter in the different years is linked.
- The subject matter for the Dutch language and arithmetic and mathematics is geared to the educational needs of individual pupils.
- The school with a substantial number of pupils with a pupil weight offers lessons materials for the Dutch language that is consistent with the educational needs of pupils with a language deficiency.
- The school has subjects aimed at improving social integration and citizenship, including the transfer of knowledge about and an introduction to the diversity in society.

[...]

Didactic performance of teachers

- The teachers create a task-oriented work climate.
- The teachers explain everything clearly.
- The teachers explicitly teach strategies for thinking and learning.

[...]

Achievements; results of pupils and their progress in development

- The results of the pupils at the end of the school period are at least in a level that may be expected from characteristics of the pupil population.
- The social skills of the pupils are on a level that may be expected.
- The results of the pupils for the Dutch language and for arithmetic and mathematics during the school period are at least on the level that may be expected on the basis of the characteristics of the pupil population.
- Pupils with specific educational needs develop according to their capabilities.
- In principle, the pupils pass through the school within the expected period of 8 years.

[...]

(Inspectie van het Onderwijs 2005, S. 18 ff.)

Das Indikatorenset ist nicht statisch, sondern befindet sich in einer konstanten Entwicklung, um sich den Veränderungen innerhalb der Gesellschaft oder neuen bildungspolitischen Entscheidungen anpassen zu können. Dementsprechend kann auch das Set der Indikatoren sich verändern. Es ist auch möglich, ein speziell auf die Bedürfnisse der Einzelschule abgestimmtes Indikatorenset zu erstellen.

Neben dem Basissatz an Indikatoren gibt es noch weitere Indikatoren, die speziell für die Beurteilung der Qualitätssicherung an der Schule entwickelt wur-

den. Diese ermöglichen es dem Inspektorat während des Schulbesuchs gezielt zu bewerten, wie die Einzelschule die Unterrichtsqualität sichert und weiterentwickelt (Inspectie van het Onderwijs 2005, S. 7f.).

The quality control indicators

Quality control: the school systematically preserves and improves the quality of its education.

- The school understands the differences in the educational needs of its pupil population.
- The school systematically evaluates the quality of its achievements every year.
- The school regularly evaluates learning and education.
- The school works systematically on activities to initiate improvements.
- The school guarantees the quality of the learning and education.
- The school transparently reports to interested parties on the quality of education and learning being realised.
- The school guarantees safety for pupils and personnel.

(Inspectie van het Onderwijs 2005, S. 20).

Im Beurteilungsrahmen der regelmäßigen Qualitätsinspektion sind zudem in Anlehnung an die Indikatoren Qualitätsstandards formuliert, mittels derer entschieden wird, ob die Qualitätsrichtlinien erfüllt sind oder nicht (Inspectie van het Onderwijs 2005, S. 8f.).

(3) Weitere Inspektion

Wenn bei der regelmäßigen Inspektion die Qualität einer Schule zweifelhaft ist, folgt eine weitere Inspektion (further inspection), die sich an den Schwächen und Problemen der vorangegangenen Evaluation orientiert. Entspricht die Schule den Qualitätsansprüchen nicht, wird die qualitätsverbessernde Inspektion eingeleitet.

(4) Qualitätsverbessernde Inspektion

Wird eine qualitätsverbessernde Inspektion durchgeführt, wird der betreffenden Schule zunächst in einem Zeitraum von maximal zwei Jahren die Möglichkeit gegeben die Qualität der Schule auf ein akzeptables Niveau zu heben. Der Schulinspektion ist es jedoch vorbehalten diesen Zeitraum zu verkürzen, falls die Inspektion aufgrund der unzulänglichen Situation schneller handeln möchte. Während der Verbesserungsphase wird die Schule weiterhin durch die Aufsichtsbehörde beobachtet, die ihre Untersuchungsergebnisse dokumentiert. Ziel dieser Inspektion ist die Beurteilung der Schule durch das Indikatorenset, unter Berücksichtigung der Fortschritte, die die Schule während der Verbesserungsphase gemacht hat.

(5) Außerplanmäßige Inspektion

Infolge einer gravierenden Beschwerde, der Aufdeckung von ernsthaften Informationen, einer Anforderung des Bildungsministers oder von Medienberichten kann eine außerplanmäßige Inspektion (incidental inspection) einer Schule durchgeführt werden.

Die Ergebnisse der Selbstevaluation der Einzelschule nehmen weiterhin eine bedeutende Position bei Schulinspektionen ein. Stellt eine Schule der Aufsichtsbehörde vertrauenswürdige Daten aus der internen Evaluation zur Verfügung, gehen diese mit in die Beurteilung ein, wodurch sich die Inspektion weniger aufwändig gestaltet.

Für die öffentliche Berichterstattung über die Qualität der niederländischen Schulen ist ebenfalls das Schulinspektorat verantwortlich. Die Veröffentlichung der Ergebnisse der Schuluntersuchungen fungiert als Informations- und Diskussionsgrundlage für die Schulen, Schülerinnen und Schüler, Eltern und andere Interessierte.

Qualitätskarten

Darüber hinaus veröffentlicht das Schulinspektorat jährlich eine Qualitätskarte (*kwaliteitskaart*) im Internet,⁶⁶ die auf Grundlage der Informationen und Schlussfolgerungen des Inspektorats strukturierte Informationen zu jeder einzelnen Schule und ihrer Qualität bereithält und somit weitere Transparenz ermöglicht. Die Qualitätskarte beinhaltet allgemeine Informationen zur jeweiligen Bildungseinrichtung: Anzahl der Schülerinnen und Schüler, die Verteilung der Lernenden auf die einzelnen Schulzweige, Anzahl der Wechsel der Schulformen, Anzahl Klassenwiederholungen und die Examensresultate. Diese Daten fungieren für die Schulen als Vergleichswerte mit anderen Schulen und für die Eltern, Schülerinnen und Schüler als Entscheidungshilfe bei der Schulwahl.

In den Niederlanden sind die Freiheit der Bildung und die Autonomie der Schule traditionell von großer Bedeutung gewesen. Die letzten beiden Jahrzehnte waren nochmals durch weitere und tiefgehende Dezentralisierungs- und Deregulierungsprozesse gekennzeichnet. Es ist jedoch nicht unerheblich zu differenzieren, auf welchen Ebenen und für welche Bereiche dezentralere Steuerungsmechanismen eingesetzt wurden. Doolaard und Kaestanje (2001), zwei niederländische Bildungsexperten, zeigen auf, dass hinsichtlich der finanziellen

⁶⁶ Ansicht unter: <http://www.kwaliteitskaart.nl/kwaliteit2003/> (Stand 12.07.07).

Verantwortungsübernahme den Schulen durch die Pauschalfinanzierung größere Freiräume eingeräumt wurden.

Das gilt jedoch nicht für den pädagogischen Bereich. Das Bildungsministerium hat diesbezüglich verschiedenste Entscheidungskompetenzen „nur“ bis auf die Mesoebene, also die mittlere Verwaltungsebene verschoben, woran die Einzelschule nicht direkt Anteil hat. Auch van Ackeren weist darauf hin, dass „neben den zentral vorgegebenen Elementen einer Wirkungssteuerung [...] sich momentan auch eine Beschneidung bislang zuerkannter schulischer Freiheiten im Bereich der Kontext- und Prozess-Steuerung [zeigt]“ (2003, S. 106): „In the curricular domain there has been a gradual move towards centralisation“ (Hendriks/Doolaard/Bosker 2001, S. 1).

3.3 Schweden

Das grundlegende Ziel der schwedischen Bildungspolitik ist es, allen Schülerinnen und Schülern einen gleichberechtigten Zugang zu gleichwertiger Bildung zu ermöglichen, unabhängig von Geschlecht, Wohnort und sozialen wie auch finanziellen Bedingungen (Daun 2004, S. 62). Die Politik ist geprägt von den egalitären Grundsätzen des schwedischen Wohlfahrtsstaates. Bereits 1842 ist in Schweden eine vierjährige obligatorische Primarschule eingeführt worden, die danach sukzessive ausgebaut wurde. Sie ist die Basis für den Grundsatz der gleichen Bildung für alle Kinder. „A hallmark of the Swedish welfare state and certainly also its education policy is a firm and overriding preference for universal and equal entitlements of citizenship, as opposed to targeting policy to those in need“ (Nicaise/Esping-Andersen/Pont et al. 2005, S. 14).

Diese ursprünglich sozialdemokratische Position markiert immer noch einen starken bildungspolitischen Konsens, selbst wenn seit Herbst 2006 eine neue konservativ-liberale Regierung im Amt ist. Anders als in den Vereinigten Staaten stehen sich nicht zwei Lager mit grundsätzlich anderen Erziehungsphilosophien gegenüber. Das Prinzip der Gleichheit im Bildungswesen ist offiziell nicht strittig, auch wenn zu Beginn der neunziger Jahre durch neo-liberale Politikelemente eine gewisse Polarisierung eingesetzt hat (Lundahl 2002, S. 626). Im OECD-Länderbericht *Equity in Education* werden drei Dimensionen des Prinzips der Gleichheit unterschieden, nämlich Chancengleichheit im Sinne von „*equivalent outcomes*“, die Gewähr sozialer Sicherheit (*trygghet*) sowie eine starke Individualisierung aller pädagogischen Maßnahmen (Nicaise/Esping-Andersen/Pont et al. 2005, S. 14-15).

In den internationalen Vergleichsstudien der OECD gilt das schwedische Bildungssystem als sehr leistungsfähig. Zwischen den einzelnen Schulen bestanden bis vor kurzem geringe Leistungsunterschiede und die Qualität der Abschlüsse ist nach wie vor hoch. Studien aus den neunziger Jahren zeigen, dass der Einfluss der sozialen Herkunft auf den Bildungserfolg zu diesem Zeitpunkt vergleichsweise schwach war (Shavit/Blossfeld 1993). Dazu trug nicht nur die schwedische Gesamtschule bei; gewichtige Faktoren waren auch die niedrige Armutsrate, die Einkommenssicherheit in Familien mit Kindern (Nicaise/Esping-Andersen/Pont/Turnstall 2005, S. 15) und die damalige Homogenität der Bevölkerung. Hinzu kommt die frühe Förderung und eine bildungspolitische Strategie, die auf Integration und nicht auf Selektion angelegt ist. Im Equity-Bericht der OECD wird diese Strategie so beschrieben:

„Swedish policy applies two main instruments in trying to achieve equivalent outcomes. One stresses ‚demand-side‘ policies aimed at eliminating barriers to, and inequalities of, educational attainment, in particular with regard to credit constraints and resource weaknesses in families. Not only is family support unusually generous but access to education is free throughout the system, from first grade to higher education. The second applies to educational system design. Swedish policy seeks to eliminate dead-ends and tracking of students through a large menu of second-chance options and bridges to higher levels. The stated aim is that *all* children will complete the equivalent of upper secondary school and that *all* be given the opportunity to pursue tertiary studies if they so desire“ (ebd., S. 14/15).

Dieses System ist allmählich entwickelt worden. Im 19. und in der ersten Hälfte des 20. Jahrhunderts war das schwedische System im Sekundarbereich vertikal gegliedert, nach dem Zweiten Weltkrieg löste die einheitliche *grundskola*⁶⁷ alle anderen Schulformen ab, seit 1972 besteht eine neunjährige Schulpflicht für alle Kinder, die bis zur Einführung von unabhängigen Schulen in der Gesamtschule verbracht wurde. Die bildungspolitischen Reformen seit 1990 haben drei Richtungen: Verlagerung der Kompetenz nach unten, neue Formen der Systemsteuerung und Einführung von reguliertem Wettbewerb.

⁶⁷ Die schwedische *grundskola*, die dem Modell einer Gesamtschule entspricht, umfasst neun Schuljahre. Sie entspricht damit nicht dem deutschen Grundschulmodell, das nur die ersten vier Schuljahre vorsieht. Im Folgenden wird weiterhin der schwedische Begriff *grundskola* verwendet, um keine falschen Assoziationen mit dem deutschen Verständnis von Grundschule aufkommen zu lassen.

3.3.1 Der Wandel der Bildungspolitik

Das Postulat der Egalität seit Gründung der *folkskola* (1842) und seit Einführung des gleichen Einschulungsalters mit sieben Jahren (1858)⁶⁸ wurde mittels einer zentralstaatlichen Steuerung des schwedischen Schulwesens realisiert. Bis weit in die achtziger Jahre des 20. Jahrhunderts hinein verfügte Schweden über zentralistisch organisierte Verwaltungsstrukturen, die auch für das Bildungssystem vorgegeben waren. Durch tiefgreifende Umstrukturierungen erfolgte Ende der achtziger Jahre jedoch ein schneller und intensiver Dezentralisierungs- und Deregulierungsprozess (Kotthoff 2003, S. 196).

Für die Abkehr von der seit Mitte des 19. Jahrhunderts aufgebauten und lange Jahrzehnte vorherrschenden zentralistischen Steuerung des Bildungssystems lassen sich vier allgemeine Beweggründe anführen:

- (1) *Ökonomische Gründe*: Die wirtschaftlich angespannte Situation nach der Ölkrise von 1973 führte auch zu Verkürzungen der Ausgaben im schwedischen Bildungswesen und verlangte nach einer effizienteren und effektiveren Steuerung des Schulsystems.
- (2) *Organisatorisch-administrative Gründe*: Die zentralistische Steuerung des Bildungssystems und die stets top-down vorgenommenen Schulreformen und deren Auswirkungen wurden zunehmend kritischer betrachtet und fanden immer weniger Zuspruch innerhalb der Bevölkerung.
- (3) *Pädagogische Gründe*: Die mit der Dezentralisierung und Deregulierung einhergehenden Erweiterungen der Entscheidungskompetenzen der Kommunen und Schulen sollten es ermöglichen, nachdrücklicher auf die Interessen und Bedürfnisse der einzelnen Schulen und ihrer Schülerinnen und Schüler einzugehen.
- (4) *Ideologisch-politische Gründe*: Die Hinwendung zu dezentralen Steuerungsmechanismen ging einher mit einer verstärkten Marktorientierung, die im Schulwesen eine freie Schulwahl (*free choice*), Wettbewerb zwischen den Schulen (*competition*) und die Förderung privater Schulen (*privatization*) beinhaltete (Kotthoff 2003, S. 207ff.).

Zuvor verfügte die schwedische Schulverwaltung über eine komplexe Organisationsstruktur; die wichtigsten Behörden waren das Bildungsministerium (*utbildningsdepartement*), die Generaldirektion für das Schulwesen (*skolöverstyrelsen*), der die Provinzialschulämter (*lämsskolnämnder*) untergeordnet waren, und

⁶⁸ Die Schulpflicht beginnt in Schweden bis heute mit sieben Jahren.

schließlich auf der untersten Verwaltungsebene die Gemeinden (*kommuner*). Die schwedischen Bildungsbehörden besaßen aufgrund dieser zentralistisch und bürokratisch ausgerichteten Struktur vielfältige Möglichkeiten, steuernd und regulierend in das Bildungswesen einzugreifen. Im internationalen Vergleich zeigte sich jedoch, dass die Lehrer bzw. Schulleiter über eine verhältnismäßig hohe Entscheidungsfreiheit etwa im Blick auf curriculare Entscheide verfügten. Die Binnensicht war etwas anders gelagert, Schulpraktiker sahen sich eher reguliert und die Vertreter der Schulverwaltung behaupteten die Wirksamkeit ihrer Erlasse (ebd., S. 198 ff).

Schon nach Einführung der neunjährigen Gesamtschule 1962 wurden Bedenken gegenüber der zentralen und bürokratischen Steuerung des schwedischen Bildungswesens geäußert, die sich bis in die siebziger Jahre fortsetzten und beständig durch Überlegungen zur Dezentralisierung der Schulverwaltung begleitet wurden. Eine durch die schwedische Regierung eingesetzte parlamentarische Kommission sprach sich in ihrem 1974 veröffentlichten Abschlussbericht mit dem Titel „Das Arbeitsmilieu der Schule“⁶⁹ für eine tiefgreifende Neugestaltung der schwedischen Bildungsverwaltung aus. Ausgehend von diesen Bestrebungen deutete sich Ende siebziger Jahre die Umstrukturierung der zentralistischen Schulverwaltung und eine Ablösung der *Top-down-Strategie* der schwedischen Schulreformen an, die maßgeblich durch Absichten zur Dezentralisierung bestimmt wurden (ebd., S. 200f.).

Wesentlich für diese Entwicklungen war der Widerspruch zwischen Egalität und Zentralismus. Der gleiche Zugang zu einer gleichwertigen Ausbildung bildet bis heute eines der grundlegendsten Ziele der schwedischen Bildungspolitik und ist tief im Bewusstsein der Bevölkerung verankert. Dementsprechend wuchs die Skepsis der Öffentlichkeit und auch des Staates gegenüber zentralistischen Steuerungsmechanismen (ebd., S. 203). Wichtig war auch die Erkenntnis, dass die *grundskola* in der bestehenden Form zu geringe Auswirkungen auf die Erhaltung und Förderung sozialer Gleichheit hatte. Es folgten zahlreiche Studien, Berichte und mehrere parlamentarische Untersuchungsausschüsse, die die zentralstaatliche Steuerung, Machtverteilungen und den Einfluss der Bevölkerung innerhalb des Staates analysierten.

1988/89 wurde durch das schwedische Parlament eine neue und radikal veränderte Aufteilung der Entscheidungskompetenzen zwischen dem Staat und den Kommunen beschlossen, die auch die Schulen betraf. In der Folge über-

⁶⁹ SOU (Statens Offentliga Utredningar): Skolans arbetsmiljö: Betänkande avgiven av Utredningen om skolans inre arbete – SIA. Stockholm: Allmänna Förlaget 1974.

trug die Regierung den Gemeinden die alleinige Verantwortung für das Personalmanagement und die Verwaltung des Budgets (ebd., S. 204). Damit verbunden waren einschneidende Änderungen für die Lehrkräfte und die Schulorganisation. Das schwedische Lehrpersonal, das zuvor durch den Staat angestellt war, erhielt mit den Kommunen neue Arbeitgeber, die die Arbeitsbedingungen im Rahmen der Gesetze auf ihren Bedarf und ihre Möglichkeiten zuschneiden konnten.

Im Hinblick auf die finanzielle Verantwortung erhalten die Kommunen ein Gesamtbudget zur Finanzierung ihres Schulwesens, das durch die jeweiligen kommunalen Steuereinnahmen ergänzt wird. Da die finanziellen Ausstattungen der Gemeinden aufgrund ihrer unterschiedlichen Einwohnerzahlen, Einnahmen durch Gewerbesteuer etc. divergieren, werden staatliche Ausgleichssubventionen vergeben. Zudem wurde 1993 die an bestimmte Verwendungszwecke gebundene zentralstaatliche Vergabe der Mittel durch eine schulbezogene Pauschalfinanzierung abgelöst (ebd., S. 211f.). Die Kommunen konnten damit über die Aufteilung des Budgets selbst bestimmen, wodurch unter anderem eine bessere Orientierung an den Situationen der einzelnen Schulen möglich wurde.

Die Kommunalisierung ging einher mit einer Stärkung der einzelnen Schulen. Sie erhielten neue Kompetenzen. So wird die Einstellung neuer Lehrerinnen oder Lehrer meist von den Schulleitern übernommen, denen es innerhalb vorgegebener Richtlinien sogar obliegt, mit den Bewerbern einen spezifischen leistungsbezogenen Lohn auszuhandeln (ebd., S. 212). Ein weiterer bedeutender Anhaltspunkt für die zunehmende Dezentralisierung und Deregulierung war 1980 die Einführung eines neuen nationalen Schulcurriculums für die *grundskola*, das Mitte der neunziger Jahre nochmals revidiert wurde und die Autonomie der Kommunen und der einzelnen Schulen grundlegend ausweitete. Das Bildungsministerium verzichtete bei der Ausarbeitung des neuen Curriculums auf detaillierte inhaltliche Vorgaben und verfasste stattdessen nur übergreifende Ziele, wodurch die konkrete Ausgestaltung des Lehrplans den Kommunen, den einzelnen Schulen, Schulleitern und Lehrern anvertraut wurde.

Die früheren Curricula von 1969 und vor allem 1962 sind mit einem sehr hohen Detaillierungsgrad ausgearbeitet worden und stellten komplexe Lehrwerke dar, die mit umfangreichen Vorschriften verknüpft waren, ohne Hinweise auf die Wirksamkeit zu enthalten, wie dies seinerzeit international üblich war. Curriculumrevision ging immer mit einer Erweiterung der Lehrpläne einher. Die Curricula von 1980 und vor allem von 1994 sind schmale Rahmenrichtlinien, die den Schulen größere individuelle Entscheidungsspielräume einräumten und die auf dem grundsätzlich neuen Mechanismus der Zielsteuerung

basierten (ebd., S. 202). In Interviews mit den Entscheidungsträgern herrscht nachträglich große Einigkeit über diese „transition from *governance by rules* to *governance by objectives*“ (Lundahl 2002, S. 628).

Bedeutsam für die Dezentralisierung waren auch die Abschaffung der Provinzialschulämter (*länskskolnämnder*) auf der mittleren Verwaltungsebene (1990) und die Ablösung der mit weitreichenden Befugnissen ausgestatteten und ebenfalls zentralstaatlich organisierten Generaldirektion für das Schulwesen (*skolöverstyrelsen*) durch die in allen Bereichen deutlich schlanker strukturierte Nationale Behörde für das Bildungswesen (*skolverket*) (1991). Damit wurde die Zuständigkeit und Entscheidungskompetenz drastisch verlagert, in Richtung Profession (ebd., S. 626) und weg von der Bürokratie. Eingeleitet wurde dieser Wandel noch von der Sozialdemokratie; die erste nicht-sozialistische Regierung unter Carl Bildt forcierte 1991 das Tempo und brachte auch Marktelemente ins Spiel (ebd.).⁷⁰

Daniel Kallós und Lisbeth Lundahl-Kallós, zwei schwedische Bildungsexperten, skizzierten die eng an Dezentralisierungs- und Deregulierungstendenzen gebundene Reform des Steuerungskonzepts des schwedischen Bildungswesens wie folgt:

„Decentralization was regarded as a means to accomplish what earlier centralized measures had failed to ensure, that is, a nationally common comprehensive school where equal opportunities should be guaranteed by local (municipal) adaptation of the school organization and the inner work of schools. Each municipality should politically decide upon a ‚municipal school plan‘ (in Swedish – *kommunal skolplan*) that is, create a local strategy on how to achieve the national objectives [...]. The plan should then be supplemented by the local work plan of each school“ (Kallós/Lundahl-Kallós 1994, S. 143).

Mit Antritt der bürgerlichen Regierung Bildt wurde eine zusätzliche Wende in der Bildungspolitik eingeleitet. Die gezielte staatliche Förderung privater Schulen sollte die Monopolstellung des Staates im Schulwesen abschwächen, den Wettbewerb zwischen den einzelnen Schulen fördern und den Eltern eine freie Schulwahl ermöglichen. Das Ziel war, die Qualität und Produktivität im Schulsystem zu erhöhen. „Privat“ meint in diesem Zusammenhang „unabhängig“:

⁷⁰ Bereits Finanzminister Kjell-Olof Feldt im zweiten Kabinett Olof Palme (1982) vertrat neo-liberale Standpunkte. Seine Kritik am schwedischen Sozialstaat war der Grund für heftige Auseinandersetzungen mit den linken Gruppierungen in der Sozialdemokratie und den Gewerkschaften; Feldt musste 1990 zurücktreten.

Die Regierungspolitik förderte Schulen außerhalb des staatlichen Systems, die sich als selbstständige Anbieter entwickeln konnten, jedoch gleichfalls vom Staat finanziert wurden. In diesem Sinn herrscht eine echte Wahlgleichheit. Die Zahl der Privatschulen ist gestiegen, seitdem diese Möglichkeit besteht. Sie stehen jedermann offen, weil sie nicht an religiöse oder weltanschauliche Voraussetzungen gebunden sind, wie dies in anderen Ländern mit *school choice* der Fall ist. 1991 besuchte knapp ein Prozent der schulpflichtigen Kinder und Jugendlichen diese Schulen, 2002 waren es bereits 5,7 Prozent; 2006 sind es knapp 7% in der *grundskola* und knapp 12% in der *gymnasieskola*.⁷¹ In der Region Stockholm betrug die Zahl bereits vorher rund 20% (Nicaise/Esping-Andersen/Pont et al. 2005, S. 19). Um die Größenordnung zu verdeutlichen: Im gleichen Jahr 2002 besuchten 13 Prozent der amerikanischen Schulkinder Privatschulen oder Charter Schools.

In Umfragen unterstützen 90 Prozent der Eltern das Prinzip der freien Schulwahl. Schon unmittelbar nach der Einführung von Wahlmöglichkeiten im Jahre 1993 entschieden sich knapp 20 Prozent der Eltern für eine andere Schule als die ihres Wohnortes (ebd.). Bis 2005 sind verschiedene Variablen dieses Prozesses untersucht worden. Einige Studien beschreiben höhere Effizienz und gesunkene Kosten, was von anderen Studien bestritten wird. Die durchschnittlichen Testleistungen der Schülerinnen und Schüler sind seit Mitte der neunziger Jahre leicht gestiegen; „[...] by and large the newly decentralized, choice-driven system has had small positive effects on the academic performance of most Swedish students“ (Björklund/Clark/Edin et al. 2005, S. 10). Verschiedene andere Untersuchungen haben die Auswirkungen des Wettbewerbs zwischen den Schulen auf das Leistungsverhalten der Schülerinnen und Schüler untersucht.

Ahlin (2003) untersucht die Leistungen in den drei Kernfächern des schwedischen Curriculums bezüglich Auswirkungen der Schulwahlfreiheit. Die Resultate sind wie folgt: In Mathematik gibt es einen statistisch signifikanten positiven Effekt zwischen Wettbewerb und Leistungssteigerung, in den Sprachfächern (Schwedisch und Englisch) gibt es einen solchen Effekt nicht. „A possible explanation might simply be that in a more competitive environment, a higher priority is given to mathematics at the expense of other subjects such as English or Swedish“ (ebd., S. 24). Wenn Wettbewerb mit den Charakteristika

⁷¹ National summary sheets on education systems in Europe and ongoing reforms 2006 Edition: Sweden. URL: http://www.eurydice.org/ressources/eurydice/pdf/047DN/047_SE_EN.pdf (Stand 26.04.07).

der Schülerinnen und Schüler in Verbindung gebracht wird, dann zeigt sich, dass Kinder mit Immigrationshintergrund⁷² und Kinder mit speziellen Bedürfnissen (children in need of special education) mehr als andere vom Wettbewerb profitieren. Gegenteilige Effekte zeigen sich bei Kindern aus bildungsfernen Familien (low education families). Allerdings, wie mit quantilen Regressionen gezeigt wird, gibt es keinen signifikanten Unterschied in den Auswirkungen des Wettbewerbs auf leistungsschwache und leistungsstarke Schülerinnen und Schüler (ebd., S. 23). Auf der anderen Seite besteht die Gefahr einer Noteninflation dort, wo Schulen immer mehr um Schülerinnen und Schüler konkurrieren. Cliffordsson (2004) stellt allgemein fest, dass das Notenniveau in den Jahrgangsstufen 10-12 in den vergangenen Jahren gestiegen ist, allerdings ohne dass dies Folgen für die Leistungen/Noten in den Hochschulen und an den Universitäten gehabt hätte (zit. nach Erickson/Lander 2007, S. 35).

Im Blick auf die Freistellung der Schulwahl kommt eine Studie der schwedischen Nationalen Erziehungsagentur zum Schluss, dass vor allem Mittelschichteltern in urbanen Milieus mit hoher eigener Bildung davon Gebrauch machen. Die Mehrheit der Eltern und viele Repräsentanten der lokalen Autoritäten gehen davon aus, dass die Möglichkeit der Schulwahl positiven Druck auf die Schulen ausübt und sie kundenfreundlicher werden lässt. Gleichzeitig verstärkt die Wahl die Qualitätsunterschiede zwischen den Schulen. Die Segregation nimmt zu, auch weil mit der Wahl urbaner Mittelschichteltern bestimmte normative Vorstellungen einer guten Schule und eines erfolgreichen Schülers verbunden sind. Die Gleichwertigkeit (equivalence) des Bildungsangebotes nimmt ab, wenn nur bestimmte Angebote nachgefragt werden und dafür die freie Wahl besteht (The Swedish National Agency for Education 2003).

Es ist auch bestätigt, dass die Leistungsanforderungen der unabhängigen Schulen höher sind, aber unklar ist, ob das der Qualität des Unterrichts oder der zunehmenden Selektion in Privatschulen zugeschrieben werden kann. In der Studie von Björklund/Edin/Fredriksson et al. (2004) sind es Kinder bestimmter Migrantengruppen⁷³ und wiederum Schülerinnen und Schüler aus einkommensschwachen Familien (Low-SES), die vom Qualitätszuwachs nicht pro-

⁷² Das betrifft das Fach Mathematik. Die Migrantenkinder stammen aus nicht-nordischen und Nicht-EU-Ländern (Ahlin 2003, S. 23).

⁷³ 1998 waren 19,7 Prozent der Gesamtbevölkerung Schwedens (1.746.000 von 8.852.000) ausländischer Herkunft. Zum Einwanderungsland wurde Schweden erst nach dem Zweiten Weltkrieg, begünstigt auch durch liberale Asylgesetze. Traditionell wanderten vor allem Finnen ein. Diese Zahl sinkt aber ständig, während die nicht-nordische und auch die nicht-europäische Migration zunimmt.

fitieren. Verglichen zumindest mit den Vereinigten Staaten sind die Segregationseffekte bislang zwar nicht wirklich bedrohlich, während aber auf der anderen Seite die Marktelemente kaum die hohen Erwartungen ihrer Befürworter erfüllt haben. „[...] the effects of the education reforms have been exaggerated by both sides“ (ebd.).

Dass die freie Schulwahl die Segregation verstärkt, konnte für Stockholm nachgewiesen werden (Söderström/Uusitalo 2005, 2005a). Zudem fallen im öffentlichen Bereich Schulen mit einem hohen Anteil von Kindern bestimmter Migrantengruppen und/oder Schulen in einkommensschwachen Gebieten zurück. Private Eliteschulen haben gleiche staatliche Finanzierung eingeklagt und sind vom Gericht bestätigt worden, weil die Kriterien für die bevorzugte Finanzierung benachteiligter Schulen unklar waren (Nicaise/Esping-Andersen/Pont et al. 2005, S. 21). Erstere Schulen werden von Eltern mit einem hohen Bildungsniveau nachgefragt, aber auch von bestimmten Migrantengruppen.

Die Prozesse der Dezentralisierung und Deregulierung sind nicht so zu verstehen, dass allein noch die schulische Ebene zuständig ist. Wohl sind die zentralstaatlichen Vorgaben für die Einzelschule seit 1991 erheblich reduziert worden. Dadurch entstanden erweiterte Spielräume für individuelle Profile der einzelnen Schulen, die sich auch in größeren Kompetenzen ausdrückten. Die Zuständigkeit und Verantwortung für das Curriculum, das Budget, die Organisation und die Personalentwicklung sind jedoch nur teilweise bis auf die schulische Ebene delegiert worden. Durch vielfältige Maßnahmen und Initiativen zur Dezentralisierung ist vor allem die kommunale Ebene in Schweden erheblich gestärkt worden.

Dezentralisierung ist so *Kommunalisierung*, was nicht notwendig zu mehr Selbstständigkeit für die Einzelschule führen muss. In Schweden hatte die Kommunalisierung auch erweiterte Kontrollen für die Schule zur Folge, die die frühere unregulierte Autonomie erheblich einschränken. Die kommunalen Instanzen haben eine wesentlich intensivere Verbindung zu den einzelnen Schulen als die zentralstaatlichen Organisationen (Broadfoot 1996, S. 118). Die Kommunen können sich stärker um ihre Schulen kümmern, aber kontrollieren sie damit auch intensiver als zuvor die Behörden in Stockholm. Schließlich ist es nicht uninteressant festzustellen, dass sich in den letzten Jahren – wenn auch auf der Basis eines neuen Systems – erneute Zentralisierungsbestrebungen finden. Ablesen lassen sich diese einerseits auf curricularer Ebene (Kuiper/Akker v.d./Hooghoff et al. 2005), aber auch an der Einführung eines neuen staatlichen Inspektorats als Teil von *Skolverket*. Zugleich beobachtet man auch, dass der Staat die Schulen auffordert, eigene Konzepte und Pläne für bestimmte Bereiche zu entwickeln, gleichzeitig macht er aber ziemlich genaue Vorgaben, wie

solche Konzepte aussehen sollen. Mehr Autonomie, dies dürfte eine schwedische Erfahrung sein, setzt das Gleichheitsgebot in Gefahr, da lokale Kapazitäten nie gleich ausgeprägt sind.

Für eine Einschätzung der Effekte der schwedischen Dezentralisierungs- und Deregulierungsbestrebungen muss zum einen berücksichtigt werden, dass Reformen lange Zeit in Anspruch nehmen, weil sie verschiedene Ebenen erreichen und umgestalten müssen. Zum anderen ist die Differenzierung zwischen den Wirkungen, die durch die Dezentralisierungs- und Deregulierungsprozesse angestoßen wurden, und den Auswirkungen, die durch andere Reformen oder gesellschaftliche Konstellationen entstanden sind, nicht unproblematisch. Gleichwohl kommt Kotthoff auf der Grundlage nationaler Untersuchungen und eigens durchgeführter Interviews mit Vertretern der Schulverwaltung, Schulleitern und Lehrpersonen zu einigen wesentlichen Schlussfolgerungen (Kotthoff 2003, S. 213ff.):

Die ökonomischen Beweggründe für ein dezentralisiertes und somit effektiveres und effizienteres Bildungssystem offenbaren sich als das vermutlich bedeutsamste Motiv der Reform. So sanken die Ausgaben nach der Verlagerung der Entscheidungskompetenzen auf die kommunale und die schulische Ebene etwa für die Erteilung des Unterrichts um 17 Prozent. Es deutet sich jedoch an, dass diese Kostenreduktion weniger auf eine verbesserte Effektivität der Kommunen und Schulen zurückzuführen ist als auf Sparmaßnahmen, denen die Schulen ausgesetzt wurden; in den neunziger Jahren sanken die Bildungsausgaben, die Löhne wurden eingefroren und sanken also de facto auch, während Faktoren wie die durchschnittliche Klassengröße anstiegen. Infolgedessen entstand bei Lehrerinnen und Lehrern rasch die Befürchtung, dass es sich nicht um eine pädagogisch orientierte Reform handelte, sondern eben um eine im Wesentlichen aus ökonomischen Beweggründen durchgeführte Umstrukturierung – ein Verdacht, der gegenwärtig wohl auch noch besteht.

Die zentrale Intention der Reform, den Kommunen, Schulen und insbesondere auch dem Lehrpersonal erweiterte Gestaltungs- und Handlungsspielräume zu ermöglichen und mehr lokale Initiativen auszulösen, ist verwirklicht worden. Die Verpflichtung der Schulen, Programme und Pläne zu erstellen, die ihre Ziele und Arbeitsschwerpunkte beinhalten, und somit ein individuelles „Schulprofil“ zu erstellen, hat einen ausschlaggebenden Beitrag zur Schulentwicklung geleistet und zu mehr Heterogenität und Konkurrenz auf lokaler Ebene geführt.

Die erweiterte Eigenständigkeit für die Schulen, die sich auch auf Unterrichts- und Arbeitsmethoden bezogen, wurde einerseits von den Lehrerinnen und Lehrern sehr begrüßt, andererseits ist ihre Arbeitsbelastung dadurch insge-

samt erheblich angestiegen.⁷⁴ So scheint „das Verhältnis der Lehrerschaft zu den neuen Steuerungsmechanismen von Anfang an erheblich getrübt. [...] Die durch die Erweiterung der schulischen Entscheidungskompetenzen ausgelöste erhöhte Arbeitsbelastung der Lehrer bei gleichzeitig eingeschränkter staatlicher Unterstützung des Schulwesens [war] offenbar auch noch gegen Ende der 1990er Jahre eine erhebliche Hypothek für eine produktive Auseinandersetzung mit den neuen Steuerungsansätzen von Schule“ (Kotthoff 2003, S. 215).

Hinsichtlich der Dezentralisierungs- und Deregulierungseffekte auf das Verhalten der Eltern, Schülerinnen und Schüler ergibt sich eine ambivalente Situation. In den entsprechenden Steuerungsdokumenten wurde nachdrücklich auf die Bedeutung der Eltern für das Lernen der Kinder hingewiesen und die Beteiligung der Erziehungsberechtigten am Schulleben als ein wesentliches Element für die Entwicklung der Schülerinnen und Schüler sowie der Schule formuliert (Swedish Ministry of Education and Science 1997, S. 10). Jedoch geht aus einer *Skolverket*-Studie hervor, dass die aktive Teilnahme der Eltern in den Schulen generell nicht sehr hoch ist und sich das Interesse an Kooperation meist nur auf das eigene Kind begrenzt (Skolverket 1998a, S. 150).

Ebenso zeigt sich, dass die mit der Reform des Steuerungsansatzes verbundene Hoffnung, intensiver auf die Interessen der Lernenden eingehen zu können, nur bedingt erfüllt worden ist:

„Pupils also regard their opportunities to join in decision-making on working methods and lessons content as greater than they were three years ago [...]. But even here there are courses for concern. Every fourth pupil often feels stressed. Pupils still ascribe teachers lower marks for their ability to listen and show consideration for what pupils think, along with the ability to make use of the things pupils have learned in their own leisure time, than they do for their knowledge of, and commitment to, their subjects. One pupil in ten rarely feels that school never manages to reach“ (Skolverket 1998b, S. 149).

Die mit der Dezentralisierung eingeführte freie Schulwahl (*free choice*) hat zur Profilierung der Einzelschulen beigetragen. Diese Chance wurde jedoch nicht

⁷⁴ Die bereits erwähnte Skolverket-Studie beurteilt die Situation wie folgt: „[...] teachers have come under considerable pressure. They feel that their workload has increased, they are less satisfied with the way their schools are run, and are more unhappy with their salaries and employment conditions than they were three years ago. During the last year, one teacher in three has seriously considered a career change“ (Skolverket 1998b, S. 146).

gleichermaßen von allen Schulen genutzt. Gerade in den Ballungszentren, wo im Gegensatz zu den ländlichen Gebieten eine wirkliche Schulwahl besteht, hat dies bei weniger kreativen Schulen zu einer Verminderung der Anmeldezahlen und damit verbunden zu einem geschmälernten Schulbudget geführt. Mithin vollzieht sich gegenwärtig vor allem in den städtischen Gebieten eine Segregation der Schülerschaft, die durch die Zunahme von privaten Schulen noch verstärkt wird. Festhalten lässt sich, dass einige Ziele, die an die Dezentralisierung und Deregulierung des Bildungssystems geknüpft waren, erreicht wurden, jedoch nicht ohne Begleiteffekte wie die ethnische und auch leistungsbezogene Segregation der Schülerinnen und Schüler, die entsprechend die bildungspolitische Frage aufgeworfen haben, wie weit diese Entwicklung noch dem Postulat der Chancengleichheit entspricht.

3.3.2 Verständnis und Bedeutung von Standards

In der bildungspolitischen Diskussion in Schweden findet der Begriff „Standards“ oder „Bildungsstandards“ bislang kaum Verwendung. Dennoch liegen durch das nationale Curriculum, die darin enthaltenen Mindestziele, die fachspezifischen Curricula und die landesweiten Tests eigentliche Standards vor (Arbeitsgruppe „Internationale Vergleichsstudie“ 2003, S. 172). Es sind Content Standards und nur am Rande auch Performance Standards. Für die Etablierung klarer Leistungsstandards spricht der Befund einer jüngeren Studie (Wikström 2005): Wenn Schulen unter politischem Druck stehen, dann werden sie die Durchschnittsnoten heraufsetzen (*increased grade point average*), sofern Standards fehlen oder hochgradig interpretationsabhängig sind.

Durch das neue nationale Curriculum für die *grundskola* (1994) wurde die Autonomie der Kommunen und der Einzelschule verglichen mit dem Curriculum von 1980 nochmals erheblich erweitert. Festgelegt wurden 1994 qualitative Bildungs- und Erziehungsziele (*goals to strive for*), die allgemein gelten sollen. Sie sind verknüpft mit Leistungszielen (*goals to attain*), die das minimale Niveau festlegen, das die Schülerinnen und Schüler am Ende ihrer Schulzeit erreicht haben sollen und aufweisen können. Diese Ziele sind Minimalstandards, die für alle gelten. Die damit verbundene Policy-Strategie ist deutlich professionsorientiert und setzt dezentrale Entscheidungen voraus.

Wie weitreichend die Gestaltungsspielräume auf kommunaler und auf schulischer Ebene verstanden werden, veranschaulicht der folgende Ausschnitt aus einem Bericht des schwedischen Bildungsministeriums:

„Central regulations of the schools in terms of level and years will no longer exist. Teachers, within the framework of time schedule, will themselves decide the allocation of teaching hours between different years and subjects. The only restriction is that imposed by the knowledge requirement of the syllabi and the assessment at the end of the fifth school year. Different approaches through thematic studies inter-disciplinary teaching, project work etc. will provide pupils with a better overall education enabling them to understand broader knowledge areas as a whole. Headteachers will have overall responsibility for the co-ordination of this“ (Swedish Ministry of Education and Science 1996, S. 6).

Dementsprechend wurde eine garantierte „Mindestlehrzeit“ eingeführt, die keine detaillierten Stunden- und Zeitpläne für die einzelnen Fächer und Jahrgänge mehr kennt, sondern nur noch die Angabe einer Pflichtschulzeit in Stunden vom ersten bis zum neunten Schuljahr (Kotthoff 2003, S. 210). Die Stundenpools verteilen die Schulen nach ihren Anliegen, das heißt sie setzen das Zeitbudget flexibel ein.

Das heute gültige Curriculum erstreckt sich auf 19 Pflichtfächer, die wiederum in fünf übergeordnete Fachbereiche aufgeteilt sind.⁷⁵ Für die einzelnen Pflichtfächer bestehen Fachlehrpläne, in denen die fächerspezifischen Minimalziele (*goals to attain*) weiter ausdifferenziert werden. Diese Lehrpläne wurden 2000 für die *grundskola* sowie die *gymnasieskola* revidiert, sind wenig detailliert und lassen auch auf dieser Ebene viele Freiheiten. Das Erreichen der Minimalniveaus wird getestet. Die Schülerinnen und Schüler werden zum Abschluss des fünften Schuljahres einem freiwilligen und zum Abschluss des neunten Schuljahres einem obligatorischen Leistungstest unterzogen. Dieser nationale Test erfolgt in den Kernfächern Schwedisch, Mathematik und Englisch. Obligatorisch sind zudem die nationalen Tests am Ende der Pflichtkurse in den Jahrgangsstufen 10-12. Zusätzlich können noch vor dem 9. Schuljahr oder in anderen Fächern zentrale Tests durchgeführt werden – dies zum Teil nach Vorgabe der Schulträger. Erst die Anbindung der lediglich auf allgemeinem Niveau verfassten Ziele des nationalen Curriculums für die *grundskola* an nationale Prüfun-

⁷⁵ 1. Kernfächer (Schwedisch, Mathematik, Englisch); 2. naturwissenschaftliche Fächer (Biologie, Chemie, Physik, Technologie); 3. sozialwissenschaftliche Fächer (Erdkunde, Geschichte, Religion, Gesellschaftskunde); 4. praktisch-ästhetische Fächer (Kunst, Werkunterricht, Hauswirtschaft, Musik, Sport & Gesundheit) und 5. individuelle Wahlpflichtfächer (zweite/dritte Fremdsprache, bei Bedarf Unterricht in der Muttersprache, Zeichensprache, andere Fächer, mit denen sich die Lernenden je nach Interesse/Neigung befassen möchten) (Kotthoff 2003, S. 212).

gen präzisiert die Vorgaben zu Leistungsstandards (Arbeitsgruppe „Internationale Vergleichsstudie“ 2003, S. 172f.).

Eine entsprechende Entwicklung lässt sich auch für den Lehrplan der *gymnasieskola*⁷⁶ aufzeigen. Das neue nationale Curriculum (1994) gesteht auch den weiterführenden Schulen mehr Entscheidungskompetenzen und Verantwortung zu. Es enthält, ebenso wie das Curriculum für die *grundskola*, nur allgemeine Richtlinien und fachspezifische Lehrpläne und entsprechende Minimalanforderungen für die 17 allgemeinbildenden und beruflichen nationalen Programme (Kotthoff 2003, S. 211). Auch hier kann das Zeitbudget flexibel genutzt werden, die Schulen können nach ihren Bedürfnissen vorgehen und sind in der konkreten Gestaltung des Angebotes keinen bürokratischen Vorschriften seitens des Staates unterworfen. Geprüft wird das Ergebnis, der Weg dorthin kann sehr individuell begangen werden.

In der jüngsten Vergangenheit sind die nationalen Curriculum-Dokumente in die Kritik geraten; sie seien nicht genügsam abgestimmt mit der weitreichenden Ambition einer Verlagerung der Zuständigkeiten im Schulsystem. Begründet wird dies mit mangelnder Verständlichkeit in der Formulierung, so dass Eltern, Schülerinnen und Schüler sich kaum mit diesen und den Lernzielen beschäftigen und auseinander setzen. Überdies findet die Unterscheidung zwischen *aims to strive for* und *aims to reach* kaum statt, das heißt die Ersteren werden selten berücksichtigt (Ekholm 2005). Ein zur Prüfung dieser Probleme eingesetztes Komitee schlägt denn auch vor, die Erstellung von Fachlehrplänen künftig den Schulen bzw. Kommunen zu überlassen und diese inklusive Zeitbudgets pro Fach, also nicht mehr landesweit, vorzugeben (ebd.); entsprechende Pilotversuche stehen vor dem Abschluss. Die Abschaffung von Zeitplänen käme zudem dem Fokus auf *goals* und *outcomes* entgegen, wobei diese Leistungsziele und ihre Struktur in Curriculum und Fachlehrplänen dann einer weiteren Klärung bedürften.

⁷⁶ Die schwedische *gymnasieskola* schliesst als weiterführende Schule an die *grundskola* an. Im Gegensatz zum deutschen Gymnasium umfasst die *gymnasieskola* sowohl die allgemeine als auch die berufliche Bildung. Um Missverständnissen vorzubeugen, wird auch hier im Folgenden weiterhin der schwedische Begriff *gymnasieskola* verwendet.

3.3.3 Überprüfung von Standards: Tests und Prüfungen

Das nationale Prüfungssystem soll gemäß der Nationalen Behörde für das Schulwesen (*Skolverket*)⁷⁷ den Lehrenden eine Unterstützung zur Notenvergabe bieten, über den Lernfortschritt der Schülerinnen und Schüler informieren und eine Grundlage für eine landesweit möglichst gleiche und einheitliche Beurteilung aller Lernenden bieten.

Das System unterteilt sich in verschiedene Prüfungs- und Testkomponenten, die von Universitäten in Zusammenarbeit mit Praxispersonen im Auftrag von *Skolverket* entwickelt werden:

- Fachprüfungen (*ämnesprov*) für die *grundskola*,
- diagnostisches Material bzw. Tests (*diagnostiska material*) für die *grundskola* und
- Kursprüfungen (*kursprov*) für die *gymnasieskola* (ebd., S. 244).

Die verschiedenen Tests und Prüfungen werden auf der Basis der nationalen Steuerungsdokumente wie Lehrplänen, Richtlinien, Beurteilungskriterien etc. entwickelt. Zu den Prüfungsaufgaben gehören umfassende und erklärende Durchführungs- und Auswertungsanweisungen, in denen der Bezug der Aufgaben zu den Lehrplänen hergestellt wird und authentische Lösungsbeispiele präsentiert werden. Die Auseinandersetzung mit den Aufgaben und den exemplarischen Schülerlösungen hat für die Lehrkräfte oftmals Fortbildungscharakter, weil sie daran ihre eigene Praxis reflektieren können (Erickson/Lander 2007, S. 33). Die Testaufgaben haben somit einen erheblichen Einfluss darauf, wie Lehrkräfte Vorgaben, Richtlinien und Lehrpläne interpretieren (ebd.). Die Tests haben auch einen Einfluss auf die Wahl der Unterrichtsinhalte und Arbeitsformen durch die Lehrkräfte, wie sich die Schulbuchautoren ihrerseits wiederum stark an den nationalen curricularen Vorgaben orientieren, auch noch nach 1992, als die staatliche Lehrmittelkontrolle aufgegeben wurde.

Die Teilnahme an den Prüfungen in der fünften Klasse ist, wie gesagt, freiwillig, das heißt die Entscheidung obliegt den Kommunen. Die Mehrzahl der Gemeinden beteiligt sich daran. Die obligatorischen Fachprüfungen in der neunten Klasse finden aufgrund der strikten Geheimhaltung der Prüfungsin-

⁷⁷ *Skolverket* bezeichnet das 1991 gegründete Zentralamt für Schule und Erwachsenenbildung, also die nationale Behörde für das schwedische Schulwesen. Am 1. März 2003 wurde das Amt für Schulentwicklung (*Myndigheten för skolutveckling*) ausgegliedert.

halte durch das *Skolverket* an einem landesweit einheitlichen Termin statt. Die Ergebnisse der Prüfungen werden im darauffolgenden Schuljahr den Schulen rückgemeldet und veröffentlicht (ebd., S. 245). *Skolverket* unterstützt die Kommunen bei der Erhebung der Daten und der Auswertung der Ergebnisse. Weiterhin sammelt *Skolverket* von ausgewählten Schulen die Ergebnisse der Fachprüfungen im neunten Schuljahr sowie der Kursprüfungen der *gymnasieskola*, um ein nationales Bild über die Ergebnisse dieser Überprüfung zu erhalten (ebd.).

Das *diagnostische Material* wird in der zweiten Klasse in den Fächern Schwedisch und Mathematik verwendet und dient der Erfassung des individuellen Standes der Sprachentwicklung und der mathematischen Begriffsbildung. In der siebten Klasse werden die *diagnostischen Tests* in den Fächern Schwedisch, Englisch und Mathematik eingesetzt, um den Leistungsstand der Schülerinnen und Schüler in Bezug auf die Ziele des nationalen Curriculums beurteilen zu können. Die Prüfungen und diagnostischen Tests werden jedoch zukünftig auf weitere Schuljahre ausgedehnt, um ein umfassenderes Bild zu erhalten; zudem sollen die Prüfungen für die verschiedenen Altersstufen dann obligatorisch werden.

Die *Kursprüfungen* an der *gymnasieskola* werden ebenfalls in den Fächern Schwedisch/Schwedisch als Zweitsprache, Englisch und Mathematik durchgeführt. Ferner gibt es noch durch das *Skolverket* aufbereitetes Prüfungsmaterial für die Fächer Französisch, Deutsch und Physik. Seit 2000 sind die Kursprüfungen im Fach Schwedisch für alle Schulen obligatorisch. Die Prüfungen in Englisch und Mathematik sind noch freiwillig, sollen jedoch auch zu verpflichtenden Prüfungsinhalten werden. In einer Stichprobe wertet das *Skolverket* die Resultate der Kursprüfungen aus und leitet die Ergebnisse an die entsprechenden Schulen im darauffolgenden Schuljahr weiter.

Der Equity-Bericht der OECD wertet den limitierten Einsatz von Tests im schwedischen System positiv, weil die Leistungsbeurteilung durch die Lehrkräfte Priorität habe und Tests lediglich eingesetzt würden, um die Urteile der Lehrerinnen und Lehrer zu unterstützen. Die Begründung erinnert an die Positionen der amerikanischen Berufsverbände:

„Judgments by teachers are usually built up over a period of time and a range of contexts so enabling a broad understanding of pupils' attainment to be amassed. Tests, on the other hand, are based on one-off occasions covering a limited sample of tasks. Teacher assessment is thus seen as having a higher validity value than tests“ (Nicaise/Esping-Andersen/Pont et al. 2005, S. 36).

Erst im Blick auf den Abschlusstest am Ende der neunten Klasse sind die Lehrkräfte verpflichtet, „to use the test results in awarding their final grade to students“ (ebd.). Die Testergebnisse erscheinen in den Abschlusszeugnissen zusammen mit den *grades* der Lehrkräfte (ebd.). Die Nutzung der Tests ist aber offenbar uneinheitlich, es fragt sich, ob damit nicht Ungleichheiten kaschiert werden, und schließlich ist die Fokussierung auf die drei Kernfächer ein Punkt, der von der OECD moniert wird (ebd., S. 37).

Generell wird der auch in Schweden zunehmende Einsatz von Tests – so auch eine Forderung der neuen, konservativ-liberalen Regierung – mit der Frage konfrontiert, ob damit formative oder summative Ziele verfolgt werden; gemäß der Aussage von Erickson/Lander (2007, S. 33) wollen die nationalen Tests in Schweden „im Grunde beides“ sein. Verwiesen wird auf eine OECD-Studie zur formativen Leistungsbeurteilung (OECD 2005a). Aus dieser Studie leitet sich folgender Befund ab, der vor amerikanischen „high stakes“ warnt:

„Given parental choice, the growth of competition between schools, and a possible increase of public distrust in teachers, the danger that summative assessment designed for *summing up* student achievement may be dominated by summative assessment for *checking up* purposes is a strong possibility. In that case, teachers may feel compelled to teach to tests: this can lead to a narrowing of the curriculum, students becoming more focused on test performance at the expense of learning goals, and, as research suggests, decreasing motivation and increased drop-out as a consequence“ (Nicaise/Esping-Andersen/Pont et al. 2005, S. 37).

2003 hat die schwedische Regierung *Skolverket* beauftragt, die Möglichkeit zu untersuchen, wie ein nationales Bildungsmonitoring basierend auf Testdaten in Schweden eingeführt werden kann. Damit würden Tests eine völlig neue Dimension erhalten, und der bisherige Zweck der Leistungsüberprüfung müsste neu bestimmt werden. Wenn dabei das Prinzip der Gleichheit gewahrt werden soll, dann müssen gemäß dem Bericht der OECD zwei Bedingungen erfüllt sein:

„First, considerable attention will need to be given to equity issues in the design of testing including the choice of constructs on which the assessments are based, their use of context, and the types of response they expect. Second, if teacher assessment retains a central role in the awarding of grades, as may be hoped, there is even greater urgency to provide additional exemplification ma-

terials, guidance, and moderation processes that will support consistency at national, municipality and school level“ (ebd., S. 37f.).

3.3.4 Qualitätssicherung und Evaluation

Einhergehend mit der Reform des Steuerungsansatzes, zeitlich jedoch nachfolgend, wurde in Schweden ein umfassendes System der Schulevaluation entwickelt. Dieses wird durch zahlreiche Akteure und verschiedenartige Evaluationsverfahren gelenkt und erstreckt sich auf alle Ebenen des schwedischen Schulsystems, also die Zentralebene, die Ebene der Kommunen und die der Einzelschulen. Auf allen diesen Ebenen sind Implementationsstrategien entwickelt worden, die teils mit neuen Gesetzen, teils mit Organisationsformen und teils auch mit Überzeugungsarbeit zu tun haben. Im Folgenden werden die wichtigsten Akteure und Konzepte der verschiedenen Ebenen aufgezeigt.

a) Die Akteure und Evaluationskonzepte der Zentralebene

Zu den bedeutendsten Akteuren im schulischen Evaluationssystem zählt das Ministerium für Bildung und Wissenschaft (*utbildningsdepartementet*). Ihm und so dem Staat obliegt die Gesamtverantwortung für die Qualitätsentwicklung bzw. Evaluation der Schulen, die durch die Festlegung nationaler Ziele, Ausführungen in den Schulgesetzen und Verordnungen sowie genauere Bestimmungen und Angaben in den Lehrplänen geleitet wird. Entsprechend den Dezentralisierungs- und Deregulierungsprozessen kommt dem Ministerium jedoch nur eine richtlinienbestimmende und steuernde Funktion zu, so dass es auf die Kooperation mit anderen Behörden angewiesen ist. Infolgedessen sind auf nationaler Ebene vier Hauptakteure für die schulische Evaluation zuständig:

- das Ministerium für Bildung und Wissenschaft (*utbildningsdepartementet*),
- die für die Lehrerbildung verantwortlichen Universitäten und Hochschulen,
- der schwedische Gemeindeverband (*Svenska Kommunförbundet*) und
- die Nationale Behörde für das Schulwesen (*Skolverket*).

Innerhalb des Ministeriums für Bildung und Wissenschaft ist der Bereich der schulischen Evaluation in der Schulabteilung angesiedelt, die für die *grundskola*

und die *gymnasieskola* und noch weitere weitreichende Aufgaben⁷⁸ zuständig ist und aufgrund ihrer geringen personellen Besetzung mit anderen Behörden zusammenarbeitet.

Dem Ministerium kommt vor allem eine strategische Position zu, die auch Einfluss auf die bildungspolitische Diskussion hat. Beispielsweise erlässt das Ministerium konkrete Anweisungen für die Nationale Behörde für das Schulwesen (*Skolverket*) und vergibt an die untergeordnete Behörde spezifische Evaluationsaufträge, deren Durchführung kontrolliert wird. Zudem werden die mittel- bis langfristigen bildungspolitischen Ziele des Ministeriums in Entwicklungsplänen (*utvecklingsplaner*) festgelegt, die alle zwei bis drei Jahre erscheinen und die zukünftigen Schritte im Bildungssystem skizzieren.

Für die Evaluation des Schulsystems ist vor allem der Entwicklungsplan für die Vorschule, Schule und Erwachsenenbildung von Bedeutung. Er ist überschrieben mit *Qualität und Gleichheit* und heißt im schwedischen Original: Utvecklingsplan för Förskola, Skola och Vuxenutbildning: Kvalitet och likvärdighet (Regeringens Skrivelse 1996/97). In diesem Plan sind die beiden wichtigsten Grundsätze der schulischen Evaluation niedergelegt, die bereits angesprochen wurden: Steuerung durch Ziele (*steering by goals*) und Steuerung durch Resultate (*steering by results*). Diese Grundsätze gelten für alle drei Ebenen, die nationale, die kommunale und die schulische. Daraus folgt, dass auf den verschiedenen Ebenen Ziel- und Steuerungsdokumente auszuarbeiten sind, deren Umsetzung durch Evaluationen geprüft werden sollen.

Wie die konkrete Verwaltung des Schulwesens, so liegt auch die Organisation und Durchführung der schulischen Evaluation in der Verantwortung der Kommunen. Um das übergeordnete Ziel einer möglichst gleichwertigen Qualität des Unterrichts zu erreichen und die diesbezügliche Qualität zu sichern, ist *Skolverket* als zentralstaatliche Aufsicht eingesetzt worden. Sie analysiert und bewertet die qualitativen Erträge der Aktivitäten auf kommunaler Ebene, greift aber nicht direkt ein.

Ein weiterer wesentlicher Akteur sind die Universitäten und Hochschulen, die seit 1977 für die Lehrerausbildung und -fortbildung zuständig sind. Obwohl auch die Universitäten durch den Dezentralisierungs- und Deregulierungspro-

⁷⁸ Zu diesen Aufgabenbereichen zählen: Kindertagesstätten, Kindergärten und Vorschulklassen; Beaufsichtigung aller Evaluationsmaßnahmen im Bildungswesen; Fragen der Budgetierung sowie ökonomische und statistische Analysen; Aus- und Fortbildung des Lehrpersonals und von Schulleiterinnen und Schulleitern; Evaluation und Analyse der Aktivitäten der Kommunen im Schulwesen und Kooperation und Aufsicht über die zugeordneten Behörden (Kotthoff 2003, S. 221).

zess wesentlich erweiterte Entscheidungskompetenzen erhalten haben, wird der Bereich der *Lehrerausbildung* weiterhin zentral durch Parlament und Regierung gesteuert. Beide entscheiden darüber, welche Universitäten oder Hochschulen mit Lehrerbildungsprogrammen beauftragt werden, sie entscheiden weiter über die allgemeine Struktur dieser Programme, die Zuteilung der Fonds und so die Anzahl der Studierenden, die Zulassungsvoraussetzungen und schließlich die Ausbildungsziele (Kallós 2003, S. 2f.).

Die Reform des Steuerungsansatzes, die Orientierung an Zielen und Resultaten und auch das veränderte Evaluationssystem stellen neue und hohe Anforderungen an die Lehrkräfte und erfordern erweiterte Lehrkompetenzen. Aus diesem Grunde hat das Ministerium für Bildung und Wissenschaft auf der Basis eines parlamentarischen Berichts von 1999 und eines neuen Gesetzes vom 25. Mai 2000 eine Reform der Lehrerausbildung initiiert. Durch die Einführung eines Basiscurriculums, das grundlegende Kompetenzen beinhaltet, die sich alle Lehrerinnen und Lehrer aneignen müssen, soll eine Standardisierung der Lehrerausbildung bewirkt werden. Das Studium unterteilt sich in drei Ausbildungsbereiche: die fach- und altersspezifische Ausbildung, ein vertiefendes Studium in verschiedenen Spezialgebieten und einen allgemeinen Ausbildungsbereich; Letzterer nimmt einen hohen Stellenwert innerhalb des Studiums ein und ist für alle Lehramtstudierenden verpflichtend. Es gibt statt wie bisher acht verschiedene Abschlüsse nur noch einen.

Die Reform rückt ab von der bisherigen Aufteilung, die Lehrerbildungsprogrammen weltweit eigen ist. Studiert werden üblicherweise Fächer, Methoden, Pädagogik (educational science), dazu kommen Praktika. Zwischen diesen Elementen bestehen so gut wie keine Verbindungen. Die neue Struktur der schwedischen Ausbildung sieht drei Lernbereiche (domains) vor, die integrativ verstanden werden und praxisorientiert sind. Der größte Anteil ist die professionelle Allgemeinbildung, auf Schwedisch *Allmänt utbildningsområde*. Sie umfasst 60 von maximal 180 Kreditpunkten und hat zwei Schwerpunkte, den Aufbau einer gemeinsamen Wissensbasis für alle Lehrkräfte sowie interdisziplinäre Fachstudien. Mindestens 10 der 60 Kreditpunkte entfallen auf Kurse, die im Praxisfeld angeboten werden. Der zweitgrößte Bereich mit 40 Kreditpunkten wird „Orientierungsbereich“ oder „*Inriktning*“ genannt. Er umfasst Studien, die auf die Unterrichtsfächer ausgerichtet sind und mit Altersgruppen oder Schulstufen gekoppelt werden können. Eine davon losgelöste Fachwissenschaft wie in Deutschland gibt es in der schwedischen Lehrerbildung nicht. Der dritte Lernbereich der *Specialisering* dient der Vertiefung des erworbenen Wissens und wird auch in der Weiterbildung angeboten, um Synergieeffekte zu erreichen (Kallós 2003, S. 4f.).

Die *Lehrerfortbildung*, die eigentlich im Verantwortungsbereich der Kommunen liegt und auch weiterhin durch diese organisiert und finanziert wird, soll ebenfalls umstrukturiert werden. Der Begriff der Lehrerfortbildung ist gemäß einem Bericht des Ministeriums für Bildung und Wissenschaft in „Kompetenzentwicklung“ umbenannt worden. Hinter dieser Namensänderung steht ein Konzept der Personalentwicklung im Berufsfeld, das neben der Verarbeitung der Erfahrungen im Praxisfeld vertiefende Studien, Methodenentwicklung, Teamwork und generell Anbindung an die Forschung umfasst (Utbildningsdepartementet 2000, o.S.; vgl. auch Kotthoff 2003, S. 225). Im Mittelpunkt stehen nicht Fächer, sondern die fortlaufende professionelle Qualifizierung der Lehrkräfte (Smitt 2002). Die neue Form der Lehrerbildung orientiert sich so an den Erfordernissen der einzelnen Schulen und lässt eine intensivere wissenschaftliche Ausrichtung erkennen.

Den *Schulleitern*⁷⁹ bzw. der *Schulleiterausbildung* kam bereits mit Beginn der Dezentralisierungs- und Deregulierungstendenzen Ende der siebziger Jahre eine für die Schulentwicklung und -qualität bedeutende Position zu. In einem Interview bezeichnete Mats Ekholm, ehemaliger Generaldirektor von *Skolverket*, die Ausbildung der Schulleiter als einen entscheidenden Aspekt für die Umsetzung der Reform des Steuerungsansatzes und des Evaluationssystems:

„When we started in the 70s to this work, we also used the school leader education in Sweden. So in 1976 the parliament took the decision that we are going to go through all school directors, all directors in our schools and the assistant directors with the training programme. One of the most important components of the programme was: How do you evaluate your schools? [...] To understand the Swedish situation you need to look at what was done between 1976 and 1987. That is the period it took to brainwash – no, to help the directors to understand that evaluation is good and how to do it, how to do it with teachers in a culture that hates to be evaluated. So we put that into the heads of the schools. [...] So when you come to the end of the last century, then you can say, well, something happened over the 25 years, and that is, people are not that afraid of being evaluated or evaluating themselves. They don't do it enthusiastically, they don't do it very professionally, but they see that it has some good in it“ (zit. n. Kotthoff 2003, S. 228).

⁷⁹ In Schweden wird zwischen *Schulleitern* und *Rektoren* unterschieden. Rektoren sind für mehrere Schulen zuständig, Schulleiter nur für eine. Die Ausbildung umfasst beide Positionen.

Rektoren bzw. Rektorinnen und Schulleiterinnen und Schulleiter sind für die Entwicklung der Schulqualität maßgeblich verantwortlich. Sie werden seit 1975 eigens ausgebildet, die Ausbildung ist obligatorisch und umfasst drei Jahre. Die Leitung unterscheidet sich durch die Ausbildung und ihren Status vom Kollegium, Schulleiterinnen und Schulleiter haben keine Lehrverpflichtungen und können sich so ganz der Führung ihrer Schule widmen. Dazu gehören die Entwicklung und Ausgestaltung des Schulcurriculums sowie die Selbstevaluation, wobei der Führungsstil kooperativ ist.

Ein weiterer Hauptakteur der Zentralebene ist der schwedische Gemeindeverband (*Svenska Kommunförbundet*), der alle 289 eigenständigen Kommunen vertritt. Der Verband unterstützt und berät Gemeinderäte, Verwaltung und das Leitungspersonal der Kommunen. Hinsichtlich der schulischen Evaluation und der Entwicklung von Schulqualität engagiert sich der Gemeindeverband insbesondere durch zwei Initiativen (Kotthoff 2003, S. 229ff.):

- „Qualitätsauszeichnung schwedische Schule“ (*Kvalitetsutmärkelsen Svensk Skola*): Diese Anerkennung wird alle zwei Jahre an eine oder mehrere Schulen vergeben. Entscheidende Kriterien für diese Auszeichnung sind unter anderem die Qualität der Schulleitung, der schulischen Datenerhebung und -analyse, der dokumentierten Ziele und Programme, die Kompetenzentwicklung des Lehrpersonals, die Kooperation inner- und außerhalb der Schule. Um daran teilnehmen zu können wird eine Gebühr von 5000 schwedischen Kronen (ca. 540 € bzw. 845 SFr.) erhoben, allerdings erhalten die Schulen dafür einen Abschlussbericht, der eine umfassende Analyse zur Entwicklung der Schule und der Schulqualität enthält. Durch diese Initiative soll den Schulen ein Anreiz geboten werden, die schulische Qualitätsentwicklung systematisch und auf Ziele hin ausgerichtet voranzutreiben.
- „Qualitätsentwickler“ für die Schule und Vorschule (*Kvalitetsutvecklare för skolan/förskolan*): Durch gezielte Ausbildung von Qualitätsentwicklern soll ein „Expertenstab“ zur Verfügung stehen, der die Schulevaluation der Kommunen systematisch aufbaut und weiterentwickelt.

Nicht unerheblich für das Engagement und die Initiativen des Gemeindeverbands im Bereich der schulischen Evaluation ist, dass seine Aktivitäten auf Angeboten basieren und die Teilnahme daran nicht verpflichtend ist. Zudem haben diese Initiativen keine Kontrollfunktion, sondern zielen nur auf die Unterstützung der Kommunen und einzelnen Schulen bei der Qualitätsentwicklung ab.

Der bedeutendste Akteur für die schulische Evaluation auf der nationalen Ebene ist die Nationale Behörde für das Schulwesen (*Skolverket*), die als zentrales Steuerungsorgan aktiv an der Sicherung der Schulqualität beteiligt ist. Die Hauptaufgabe der Behörde besteht darin, sich detaillierte Kenntnisse und datenbasierte Einsichten über die Organisation und Funktion des gesamten Bildungswesens zu verschaffen. Damit soll gewährleistet werden, dass die Entwicklung des schwedischen Schulwesens mit den Vorgaben und Zielen der nationalen Bildungspolitik übereinstimmt. Die Rolle von *Skolverket* im schwedischen Schulwesen wird allgemein so definiert:⁸⁰ Es ist Aufgabe der Behörde

- to define goals in order to administrate,
- to inform in order to influence,
- to review in order to improve.

Der Auftrag ist nicht verknüpft mit Konsequenzen oder Sanktionen. Wohl gibt der Staat die Ziele und Richtlinien der nationalen Bildung vor. Aber er vermeidet detaillierte Regulierungen, um den Gemeinden und den einzelnen Schulen so viel Freiheit wie möglich zu geben, ihre Arbeit eigenständig zu formulieren. Die Ziele sollen die örtliche Administration erleichtern und zugleich sicherstellen, dass im schulischen Unterricht Notwendigkeiten aufgrund der gesellschaftlichen Entwicklungen sowie der Arbeitswelt angemessen berücksichtigt werden, ebenso die persönlichen Bedürfnisse der Lernenden und das allgemeine Ziel der „greater equality in education“. Dazu betreibt *Skolverket* die permanente Reform der Curricula, setzt Regeln für die Graduierung⁸¹, legt Kriterien für die Gründung der unabhängigen Schulen fest und bestimmt über die Vergabe bestimmter Geldmittel.

Die Agentur informiert, damit der Prozess der Schulentwicklung möglichst transparent vollzogen werden kann. Zwei hauptsächliche Adressaten werden genannt: die Lehrkräfte in den Schulen einschließlich der mit ihnen kooperierenden Behörden auf der einen, die Eltern auf der anderen Seite. Die Schulwahl

⁸⁰ The Swedish National Agency for Education – *Skolverket*. <http://www.skolveret.se/sb/d/353> (Stand 01.08.06). Die folgenden Zitate stammen aus diesem Dokument.

⁸¹ Seit 1998 werden drei Grade verwendet, *accepted, well accepted, very well accepted*. Die fachspezifischen Richtlinien enthalten Kriterien für die Benotung und die nationalen Tests Beispiele dafür, welches Wissen Schülerinnen und Schüler vorweisen müssen, um ein bestimmtes Prädikat zu erreichen. Es ist allerdings aus Untersuchungen bekannt, dass nicht sämtliche Lehrpersonen diese Dokumente in idealtypischer Weise für ihre Benotungen verwenden (Ekholm 2005).

setzt möglichst gute Kenntnisse über die örtliche und regionale Situation voraus, die *Skolverket* mit Hilfe von Reviews und Evaluationen zur Verfügung stellt. Weiter dienen guten Kenntnisse auch der Entwicklung der von den Eltern gewählten Schule. Auch hier wird auf Kooperation gesetzt: „With this knowledge to support them, parents can engage in a constructive dialogue with teachers and school administrators how their individual school can be developed.“

Evaluationen durch *Skolverket* dienen der Verbesserung der Qualität. Damit kommen verschiedene Review-Methoden zur Anwendung, die je andere Zwecke erfüllen. Reviews werden auf drei Ebenen durchgeführt, mit Hilfe von Inspektoraten, als nationale Evaluationen und in der Beteiligung an internationalen Leistungsvergleichen. Die Datenerhebungen sind in jedem Falle mit Rückmeldungen verbunden, die so beschrieben werden. „Reviewing to effect improvements can involve:

- Establishing, by way of educational inspection, whether and how well an educational activity is functioning in relation to the regulations set out in the Education Act and school curriculum, and drawing attention to areas where a municipality or board of an independent school need to invest more efforts in their own development work.
- Focusing on, by way of national evaluations, areas where development is needed at the national level, as well as providing the underlying basis for this development, e.g. to help school principals and supervisors in their efforts to lead and rejuvenate activities at the local level.
- Participating in international evaluations in order to gain more indepth knowledge on comparable education systems and on how other countries have dealt with areas similar to those needing improvement in the Swedish education system.“

Skolverket gliedert sich heute in drei Departemente, zu denen noch die Administration gerechnet werden muss. Die drei Departemente haben je verschiedene Untereinheiten. Das erste Departement befasst sich mit der Programm- und Curriculumentwicklung, das zweite mit Rahmenbedingungen einschließlich der unabhängigen Schulen und das dritte mit Inspektionen, wobei hier landesweit fünf verschiedene Agenturen tätig sind.⁸² Hierbei sind gewisse Tendenzen zu erneut stärkerer zentraler Kontrolle – bezogen auf die Einlösung des Rechts

⁸² Inspektoratseinheiten von *Skolverket* gibt es derzeit in Göteborg, Linköping, Lund, Stockholm und Umea.

aller auf einen gleichen Zugang zu hochwertigem Unterricht – erkennbar. Seit 2003 hat nämlich die neue staatliche Inspektionsagentur als Teil von *Skolverket* unter dem Leitsatz „Inspecting for Improving“ damit begonnen, flächendeckend sämtliche Schulen einer regelmäßigen Inspektion zu unterziehen.⁸³

b) Die Akteure und Evaluationskonzepte der Mesoebene

Ausgangspunkt für die Einführung der flächendeckenden Schulinspektion dürften Schwierigkeiten sein, die sich im Zuge der Kommunalisierung des Schulwesens, worunter auch die große Verantwortung bezüglich Qualitätssicherung und -entwicklung fällt, ergeben haben. Verschiedene Kommunen waren angesichts dieser neuen vielfältigen Aufgaben überfordert; zudem war die professionelle Akzeptanz dieser Behörden von Lehrerseite nicht überall im nötigen Maß gegeben. Ekholm kommt neuerdings gar zum Schluss: „Allerdings haben inzwischen die Politiker auf Landesebene immer mehr ihr Vertrauen in kommunale Steuerungssystem verloren. Neue Formen staatlicher Steuerung sollen die Probleme bei der Steuerung durch die Gemeinden kompensieren“ (2007, S. 34).

Die Aufgaben der Kommunen hinsichtlich der schulischen Evaluation sind in den Entwicklungsplänen von 1997 und 1999 (Regeringens Skrivelse 1996/97; Regeringens Skrivelse 1998/99) verbindlich festgehalten. Wie auf der Zentralebene gilt für die Mesoebene die Steuerung durch Ziele und Ergebnisse (*steering by goals – steering by results*) als operierender Steuerungsmechanismus des kommunalen Schulwesens. Dementsprechend sind die Kommunen verpflichtet, zielorientierte *Schulpläne* vorzulegen, in denen sie aufzeigen müssen, wie die nationalen Ziele erreicht werden sollen. Neben deren Implementation sind die Gemeinden auch dafür verantwortlich, die in den Schulplänen festgehaltenen Ziele zu evaluieren und in einem jährlich veröffentlichten *Qualitätsbericht* die Fortschritte und Schwächen aufzuzeigen (Kotthoff 2003, S. 257).

Im Entwicklungsplan wird zwar nachdrücklich die Notwendigkeit einer kontinuierlichen schulischen Evaluation auf kommunaler Ebene festgehalten, hinsichtlich der Organisation und Durchführung ist der Plan, der nur wenig regeln will, naturgemäß aber sehr abstrakt. Die Kommunen sorgen für unterschiedliche Formen der Konkretisierung, was dann aber die Vergleichbarkeit fraglich macht. 1999 haben weniger als die Hälfte (45%) der Kommunen den obligatorischen Qualitätsbericht vorgelegt, wobei sich diese Zahl aber inzwischen ver-

⁸³ Zu den Zielen, Vorgehensweisen und Methoden der Inspektion siehe: <http://www.skolverket.se/sb/d/160> (Stand 26.04.07).

bessert hat. Sie belegt aber, dass der Implementationsprozess auf lokaler Ebene an vielen Stellen sehr schwierig war.

Wie die konkrete Ausgestaltung des Schulplans vorgenommen wurde, lässt sich beispielhaft an der Kommune Uppsala und deren Methoden und Instrumenten schulischer Evaluation zeigen (ebd., S. 266ff.). Der Schulplan von Uppsala fungiert als das wesentliche kommunale Steuerungsinstrument für das Schulwesen. Gleichzeitig ist er für die schulpolitische Auseinandersetzung der Kommune eine wichtige inhaltliche Orientierung, die Schulleiter, Lehrkräfte, Schüler, Eltern und Öffentlichkeit nutzen können. Transparenz wird aber erst geschaffen, wenn Plan und Evaluation in ein Verhältnis gesetzt werden.

Im Blick auf die Evaluation werden im Schulplan von Uppsala vier Punkte festgehalten:

- (1) Die schulische Evaluation ist auf allen Ebenen des kommunalen Schulsystems durchzuführen (Lernende, Lehrende, Schulleiter und -rektoren, kommunale Ausschüsse etc.).
- (2) In der *grundskola* sollen kontinuierliche Aufsichtsbesuche absolviert werden.
- (3) Für die schulische Evaluation und Aufsicht sind Pläne und Richtlinien auf allen Ebenen des kommunalen Bildungswesens zu erstellen, deren Ziel es ist, positiv auf die kontinuierliche Schulentwicklung zu wirken.
- (4) Die Schulen und lokalen Kommunalausschüsse werden dazu verpflichtet, abgestimmt auf die einzelnen Schulstunden in Programmen und Plänen darzulegen, wie sie die Erreichung der anvisierten Ziele evaluieren wollen.

c) Schulische Evaluation auf der Ebene der Einzelschule

Im schwedischen System soll der Einzelschule und ihren Evaluationsmaßnahmen eine beträchtliche Bedeutung zukommen. Bereits im Entwicklungsplan von 1997 wurde festgehalten, dass in einem zielgesteuerten Bildungssystem Evaluation und Qualitätsentwicklung vor allem in der Verantwortung jeder einzelnen Schule liegen müsse. Die Daten sollten für die Entwicklung der eigenen Schule eingesetzt werden und der Entwicklung der Schülerinnen und Schüler zugute kommen (Regeringens Skrivelse 1996/1997, S. 102, zit. n. Kotthoff 2003, S. 283). Bedeutend für die Evaluation in den Schulen sind, wie gesagt, die gestärkte Rolle des Schulleiters bzw. der Schulleiterin, die Erstellung eines *lokalen Arbeitsplans* und die regelmäßige Anfertigung eines schulischen *Qualitätsberichts*; die enthaltenen Informationen werden im Internet öffentlich zugänglich gemacht.

Durch den regelmäßig, einmal jährlich zu erstellenden Qualitätsbericht legt die Schule Rechenschaft darüber ab, wie erfolgreich sie die im lokalen Arbeitsplan definierten Ziele umgesetzt hat. Grundlage für diesen Bericht sind unter anderem die Testergebnisse der Schülerinnen und Schüler, die Mitarbeitergespräche und die durch die Unterrichtsbesuche erhobenen Daten des Schulleiters. Die Erstellung des Qualitätsberichtes liegt ebenfalls im Verantwortungsbereich der Schulleiterin oder des Schulleiters. Zu dessen Verfertigung hatte *Skolverket* zwar bereits zu Beginn allgemeine Richtlinien entwickelt, da die Durchführung und Auswertung der innerschulischen Evaluation im Verantwortungsbereich der einzelnen Schule liegt, ergaben sich diesbezüglich aber beträchtliche Unterschiede zwischen den Schulen (Kotthoff 2003, S. 283). In Reaktion darauf wurden inzwischen von staatlicher Seite strengere Richtlinien und vermehrte Vorgaben erlassen, etwa bezüglich der im Bericht zu verwendenden Indikatoren.

Der lokale *Arbeitsplan*, an dem sich die Evaluation orientiert, ist unter Mitarbeit der Lehrenden und Lernenden zu erstellen und sollte sich ausdrücklich auf die kommunalen wie nationalen Ziele beziehen (Alignment) und zum Beispiel Angaben zu den Methoden und Instrumenten der innerschulischen Evaluation machen. Damit der lokale Arbeitsplan und die darin festgehaltenen Absichten nicht nur ein abstraktes Dokument ohne Wirkung auf den Schulalltag bleibt, hat *Skolverket* Maßnahmen zur Umsetzung mit der Funktion des Schulleiters verbunden:

„Die Verantwortung des Rektors für *follow-up* und Evaluation umfasst [...] nicht nur den Arbeitsplan, sondern bezieht sich auf den Betrieb als Ganzes. Es ist die Verantwortung des Rektors, sicherzustellen, dass *follow-up* und Evaluation natürliche Bestandteile der schulischen Arbeit sind, über die Lehrer und Schüler übereinkommen, und dass die national festgesetzten Ziele erreicht werden. Kontinuierlicher Kontakt mit dem Personal, z.B. in Form von geplanten Mitarbeitergesprächen, ist wichtig, um den Bedarf an Betriebs- und Kompetenzentwicklung oder anderer Maßnahmen dafür, dass die Schüler die aufgestellten Ziele erreichen, festzustellen. Kontinuierliche Besuche in der Unterrichtssituation sind eine wichtige Ergänzung zu solchen Gesprächen. Sie haben auch große Bedeutung für die Motivation des Personals“ (Skolverket 1998c, S. 19, zit. n. Kotthoff 2003, S. 282).

3.4 England

Im europäischen Vergleich hat das englische Bildungssystem die wohl radikalsten Veränderungen in Richtung Output-Orientierung erlebt (Demaine 2002 und zahllose andere). Das betrifft den öffentlichen Bereich, nicht den privaten der Independent Schools, der keiner staatlichen Bildungspolitik unterworfen ist. Anders als in Schweden werden die unabhängigen Schulen nicht staatlich finanziert, und anders als in den Niederlanden sind diese Schulen überwiegend keine, die bestimmten Konfessionen angehören. Auch in England dominiert der öffentliche Sektor, die traditionellen Eliteschulen bilden nur einen Teil der unabhängigen, die wiederum etwa 8% der Schülerschaft stellen.

Die Veränderungen des öffentlichen Schulwesens gehen zurück auf die siebziger Jahre, als Margaret Thatcher im Kabinett von Edward Heath Secretary of State for Education and Science war (1970-1974). Bildungsreform war traditionell ein linkes Thema, das auf Chancengleichheit und den Ausbau des staatlichen Sektors gerichtet war. Ein vergleichbar erfolgreiches Konzept der Konservativen war nicht verfügbar. Thatcher richtete sich als Ministerin mit ihren Reden vor allem provozierend gegen die Gesamtschulen, ihre Maßnahmen aber leiteten keinen grundsätzlich anderen Kurs ein. Thatcher war zum Beispiel gegen die Fortexistenz der Grammar Schools und für die Stärkung der lokalen Schulaufsicht.

In der Opposition (1974-1979) wurde versucht, die konservative Schulpolitik offensiver zu gestalten, ohne damit bereits einen wirklichen Umbau des Bildungssystems zu verbinden. Das geschah erst im Laufe der ersten Regierung Thatcher (1979-1983), als neo-liberale Konzepte an Einfluss gewannen. Mitte der achtziger Jahre wurde das Marktmodell maßgebend, allerdings nicht in dem Sinne, dass freie Bildungsmärkte entstanden wären, die private Unternehmen, frei erwerbbar Produkte und Kunden vorausgesetzt hätten. Mit dem von Kenneth Baker lancierten Education Reform Act von 1988⁸⁴ bildete sich ein gemischtes System aus Schulwahl, gemäßigttem Wettbewerb und starker Output-Orientierung. Dieses System wird in der Literatur oft als „Quasi-Market“ bezeichnet. Kennzeichnend dafür ist auch die Etablierung neuer staatlicher Behörden sowie auf der anderen Seite die Übernahme von Kommunikationsformen aus der Wirtschaft, wie sie zum Beispiel Rankings darstellen (Stanfield/Tooley 2005).

⁸⁴ Zur Vorgeschichte des Gesetzes vgl. Flude (1989).

3.4.1 Der Wandel der Bildungspolitik

Das englische Bildungswesen war im 20. Jahrhundert lange charakterisiert als „national system, locally administred“. Die lokale Ebene ist 1902 mit der Gründung der *Local Education Authorities* (LEA) neu gestaltet worden. Grundlage dafür war ein Erziehungsgesetz, das unter dem konservativen Premierminister Arthur Balfour erlassen wurde. Das Gesetz erlaubte erstmalig allen Elementarschulen (Board Schools) eine Finanzierung aus örtlichen Steuern. Zu diesem Zweck wurden auch die Konfessionsschulen in das staatliche System integriert. Die früheren, gewählten School Boards wurden aufgelöst, die Macht lag nunmehr in Händen der LEAs, also seinerzeit auf der Ebene der Counties, Grafschaften oder Distrikte. Das neue System betraf England und Wales, wurde 1904 auf London übertragen und blieb bis 1965 unverändert. Danach wurden verschiedene Verwaltungsreformen durchgeführt, die Macht der LEAs war aber bis zum Education Reform Act 1988 ungebrochen. Die LEAs sind zuständig für alle staatlichen Schulen in ihrem Bereich, sorgen für den Unterhalt und die Betriebsmittel, stellen die Lehrkräfte an⁸⁵ und legen die Schülerzahlen fest.

Ein wesentliches Merkmal des englischen Bildungssystems war so seit Beginn des 20. Jahrhunderts die dezentrale Schulverwaltung. Von den drei Ebenen, Zentralverwaltung, LEAs und Einzelschulen, war die mittlere die dominante. Das Ministry of Education⁸⁶ löste 1944 das englische Board of Education ab, hatte aber gegenüber den Bezirken bis 1988 wenig Steuerungsmacht. Die einzelnen Schulen wurden von den LEAs regiert, wobei im Sekundarbereich seit dem 19. Jahrhundert immer Schulleitungen bestanden, zunächst oft in Form von *senior teachers*. Die klassischen Grammar Schools hatten Headmaster in starker Funktion, ebenso die unabhängigen Schulen. Die *National Association of Head Teachers* (NAHT) wurde 1897 in Nottingham gegründet und repräsentiert heute etwa 30.000 Schulleiterinnen und Schulleiter. Die nationale Schulleiterkonferenz der unabhängigen Schulen geht auf das Jahr 1869 zurück.

Seit Mitte der siebziger Jahre häuften sich in England die Klagen über sinkende Schülerleistungen und über zu große Unterschiede zwischen den einzelnen Schulen. 1965 waren Gesamtschulen eingeführt worden, die das während des Zweiten Weltkriegs gegründete dreigliedrige System auf der Sekundarstufe I ablösen. Der konservative Erziehungsminister Richard Austen Butler brachte 1944 im Parlament ein Gesetz durch, das erstmalig eine kostenlose Ausbildung

⁸⁵ Ausgenommen die Foundation Schools.

⁸⁶ Heute Department for Education and Skills (DfES).

auch in der Sekundarschule garantierte. Gleichzeitig wurde die Sekundarschule in drei Typen⁸⁷ untergliedert, in die die Schülerinnen und Schüler nach der Eleven-Plus-Prüfung aufgeteilt wurden. Unterschieden wurden die *Grammar Schools*, also die englischen Gymnasien, die *Secondary Modern Schools* sowie die *Secondary Technical Schools*, ein neuer Typus. Das Gesetz sah eine Aufteilung nach den „abilities and aptitudes“ der Schülerinnen und Schüler vor, also nicht nach ihrer Herkunft. Faktisch blieb die Zahl der Grammar Schools erhalten, Technical Schools auf der Sekundarstufe I wurden nur wenige gegründet und die weit-aus größte Zahl der Schülerinnen und Schüler besuchte die Secondary Modern Schools (Dunford/Sharp 1990, S. 17ff).

Mit dem Ministry of Education Circular 10/65 forderte die Labour-Regierung Wilson die LEAs auf, die Sekundarschulen in ihrem Bereich so zu reorganisieren, dass echte, nicht-selektive Gesamtschulen nach dem Vorbild der amerikanischen High school entstehen. Die Labour-Regierung wollte damit eine deutliche Verbesserung der Chancengleichheit erreichen und verstand Bildungspolitik als aktive Sozialpolitik. Die Schulen sollten für eine angemessene soziale Mischung der Schülerschaft sorgen und die Vorteile der Herkunft beenden. Die erste dieser Schulen, die Kidbrooke School in Greenwich, war bereits 1954 eröffnet worden. Bis 1975 hatte die Mehrzahl der LEAs die selektive Eingangsprüfung am Ende der Primarschule abgeschafft und Gesamtschulen eingerichtet. Meistens handelte es sich um Fusionen aus Grammar Schools und Secondary Modern Schools. In bestimmten ländlichen Bezirken blieben die Grammar Schools erhalten, bildeten aber sehr schnell eine Minderheit. Die LEAs vollzogen so die Politik der Labour Party.

Heute besuchen etwa 90% der englischen Schülerinnen und Schüler Comprehensive Schools. Allerdings sind viele dieser Schulen inzwischen Spezialschulen,⁸⁸ was ihnen erlaubt, bis zu 10% ihrer Lernenden selbst auszuwählen. Daneben existiert in England und darüber hinaus im Vereinigten Königreich ein Sektor von *Independent Schools*, die nicht aus öffentlichen Mitteln finanziert werden, im Blick auf ihre Organisation und Entwicklung unabhängig sind und nur einer besonderen Aufsicht, dem Independent Schools Inspectorate in London, unterliegen. In England umfasste dieser Sektor im Januar 2004 gemäß der amtlichen Statistik 2328 Schulen mit 599.630 Schülerinnen und Schülern.⁸⁹ Der

⁸⁷ Tripartite system.

⁸⁸ Diese Schulen bieten ein erweitertes Curriculum und andere Fächerkombinationen an.

⁸⁹ Die meisten Schulen, aber nicht alle, gehören dem Independent School Council an. Der Council zählt für das Vereinigte Königreich im Jahr 2006 etwa 2500 unabhän-

private Sektor wächst, wenngleich nicht übermäßig: 1994 waren es 2266 Schulen mit 544.850 Schülerinnen und Schülern (Statistics of Education 2004, S. 72). Diese Schulen heißen auch, entgegen dem üblichen Sprachgebrauch, „*public schools*“. Der Privatschulbereich unterhält sich weitgehend selbst und bezieht nur wenig staatliche Unterstützung.⁹⁰

Die Einführung von Gesamtschulen im staatlichen Sektor war Mitte der siebziger Jahre weitgehend abgeschlossen. Die vier Jahre konservativer Regierung hatten daran nichts geändert, auch weil die LEAs autonom vorgehen konnten. Zentrale Steuerungsmechanismen und alternative politische Konzepte gab es nicht. Für einen neuen Ton in der Bildungspolitik sorgte dann die Labour Party. James Callaghan, der frühere Außenminister, wurde am 5. April 1976 Nachfolger von Harold Wilson im Amt des Premierministers. Callaghan hielt ein halbes Jahr nach Amtsantritt eine Rede im Ruskin College der Universität Oxford. Diese Rede markiert den Beginn der englischen Debatte über *Bildungsstandards*, die also, ähnlich wie in Schweden, von sozialdemokratischer Seite angestoßen wurde. In Callaghans Rede wurde erstmalig der Ausdruck „higher standards“ benutzt, erstmalig war von „efficiency“ die Rede und erstmalig wurde das öffentliche Interesse an Rechenschaftslegung der Schulen herausgestellt.

Die gesellschaftliche Aufgabe der Schule wurde vor dem Hintergrund einer Defizitanalyse bestimmt, die herausstellte, dass Schulabsolventen nicht mehr über bestimmte Basisfertigkeiten (basic tools) verfügten, dass Akademiker eher einen Posten im öffentlichen Dienst als in der Wirtschaft anstrebten und dass mit den neuen informellen Methoden des Unterrichts nicht mehr genügend Sorgfalt und Tiefe des Verstehens erreicht werden (Callaghan 1976, S. 3). In Reaktion auf diese Defizite wurden ein „basic curriculum with universal standards“ gefordert, eine grundlegende Veränderung der Inspektorate sowie eine transparente Rechenschaftslegung über die Nutzung der Ressourcen und damit über das Erreichen oder Verfehlen der Ziele. Die Lehrerschaft, so Callaghan, könne sich nicht hinter die Persönlichkeitspädagogik zurückziehen.

gige Schulen mit etwa 620.000 Schülerinnen und Schülern (ISC 2006, S. 3). Dazu werden noch rund 130.000 Lernende in über 500 Preparatory Schools gerechnet. Die offizielle Definition lautet so: „An independent school is legally defined as any institution providing full time education for five or more pupils of compulsory school age which is not maintained by a local education authority nor a non maintained special school“ (Statistics of Education 2004, S. 10).

⁹⁰ 32,4% der Schülerinnen und Schüler in den ISC-Schulen erhalten Unterstützung in irgendeiner Form, meistens von der Schule selbst. Die LEAs tragen lediglich mit 3,3% dazu bei (ISC 2006, S. 9).

Die Ziele der Erziehung hatten zwei Spitzen, zum einen die Eingliederung in die Gesellschaft und zum anderen die Vorbereitung auf den Arbeitsmarkt.

„Both of the basic purposes of education require the same essential tools. These are basic literacy, basic numeracy, the understanding of how to live and work together, respect for others, respect for the individual. This means requiring certain basic knowledge, and skills and reasoning ability. It means developing lively inquiring minds and an appetite for further knowledge that will last a lifetime. It means mitigating as far as possible the disadvantages that may be suffered through poor home conditions or physical or mental handicap“ (ebd., S. 4).

Die Vision von Labour war also eine effiziente Gesamtschule, die für Chancengleichheit sorgt, aber auch ressourcengerecht verfährt, also nicht einfach immer mehr Mittel erhält, ohne nachzuweisen, was mit den Mitteln erreicht wird. Daraus ergab sich eine Bestimmung der Aufgaben der Zukunft, die verblüffend genau antizipiert, was später tatsächlich realisiert wurde. Am Ende seiner Rede sagte Callaghan:

„Let me repeat some of the fields that need study because they cause concern. There are the methods and aims of informal instruction, the strong case for the so-called ‚core-curriculum‘ of basic knowledge; next, what is the proper way of monitoring the use of resources in order to maintain a proper national standard of performance; then there is the role of the inspectorate in relation to national standards; and there is the need to improve relations between industry and education“ (ebd., S. 5).

Callaghans Rede hatte Voraussetzungen, die mit dem Erfolg der *progressive education* in der linken Schulreformbewegung seit Ende des Krieges zu tun haben. Die Rede ist keine direkte Vorbereitung der konservativen Bildungspolitik in den achtziger Jahren (Batteson 1997, Chitty 1998), aber sie stellt eine Absage an die „anti-autoritären“ Konzepte der Schulreform dar, die als bürgerlich galt und mit Labours Arbeitsmarktpolitik unverträglich war. Zudem hatte die Rede einen bestimmten politischen Kontext; sie war die Antwort auf eine zeitgenössische Problemwahrnehmung, die so etwas wie einen Kulturkampf darstellte (Riley 1998).

Der Literaturkritiker Tony Dyson hatte seit 1969 zusammen mit dem Anglisten C.B. Cox *Black Papers on Education*⁹¹ veröffentlicht, in denen die Kindzentrierung der englischen Reformpädagogik sowie der Aufbau von Gesamtschulen einer scharfen Kritik unterzogen wurden. Cox, Professor für Englisch in Manchester, war später als Präsident des *National Council of Educational Standards* einer der führenden Vertreter der englischen Standards-Bewegung. Die *Black Papers* lösten eine heftige Diskussion aus, die zusätzlich munitioniert wurde durch Vorfälle in der William Tyndale Junior and Infants Primary School in Islington, Nord-London. Diese extrem libertäre Schule⁹² galt den Kritikern als Beweis, dass etwas gegen die radikale Kindzentrierung in den öffentlichen Schulen unternommen werden müsse. Ein Mittel war Zielsteuerung, zu unterscheiden von den reformpädagogischen Idealen des persönlichen Wachstums, die in verschiedenen „experimentellen“ Schulen vor allem in London handlungsleitend wurden.

Die *William Tyndale Affair* (1974-1976) war der Anlass für Politik und Zentralverwaltung, stärker zu intervenieren und den öffentlichen Schulen Standards vorzugeben (Davis 2002). Auslöser war der Untersuchungsbericht des Richters Robin Auld über die Schule in Islington, der 1976 vorlag und für die Inner London Education Authority verfasst worden war. Der Auld-Report hob vor allem die Mängel der Schulaufsicht hervor und löste damit eine scharfe Diskussion über das Verständnis der Inspektorate aus.⁹³ Auch Vermutungen über einen Zusammenhang zwischen dem Rückgang der Schülerleistungen und der progressiven Pädagogik bestimmten die Diskussion. Im gleichen Jahr 1976 erschien Neville Bennetts Buch *Teaching Styles and Pupils Progress*, in dem die offenen Methoden des informellen Lernens im Primarschulbereich einer scharfen

⁹¹ Cox, C.B./Dyson, A.E. (Hrsg.) (1969a): *Fight for Education: A Black Paper*. London: Critical Quarterly Society. Cox, C.B./Dyson, A.E. (Hrsg.) (1969b): *Black Paper Two: The Crisis in Education*. London: Critical Quarterly Society. Cox, C.B./Dyson, A.E. (Hrsg.) (1971): *The Black Papers on Education*. London: Davis-Poynter Ltd.

⁹² Im Frühjahr 1974 übernahm Terry Ellis die Schulleitung. Er und sein Stellvertreter Brian Haddow führten neue Methoden des freien Lernens ein und spalteten das Kollegium. John Davis (2002, S. 285) beschreibt die Praxis als „very intense form of child-centered education“.

⁹³ In den sechziger Jahren ging die Zahl der Inspektorate drastisch zurück und die Schulaufsicht verstand sich immer mehr als „teacher-support“ (Davis 2002, S. 279).

Kritik unterzogen wurden.⁹⁴ In diesem Kontext empfahl Premierminister Callaghan die Einführung von Standards. Er grenzte sich wohl von den *Black Papers* ab, in denen er die Verteidigung von Klassenprivilegien erkannte, aber er argumentierte klar für eine neue Philosophie der Steuerung.

Pädagogischer Ausgangspunkt der Diskussion war der *Plowden-Report*⁹⁵ von 1966. Unter der Leitung von Bridget Plowden, seinerzeit Direktorin der Trust House Forte Ltd., erarbeitete das Central Advisory Council for Education für England im Auftrag des Erziehungsministers einen Bericht⁹⁶ über den Zustand und die Entwicklung des englischen Primarschulwesens. Der Auftrag wurde im August 1963 erteilt, der Bericht selbst war Ende Oktober 1966 fertig gestellt. Der Rat, der ihn verfasste, bestand am Ende aus 25 Mitgliedern, die sämtlich dem Bildungssystem nahe standen. Neben Lehrkräften, Schulleitern und Vertretern der Verwaltung waren Eltern im Council, verschiedene Professoren und Journalisten, aber weder Vertreter der Wirtschaft⁹⁷ noch der Politik.

Der Report berichtete über den Zustand der englischen Primarschulen und gab darauf bezogene Empfehlungen ab. Ausgehend von einer entwicklungspsychologischen Perspektive, die wesentlich Jean Piaget verpflichtet war, forderte der Report die Abschaffung der Eleven-Plus-Prüfung und so der Eingangselektion für die Sekundarschule. Weiterhin sollte die Primarschule nicht resultatorientiert und so ohne Leistungsdruck arbeiten, jedes frühzeitige Streaming der Schülerinnen und Schüler, ob nach Begabung, Leistung oder Fähigkeiten, sollte vermieden werden, die Lerngruppen sollten möglichst lange homogen bleiben. Empfohlen wurden des Weiteren die Einführung von Gesamtschulen sowie ein lehrergeleitetes, möglichst freies Curriculum. Der Report erinnerte auch daran, dass das System der Bezahlung der Lehrkräfte nach ihrer Leistung 1898 abgeschafft worden war (The Plowden Report 1967).⁹⁸

⁹⁴ Eine Nachanalyse der empirischen Daten Jahre später zeigte allerdings unterschiedliche Ergebnisse (Aitken/Bennet/Hesketh 1981). Diese Differenzierung wurde öffentlich allerdings nicht mehr wahrgenommen.

⁹⁵ The Plowden Report (1967): Children and their Primary Schools. A Report of the Central Advisory Council for Education. London: Her Majesty's Stationery Office.

⁹⁶ Es war der erste Report über die Primarschulen seit dem *Hadon's Report* von 1931.

⁹⁷ Der einzige Ökonom, I.C.R. Byatt von der London School of Economics, wechselte 1967 in den öffentlichen Dienst.

⁹⁸ Ein Teil des Lohnes der Lehrkräfte war bis Ende des 19. Jahrhunderts leistungsbezogen. Basis für diesen Teil der Bezahlung waren die Prüfungen am Ende des Schuljahres, die von den königlichen Inspektoren abgenommen wurden. Mit dem *Handbook of Suggestions for the Considerations of Teachers*, das das Board of Education

Bereits der *Hadow-Report*⁹⁹ von 1931 hatte ein kindzentriertes, individualisiertes Lernen in der Primarschule empfohlen, allerdings ohne große Wirkungen. Erst der *Plowden-Report* stieß auf breites Echo und wurde auch von der Lehrerschaft beachtet. Das Prinzip des „entdeckenden Lernens“ und das Bild des forschenden Kindes waren nicht mehr lediglich eine Provokation, sondern ein Aufruf zur Veränderung, der besser als 1931 psychologisch abgestützt war. Über die Richtung der Reform ließ der Report keinen Zweifel. Nach einer Kritik an den vagen Bildungszielen der öffentlichen Schule, die kaum mehr darstellten als „expressions of benevolent aspiration“ (The Plowden Report 1967, Abschnitt 497), heißt es:

„[...] one of the main educational tasks of the primary school is to build on and strengthen children’s intrinsic interest in learning and lead them to learn for themselves rather from fear of disapproval or desire for praise“ (ebd., Abschnitt 532).

Die Lehrkräfte sollten nicht von dem ausgehen, was messbar ist oder so erscheint; vielmehr sollten die Entwicklung des Kindes, die Freude am Lernen und der individuelle Lernfortschritt im Mittelpunkt stehen. Die Folgen für das Curriculum werden ausführlich dargelegt (ebd., Kapitel 17). Es ging nicht um einen Gegensatz zwischen „child“ und „curriculum“, sondern um einen kindgerechten Unterricht anspruchsvoller Gehalte.¹⁰⁰ Rückblickend schrieb Bridget Plowden (1987), dass die Kommission wohl das Lernen in der Schule individualisieren wollte, aber nicht zulasten des Wissens oder der praktischen Fähigkeiten der Kinder.

Im zweiten Kapitel des Reports wird auch etwas über den Nutzen von Tests gesagt. IQ-Tests werden als Momentaufnahmen in der Entwicklung des Kindes verstanden, die keine wirkliche Prognose erlauben, wohl aber bei der Diagnose helfen können. Zusammenfassend heißt es: „Though IQ scores are a useful rough indication of potential ability, they should not be treated as infallible predictors. Judgements which determine careers should be deferred as long as pos-

1905 herausgab, wurde die Autonomies der Lehrkräfte sukzessive erhöht (Board of Education 1905).

⁹⁹ Board of Education/Hadow, W.H. (1931): Report of the Consultative committee on the primary school. London: His Majesty’s Stationery Office.

¹⁰⁰ „The primary school curriculum must touch on the scientific and mathematical knowledge on which the modern world depends and in which children are particularly interested“ (The Plowden Report 1967, Abschnitt 875).

sible“ (ebd., Abschnitt 75 (d)). Die zentrale Konsequenz der Analyse waren also *nicht*, wie später, Tests, wobei spezielle Leistungstests gar nicht angesprochen wurden.¹⁰¹ Zu den praktischen Vorschlägen zählten die Flexibilisierung der Unterrichtszeit, ein integriertes statt fächerzentriertes Curriculum, erfahrungsbezogene Arbeit in Projekten, der Einbezug der schulischen Umwelten, Methoden des entdeckenden Lernens sowie neue Formen der Beschreibung von Lernfortschritten (ebd., Kapitel 16).

Am Ende des 16. Kapitels gebrauchen die Autoren des *Plowden-Reports* die beiden Begriffe „Standards“ und „Evaluation“, und zwar in Form einer abschließenden Negation:

„We have considered whether we can lay down standards that should be achieved by the end of the primary school but concluded that it is not possible to describe a standard of attainment that should be reached by all or most children. Any set standards would seriously limit the bright child and be impossibly high for the dull. What could be achieved in one school might be impossible in another“ (ebd., Abschnitt 551).

„Expressionist education“ nannte der konservative Philosoph Roger Scruton (1987) zwanzig Jahre später die Vorschläge des *Plowden-Reports*, die als Triumph der „educationalists“ bezeichnet wurden. Die reformpädagogische Psychologie habe nur dazu beigetragen, die Leistungen noch weiter abzusenken. Solche wenig belegten Thesen wurden in den Jahren nach Amtsantritt von Margaret Thatcher (1979) öffentlich angenommen und bestimmten die Diskussion. Der *Plowden-Report* schien ein fernes Ereignis aus einem Jahrzehnt zu sein, das keine Verbesserung gebracht hatte. Das konnte ohne Datenbasis leicht behauptet werden und war besonders für das bürgerliche Publikum eingängig. Welche Effekte der Report tatsächlich hatte, ist nie untersucht worden, während auch gesagt werden muss, dass er ohne eine entwickelte Implementationsstrategie veröffentlicht wurde. Wenn später darauf verwiesen wurde, in Deutschland und Österreich sei der wegweisende *Plowden-Report* verpasst worden (Gruber 1987), dann unterstellt das eine Nutzung in England, die angesichts des Politikwechsels in den achtziger Jahren vermutlich eher überschätzt wurde.

¹⁰¹ Der SAT-Leistungstest ist unter der Bezeichnung *Scholastic Achievement Test* 1901 an verschiedenen amerikanischen Colleges und Universitäten als Eingangstest eingeführt worden.

1984 veröffentlichte das 1977 gegründete Adam Smith Institute den *Omega Report: Education Policy*. Dieser Bericht war das Gegenstück zum *Plowden-Report*, er war weder entwicklungspsychologisch noch reformpädagogisch ausgerichtet, sondern vertrat eine markttheoretische Optik. Das zentrale Stichwort im *Omega-Report* war „*producer capture*“. Das Stichwort wird so erläutert: Fast-Monopolisten wie das staatliche Bildungswesen in England tendieren dazu, das Interesse der Produzenten und nicht der Konsumenten im Auge zu haben. Die Folgen sind „*employment laxity, giantism and resistance to change*“ (Adam Smith Institute 1984, S. 3). Das System ist träge, schwerfällig und bedient sich selbst, ohne auf den Output oder die Kunden zu achten. Abhilfe schaffen würde ein System, das den Eltern die freie Schulwahl erlaubt, die Finanzierung auf Bildungsgutscheine umstellt und die Schulen wie selbstständige Unternehmen betrachtet. Die Schulleitungen wären frei, die Lehrkräfte einzustellen oder zu entlassen, Eltern wären die Kunden der Schulen und die Löhne der Lehrkräfte würden sich nach ihrer persönlichen Leistung bemessen (ebd., S. 7). Wer mehr leistet, bekommt auch mehr Lohn, wobei jede Schule die Höhe des Lohnes für sich festlegen kann.

Dieses radikale Manifest stieß – das war Teil der Strategie – auf erbitterte Kritik der Linken, aber es setzte sich im Blick etwa auf die Liberalisierung der Kundenmacht durch Vouchers oder die sofortige und gänzliche Privatisierung der Schulen auch innerhalb der konservativen Partei nicht durch. Doch die Systemfrage war gestellt und so auch die Frage, wie die weitere Entwicklung des englischen Bildungswesens vor sich gehen sollte. Zudem war die Ablösung pädagogisch-psychologischer Konzepte der Schulreform durch ökonomische zu beobachten, die in Teilen der englischen Öffentlichkeit mehr Plausibilität erzielten, was nicht einfach als Neuerungssucht – *neophilia* (Winch 1987, S. 6) – abgetan werden kann. Die „*educationalists*“ verloren zusehends an bildungspolitischem Einfluss, weil sie der Idee der Output-Steuerung nichts entgegenzusetzen hatten.

Im Umfeld der Parlamentswahlen vom Juni 1987 erschien Stuart Sextons Buch *Our Schools: A Radical Policy*,¹⁰² in dem eine Strategie der schrittweisen und längerfristig angelegten Privatisierung entwickelt wurde. Grundlage der Strategie war eine Veränderung der Bildungsfinanzierung. Schulen, die selbständig werden wollen, würden in Zukunft direkte Zuwendungen (*grants*) von der

¹⁰² Stuart Sexton war zu diesem Zeitpunkt Leiter der Education Unit im Institute of Economic Affairs. Er schrieb Reden für Margaret Thatcher und war Berater der beiden Erziehungsminister Mark Carlise (1979-1981) und Keith Joseph (1981 bis 1986).

Zentralregierung erhalten, wären also im Blick auf das Budget nicht mehr von den LEAs abhängig. Sextons zentrales Argument war, dass sich die Öffentlichkeit nicht im Klaren darüber sei, wie teuer und uneffizient die staatlichen Schulen tatsächlich seien, die weder kostenbewusst arbeiteten noch sich an ihren Leistungen messen lassen wollten. Vor allem diese Idee der Überprüfung des Einsatzes der Ressourcen und so der Transparenz setzte sich durch.

Vor diesem Hintergrund erließ die dritte Regierung Thatcher den *Education Reform Act* von 1988, der gedacht war als Einstieg in die längerfristige Entwicklung eines nationalen Bildungsmarktes. Die staatlichen Schulen sollten in eine Wettbewerbssituation versetzt werden und mit besonderer Qualität Kunden anziehen. Das umfangreiche Gesetz mit seinen insgesamt 238 Paragrafen sah sechs gravierende Neuerungen vor, die das System grundlegend verändern sollten:

- (1) Einführung eines *Nationalen Curriculums*, das bestimmte Themen und Fächer verbindlich vorschreibt. Damit sollen „knowledge, skills and understanding“ erfasst werden, über die die Schülerinnen und Schüler am Ende verschiedener Stufen ihrer Ausbildung je verfügen müssen (Art. 2);
- (2) Festlegung von drei Fächern im *Kerncurriculum* (core subjects) (Mathematik, Englisch, Naturwissenschaften) sowie weiterer *Grundlagenfächer* (foundation subjects) (Art. 3);
- (3) Unterscheidung von vier *Key Stages* der Schulkarriere (Schulbeginn bis zum siebten Lebensjahr, achtes bis elftes Lebensjahr, zwölftes bis vierzehntes Lebensjahr, fünfzehntes Lebensjahr bis zum Ende der Schulpflicht) (Art. 3,3);
- (4) die Zulassung von *Grant-Maintained Schools*. Diese Schulen haben eine eigene Leitung (governing body) und scheren aus dem System der LEAs aus (opting out) (Kap. IV des Gesetzes);
- (5) freie *Schulwahl* der Eltern für GM-Schools (school choice) (Art. 61);
- (6) Einführung einer *Schülerpauschale* (formula funding): Die Schulen erhalten umso mehr Geld je mehr Schüler sie anziehen (Art. 33,4).

Die Auswirkungen dieses Gesetzes sind als *Education in a Post-Welfare Society* beschrieben worden (Tomlinson 2001), als Versuch, die individuelle Freiheit der Konsumenten zu stärken und durch mehr Wettbewerb die Leistungen zu verbessern. Tatsächlich aber wuchs zunächst die Macht der Zentralregierung, entstanden neue Bürokratien und wurden ständig anpassende oder nachbessernde Gesetze notwendig.

Die LEAs wurden nicht aufgelöst, sondern sollten durch Anreizverschiebungen herausgefordert werden. Das Mittel war die Schaffung von GM-Schulen, die nach ihrem „opting out“ aus der lokalen Schulaufsicht direkte Mittel von der Zentralregierung erhielten. Im Jahre 1996, acht Jahre nach Erlass des Gesetzes, gab es in England 1090 GM-Schools, davon waren 60% Sekundarschulen. Der *School Standards and Framework Act* von 1998 beendete den Sonderstatus der GM-Schools. Die bestehenden Schulen konnten wählen, ob sie als *foundation schools* weitermachen oder als *maintained community schools* in die Verwaltung der LEAs zurückkehren. Foundation Schools wurden nicht mehr direkt von der Regierung finanziert, behielten aber eine gewisse Unabhängigkeit. Sie kontrollierten den Zugang, stellten die Lehrkräfte ein und sie oder eine „charitable foundation“ waren Besitzer der Grundstücke und Gebäude (Statistics of Education 2004, S. 9).

In Interview-Studien zeigte sich, dass die Motive der Schulleitungen, Lehrkräfte und Schulkommissionen für das „opting out“ in den neunziger Jahren pädagogischer Natur waren. Die Mittel sollten nicht in Tests investiert, sondern für die Verbesserung des Lernens genutzt werden (Stander 2001). Aber längst nicht alle Schulen folgten dem Anreiz, von den LEAs unabhängig zu werden. 1996 gab es in England 18.480 Primar- und 3594 Sekundarschulen mit insgesamt über sieben Millionen Schülerinnen und Schülern und mehr als 360.000 Lehrkräften (Statistics of Education 2004, S. 16). Die GM-Schulen machten daher einen Anteil von unter fünf Prozent des Gesamtbestandes aus. Das ist angesichts des hohen Aufwandes eine erstaunlich geringe Zahl, die darauf schließen lässt, dass die konservative Regierung die Zufriedenheit der Schulen mit den LEAs unterschätzt oder zu geringe Anreize geboten hat.

2004 bestanden 150 lokale Erziehungsbehörden in den acht englischen Regionen sowie in Inner und Outer London (ebd., S. 15).¹⁰³ Dieses engmaschige Netz stellt eine örtlich bewährte Form der Schulverwaltung und Dienstleistung dar, mit der die Zentralregierung neue Funktionen und generell eine eher strategische Rolle verbinden wollte. In Interviews artikulierten die Akteure Befürchtungen, dass durch den Politikwechsel eine Zentralisierung zulasten der örtlichen Flexibilisierung und Innovationsfähigkeit einsetzen werde. Tatsächlich gab es nach dem Education Reform Act 1988 keine neue Rollenbeschreibung für die LEAs. Die konservative Regierung setzte auf das „opting out“, erst La-

¹⁰³ Nicht alle diese Behörden sind staatlich. Die erste LEA wurde im Dezember 1999 privatisiert. Die Consultingfirma *Cambridge Education Associates* CEA erhielt für 105 Mio. £ einen Siebenjahresvertrag zur Führung der LEA in Islington. Die Firma heißt jetzt CEA@Islington.

bour legte eine strategische Rolle fest, die nicht zufällig auf die Implementation von Standards bezogen war (Fletcher-Campbell/Lee 2003, Introduction).

Die Studie von Fletcher-Campbell/Lee zeigt die zentrale Bedeutung der lokalen Ebene und zugleich die hoch differenten Kontexte, in denen die verschiedenen LEAs tätig sind. Die strategische Rolle der Implementation von Standards wird in aller Regel angenommen, aber lokal interpretiert. Die neue Rolle setzt sich aus vier Politikstrategien zusammen, den schulischen Entwicklungsplänen (Education Development Plans), der gerechten Budgetierung der Schulen (fair funding), dem Praxiscode der Beziehungen zwischen den LEAs und den Schulen sowie dem Standards Fund. Diese vier Instrumente fanden durchgehend Zustimmung, niemand in den visitierten LEAs stellte auch die damit verbundene Steuerungsphilosophie in Frage. Allerdings setzt das „local leadership for national priorities“ voraus, einen breiten Spielraum und eigene Visionen, wie die Prioritäten umgesetzt werden können. Die Implementation verlief und verläuft weder uniform noch linear (ebd., S. 69).

Die vier Instrumente lassen sich wie folgt näher beschreiben:

- *LEA Education Development Plan*
Basierend auf einer Überprüfung der bisher erreichten Qualität aufgrund von Leistungsdaten werden Ziele für die einzelnen Schulen sowie für die jeweilige LEA als Ganze festgelegt; die LEA stellt auch ihre Prioritäten dar, wie die Verbesserung der Schulqualität erreicht werden soll.¹⁰⁴
- *The Code of Practice on LEA-School Relations*¹⁰⁵
Gemeinsam ist beiden Seiten, dass sie für das Anheben der Standards besorgt sind. Die Schulen verwalten sich selbst, die LEAs greifen nur ein, wenn der Erfolg ausbleibt. Das beiderseitige Verhältnis ist auf Partnerschaft und Kooperation angelegt. Als gemeinsames Prinzip gilt „zero tolerance of under-performance“. Die Priorität für die LEAs ist klar „to promote high standards in education“. Mittel dafür sind die Unterstützung der Schulen in ihren Anstrengungen zur Qualitätsverbesserung, ein vergleichendes Bildungsmonitoring im Einzugsbereich sowie die Orientierung der Entwicklung an Best Practice.
- *Fair Funding*
Seit 2001 sind die LEAs mit zentralen Fonds zuständig für die Finanzierung in vier Bereichen, nämlich besondere pädagogische Bedürfnisse, Regelung

¹⁰⁴ Näheres bei Arnold (1999).

¹⁰⁵ Die Instrumente sind näher beschrieben im Papier *Schools: Building on Success* (Department for Education and Employment (DfEE) 2001).

des Schulzugangs, Schulentwicklung und strategisches Management. Die Verteilung der Mittel erfolgt nach sieben verbindlichen Prinzipien, zu denen etwa das Anheben der Standards, die Gleichheit zwischen Schultypen und die Transparenz der Budgets gehören.

- *The Standards Fund*

Seit dem Education Act 1996 gibt es einen speziellen Fonds, der Mittel bereitstellt für die Implementation der Leistungsziele. Die Mittel werden von den LEAs nach jährlich neu definierten Kategorien eingesetzt, wobei seit 2001 größere Freiheiten bestehen, wie die Mittel örtlich verwendet werden (ebd., S. 3ff.).

Die Verlagerung der Gewichte ist im Jahr 2000 in einem offiziellen Papier des *Department for Education and Employment* (DfEE) beschrieben worden. Es heißt hier, dass die einzelne Schule, nicht die LEAs, die zentrale Handlungseinheit (delivery unit) für die Systementwicklung sei. Für die LEAs müssen neue Aufgaben definiert werden, die nicht länger im Bereich Verwaltung und Service liegen. Handhabe dafür bietet der *Local Government Act* 2000, in dem alle lokalen Autoritäten, also auch die LEAs, Teil der regionalen Entwicklungspolitik werden. Das bedeutet für die LEAs Vernetzung mit lokalen Partnern im Blick auf gemeinsame Ziele. Auch können sich die LEAs an nationalen Bildungsinitiativen beteiligen oder zur Sicherung guter Rahmenbedingungen beitragen. Direkt regieren sollen sie nicht mehr: „Good schools manage themselves“ (Department for Education and Employment (DfEE) 2000, S. 7). Ein Grund für diese Entwicklung ist das lokale Schulmanagement: Die einzelnen Schulen verwalten ihre eigenen Budgets.

In der Zwischenzeit hatte das erste Kabinett Blair 1997 die Regierungsgeschäfte übernommen und mit der Publikation des Weißbuches *Excellence in Schools* die grundlegende bildungspolitische Position von Labour bekannt gemacht. Die an Wettbewerb, freier Wahl und Rechenschaftspflicht orientierte Bildungspolitik der Konservativen wurde nicht grundsätzlich in Frage gestellt, was viele Parteimitglieder und vor allem die mit Labour sympathisierenden Intellektuellen überraschte. Den Fokus des White Paper bildeten *Standards* und *Zielsteuerung*, die beide der Verbesserung der Schülerleistungen dienen sollten. Eine Rückkehr zu den Positionen des *Plowden-Reports* war weder angestrebt noch überhaupt in Reichweite des bildungspolitischen Horizonts. Die neue, stärker marktwirtschaftliche Orientierung hatte Premierminister Blair im Vorfeld der Parlamentswahlen unter den Stichworten „Third way“ und „New Labour“ proklamiert, wobei unklar war, wieweit das die Bildungspolitik tangieren würde.

Die neue Regierung übernahm die erweiterte Selbständigkeit und Eigenverantwortung der Einzelschule, die regelmäßigen national einheitlichen Leistungstests (National Curriculum Assessment), die Entwicklung eines effizienten Schulinspektionswesens sowie die umfassende Information der Eltern über die Qualität des Schulwesens und die damit einhergehende Stärkung ihrer Rechte. Die Erhöhung der schulischen Standards wurde zu einem Credo der neuen Regierung, die so etwas wie einen Kreuzzug ausrief („crusade for higher standards“).¹⁰⁶ Rechenschaftspflicht, Standardisierung und weitestgehende Transparenz der schulischen Qualität bilden bis heute Eckpfeiler der englischen Bildungspolitik.

Die Labour-Regierung setzte aber auch neue Zeichen. Die Balance zwischen externer Kontrolle und Unterstützung der Schulen sollte wieder austariert werden und damit die Interessen und Bedürfnisse der Schulen und Lehrkräfte stärkere Beachtung finden. Schulische Evaluation besteht nunmehr aus einer Kombination von externer Kontrolle mit Unterstützung und Beratung, wobei den LEAs eine qualitätsentwickelnde Funktion zugebilligt wurde. Die konservative Regierung hatte kaum Unterstützungssysteme angeboten, die Implementationstrategie waren Erlasse und neue Behörden. Das änderte sich unter Labour, als die LEAs für die Implementation von Bildungsstandards zuständig wurden. Noch am 17. November 1999 hatte der damalige Chefinspektor der englischen Schulen, Chris Woodhead¹⁰⁷, die Existenz der LEAs mit dem Argument in Frage gestellt, dass man sie auch auslagern könne. Nur wenige pädagogische Funktionen bedürften einer demokratischen Kontrolle vor Ort, der größere Rest der administrativen Arbeit ließe sich ohne Verlust auch anders organisieren.

¹⁰⁶ 1997 wurde als Ziel benannt, dass 2002 80% aller 11-Jährigen auf dem key stage 2 mindestens level 4 als Durchschnittsniveau in Englisch erreichen sollen. Lag der Anteil 1997 für Englisch noch bei 62,5%, so ist er im Jahr 2000 bereits auf 75% angestiegen. Allerdings ist die Entwicklung mit dem Vorwurf verbunden, die besseren Zahlen seien die Folge einer Preisgabe von Standards; die Anforderungen seien kontinuierlich gelockert worden (Ackeren v. 2003a, S. 78f.).

¹⁰⁷ Woodhead war von 1994 bis 2000 *Chief Inspector of Schools* in England. Er war vor allem in der Lehrerschaft umstritten, weil er progressive Unterrichtsmethoden ablehnte. Labour Minister David Blunkett erneuerte seine Anstellung 1998. Woodhead (2004) ist heute der schärfste Kritiker des Labour-Kurses in der Bildungspolitik.

Inzwischen kann ein Ergebnis der Forschung als gesichert gelten:

„National policies may only be able to be implemented when local resources (broadly interpreted) have been full taken into account. It may be a criterion of an effective national policy that it has flexibility for local implementation“ (Fletcher-Campbell/Lee 2003, S. 69).

Das erste Bildungsgesetz der Labour-Regierung war der *School Standards and Framework Act* von 1998. Hier wurde unter dem Stichwort „Parental Preferences“ die Schulwahl de facto freigestellt. Die LEAs wurden verpflichtet, auf die begründeten Präferenzen der Eltern Rücksicht zu nehmen, wobei die Wahl nicht völlig freigestellt wurde.¹⁰⁸ Mit diesem Gesetz waren die LEAs selbst gehalten, „to promote high standards in primary and secondary education“ (*School Standards and Framework Act*, Art. 5), eine Formel, die seitdem immer wieder verwendet wurde. Diesem Zweck dienen die Entwicklungspläne, die durch den Secretary of State autorisiert werden (ebd. Art. 6 und 7). Die LEAs veröffentlichen anschließend ihren Plan und machen so transparent, was sie im Entwicklungszeitraum erreichen wollen (ebd., Art. 9). Der Secretary of State erhält das ausdrückliche Recht, „to secure proper performance of LEA’s functions“ (ebd., Art. 8). Zudem können durch den Secretary of State zwischen einzelnen Schulen „education action zones“ eingerichtet werden, was die strategische Rolle der LEAs aber nicht angreift (ebd., Art 10).

Der Fokus der Systementwicklung ist deutlich die einzelne Schule, was sich auch daran ablesen lässt, dass das Bildungsministerium gegenüber schwachen Schulen in sozio-ökonomisch benachteiligten Gegenden eine entgegenkommendere Haltung einnimmt als die konservative Regierung. Das Green Paper *Every Child Matters* von 2003 formuliert sogar eine klare Politikstrategie, die darin besteht, in lokalen Netzwerken alles zu unternehmen, dass jedes Kind und jeder Jugendliche seine Potenziale entfalten kann, ohne durch soziale, mentale oder ökonomische Handicaps behindert zu sein.¹⁰⁹ Das Weißbuch *Higher Standards, Better Schools For All* (2005) schloss hier an und formulierte spezielle Be-

¹⁰⁸ Die Pflicht der LEAs, den Präferenzen der Eltern zu folgen, ist nicht gegeben, wenn damit der effektive Einsatz der Ressourcen gefährdet ist, wenn die bevorzugte Schule den staatlichen Vorschriften widerspricht und wenn die Zulassung nur aufgrund bestimmter Fähigkeiten (einschließlich Hochbegabung) erfolgen soll (*School Standard and Framework Act* 1988, Art. 86).

¹⁰⁹ Treasury (2003): *Every child matters*. London: TSO. Siehe auch http://www.everychildmatters.gov.uk/_files

dingungen für die Schulen zur Entfaltung der Potenziale ihrer Schülerschaft (Department for Education and Skills (DfES) 2005). Unter der Labour-Regierung wurden auch die Leistungen der Lehrerschaft positiver in der Öffentlichkeit dargestellt als in der Vorgängerregierung. An der grundlegenden Steuerungsphilosophie änderte sich dadurch jedoch nichts.

Die charakteristische Mischung der englischen Bildungspolitik aus Top-down-Strategien und Bottom-up-Elementen blieb nicht nur erhalten, sondern wurde verstärkt. Fallstudien zeigen, wie die Labour-Regierung neue Ebenen der Beteiligung auf der Ebene der Schulen eingeführt, die Schulfinanzierung verändert und zugleich die zentrale Ebene gestärkt hat (Bache 2003). Die Pluralisierung vor Ort reflektiert keine pluralistische Machtstruktur (Wilson 2003), sondern setzt eine starke Steuerung von oben voraus. Die entscheidenden Parameter der Macht im Bildungsbereich blieben in der Hand der Zentrale, während auf der anderen Seite die Freiräume vor Ort erweitert werden. Das lässt sich auch an den jüngsten Plänen der Regierung Blair zeigen. Am 28. Februar 2006 veröffentlichte die damalige Erziehungsministerin Ruth Kelly ein neues Erziehungsgesetz, in dem bereits die Wiedereinführung von GM-Schulen vermutet wurde.¹¹⁰

Alle Schulen können neu sogenannte *trust schools* werden und so den Status der *foundation schools* erreichen, also der ehemaligen GM-Schulen. Gegenüber den Foundation-Schools erhalten die Trust-Schools eine neue Form der Regierung. Sie können sich externe Partner suchen, die im Governing Body (GB) der Schule die Mehrheit bilden, wobei ein Drittel des GBs mit Vertretern der Elternschaft besetzt sein muss. Die GBs sind für die Anstellung der Lehrkräfte sowie des nicht-pädagogischen Personals der Schule zuständig und sie legen die Aufnahmebedingungen für die Schülerinnen und Schüler fest.¹¹¹ Die Regierung geht davon aus, dass externe Partner mit Verfügungsmacht sich direkt auf die Schulqualität auswirken.¹¹² Ein Trust kann auch mehrere Schulen umfassen, und alle „local stakeholders“ können sich daran beteiligen. Business Foundations oder Universitäten können also (auch zusammen) Trust-Schools formen.

Die neue strategische Rolle der LEAs wird in einem erläuternden Papier zum Gesetz nochmals so gefasst:

¹¹⁰ The Guardian Wednesday March 1, 2006.

¹¹¹ Dabei gilt ein enger Rahmen, der im School Admission Code von 2003 festgeschrieben ist.

¹¹² http://www.findoutmore.dfes.gov.uk/2006/02/trust_schools.html (Stand 04.08.06).

„Local authorities will take in a new strategic role, with duties to promote choice, diversity, high standards and for the first time, the fulfilment of *every child's* educational potential. They will also respond to parental concerns about the quality of local schools – and, in doing so, they will have new powers to intervene earlier when performance is poor. As the commissioner of school places, local authorities will be able to propose expansions to all categories of school, set the terms for school competitions and take all decisions relative to school organisation“ (Department for Education and Skills (DfES) 2006a, S. 1; Hervorhebung d. Verf.).

Außerdem müssen die LEAs für einen fairen Zugang (fair access) sorgen und verhindern, dass die Aufnahme der Schülerinnen und Schüler einzig nach Talent erfolgt (selection by ability), weil sich bestimmte Schulen dann die Besten aussuchen würden und eine soziale Mischung nicht zustande käme. Interviews mit Eltern oder Schülerinnen und Schülern vor Aufnahme bleiben verboten (outlawed). Fairer Zugang soll auch dadurch erreicht werden, dass die LEAs für die Kinder der „most disadvantaged families“ einen freien Transport zur Schule ihrer Wahl und zudem allen Eltern spezielle Beratungsdienste für die Entscheidung anbieten, welche Schule sie für ihr Kind präferieren (ebd., S. 2). Schließlich übernehmen die LEAs Verantwortung „for making sure young people have a range of exciting and positive things to do in their spare time“ (ebd.).

School choice wurde zum überragenden Thema am Ende der zweiten Regierung Blairs. Der Premier selbst zieht in einem Vorwort zum Weißbuch *Higher Standards, Better Schools for All* (2005) eine überaus positive Bilanz seiner Regierung im Bildungsbereich. Ausgehend von der Rede James Callaghans 1976 sei in dem gesamten Vierteljahrhundert Schulreform viel erreicht worden, wobei die Verbesserungen seit 1997 naturgemäß besondere Hervorhebung erfahren. Dazu zählen die erweiterte Verantwortung der Schulen etwa im Primarbereich für „literacy and numeracy“, die Erhöhung der Freiheiten und im Gegenzug die Verbesserung der Kontrollen, die dazu geführt hat, schlechte Schulen gezielt herausfordern zu können; weiter die Revitalisierung der „comprehensive education“, die Anhebung der Lehrergehälter und die Verbesserung der Ausstattung der Schulen. Die vergleichsweise hohen Investitionen des Staates im Bildungsbereich¹¹³ werden mit einem signifikanten Anstieg der Schülerleistungen

¹¹³ Die Gesamtausgaben stiegen von 35 Bio. £ (1997/1998) auf 51 Bio. £ (2004/2005) (Department for Education and Skills 2005, S. 13f.). Bis zum fiskalischen Jahr 2007/2008 ist eine Steigerung auf über 60 Bio. £ vorgesehen (ebd., S. 154).

bis 2005 gerechtfertigt, der vor allem mit einer Steigerung der Unterrichtsqualität erklärt wird (Department for Education and Skills (DfES) 2005, S. 1f.).

Die Gewinne der Reform werden der erhöhten Freiheit und der neuen Flexibilität zugeschrieben, während auf der anderen Seite immer noch zu wenig Wahlmöglichkeiten bestünden und die Standards nicht hoch genug seien (ebd., S. 3). Hier kommt die *school choice* ins Spiel:

„Parent choice can be a powerful driver of improved standards. Performance tables and inspections have given many parents the information that has enabled them to make objective judgments about a school’s performance and effectiveness. This has been an important pressure on weaker schools to improve“ (ebd.).

Im Weißbuch selbst wird die Stärkung der Kundenmacht (Eltern) so gefasst: Die Eltern erhalten im gesamten Schuljahr regelmäßige Berichte über den Leistungsstand ihrer Kinder, sie nehmen über Elternräte (Parent Councils) direkten Einfluss auf bestimmte Entscheidungen der Schule, etwa bezogen auf die Schulspeisung, die Schuluniformen oder die Disziplin; die Eltern können des Weiteren lokale und nationale Beschwerdedienste (Complaints Services) aufsuchen, sie haben Zugang zu allen relevanten Informationen über die Schule und sie erhalten finanzielle Unterstützung, wenn sie eine neue Schule suchen (ebd., S. 9). Das Ziel ist die Schaffung eines staatlichen Systems, das – wie in Schweden (ebd., S. 4) – echte Wahlmöglichkeiten bietet.

Ein Strategiepapier des Department for Education and Skills (DfES) sowie der Strategy Unit des Premierministers vom Juni 2006 vergleicht die Entwicklung in Richtung School Choice in sieben Ländern¹¹⁴ und zieht daraus folgende Schlüsse: „For choice-based systems to be successful in raising educational standards *for all*

- parents – particularly parents of disadvantaged children – need to be provided with high quality information, guidance and advice about the options open to them;
- help with the costs of making choices, such as covering the costs of transport to alternative schools, should be provided for those who need it;

¹¹⁴ Schweden, Vereinigte Staaten, Neuseeland, Niederlande, Dänemark, Norwegen und Finnland.

- schools should be prevented from „cream skimming“ (selecting the cheapest to teach or most able to learn) by putting in place funding regimes that reflect the higher costs of teaching some groups and by using regulation and statutory guidance to prevent unfair selection; [...]
- and above all, early effective action should be taken to tackle failing or poorly performing schools and increase the supply of good schools. Ultimately, the greatest safeguard against adverse effects on social inequalities and segregation is to make sure that there are more good schools“ (Department for Education and Skills/Prime Minister’s Strategy Unit 2006, S. 5).

Das Strategiepapier schlägt auch vor, dass die Finanzierung der Schulen (funding) den Wahlen der Eltern folgt, „so good performers are rewarded and can expand, and poor performers penalised“ (ebd.). Wenn Schulen keine finanziellen Implikationen spüren, wenn sie für Schülerinnen und Schüler nicht attraktiv sind, dann funktioniert das ganze System nicht (ebd.).

Die politischen Vorschläge des Strategiepapiers beziehen sich auf die Verbesserung der Beratung der Eltern, auf einen freien Transport für Kinder aus ärmeren Familien und weiter auf „one-to-one tuition“ in den Fächern Englisch und Mathematik in Schulen mit den schlechtesten Schülerleistungen, um zu verhindern, dass diese Lernenden ständig weiter zurückbleiben. Gefordert ist ein Verhaltenscode, wie das „cream skimming“ verhindert werden kann. Gute Schulen müssten expandieren können, schlechte Schulen geschlossen werden. Die Selbstregierung der Schulen soll den erwähnten Trust-Status erhalten, und für die LEAs soll eine Rolle gefunden werden, „commissioning rather than providing education, to support parents’ choices and diversity of provision“ (ebd., S. 5f.).

Rückblickend waren unter den konservativen Regierungen die LEAs wohl herausgefordert, aber nicht wirklich gefährdet. Der Opting-Out-Prozess betraf die weitaus meisten Schulen nicht, die lokale Schulverwaltung blieb weitgehend unverändert erhalten und wurde erst unter New Labour umgebaut. Gestärkt wurden dadurch auf der einen Seite die Zentralebene, also das Department for Education and Skills sowie die unabhängigen Behörden, auf der anderen Seite die einzelnen Schulen und nunmehr auch Akteure wie die Eltern. Zwar wurden durch die Einführung des National Curriculum und von *national assessments* Bereiche der Autonomie der Schulen beschnitten; dies betrifft jedoch in erster Linie die Selbständigkeit der Lehrpersonen, während die Selbständigkeit der Einzelschule in Fragen der internen Organisation und des Unterrichts insgesamt zugenommen hat. Insbesondere durch die Einführung des *local management of*

schools (LMS) wurden die finanziellen Kompetenzen der Einzelschulen bzw. der *governing bodies* beträchtlich erweitert.

Die Auswirkungen erweiterter schulischer Entscheidungsspielräume sind, wie verschiedene Studien belegen, grundsätzlich anders für Schulleitungen und die *governing bodies* als für die Lehrpersonen. Die mit der Autonomisierung einhergehenden Test- und Accountability-Mechanismen stellen vor allem für die Lehrerinnen und Lehrer eine Einschränkung pädagogischer Freiheiten dar und sind mit Mehrbelastungen verbunden. Die Frage, ob die Ausweitung schulischer Entscheidungsspielräume zu einem effizienteren Umgang der Einzelschulen mit ihren finanziellen Ressourcen führt und die Motivation der Mitarbeitenden erhöht, wird kontrovers diskutiert. Wiederum sehen die Schulleiterinnen und Schulleiter und *governing bodies* diese Veränderungen positiv. Die Kehrseite des erhöhten Kostenbewusstseins schlägt sich aber in den Personalanstellungen, Entlassungen, den Arbeitsverträgen und Lehrergehältern nieder. Seit diese Kompetenzen an die Schulen delegiert worden sind, sind Lehrerinnen und Lehrer ohne den vermittelnden Einfluss ihrer LEA in etwaigen Auseinandersetzungen mit der Schulleitung oder dem *governing body* unmittelbar deren Sanktionen ausgesetzt. Dies ist auch vor dem Hintergrund zu sehen, dass die Stärkung der schulischen Selbstverantwortung die funktionale Differenzierung zwischen Leitungspersonal und dem unterrichtenden Personal verschärft hat.

Die Steuerungsphilosophie und die Ziele hinter den Reformen haben sich unter New Labour nicht grundsätzlich geändert: Auf der Agenda steht weiterhin die Steigerung der Schülerleistungen bzw. die Beibehaltung der erreichten Zuwächse, wenngleich hierbei die Leistungsschere und der bestehende Zusammenhang zwischen sozio-ökonomischen Merkmalen und Leistung vermehrt Beachtung gefunden haben. Verstärkt gilt das Augenmerk aktuell dem Abbau von Bürokratie und Aufwand, so etwa im Bereich des Inspektionswesens (s.u.); zentral ist schließlich weiterhin die Umsetzung der Versprechen von mehr Wahlfreiheit und damit die Berücksichtigung der individuellen Bedürfnisse der Abnehmer (Department for Education Skills (DfES) 2004b). Wenn heute zudem vermehrt von „personal learning“ (Department for Education Skills (DfES) 2005), Kapitel 4) die Rede, bedeutet das nicht die Rückkehr zum *Plowden-Report*. Eher ist die Leistungsmaximierung jedes einzelnen Kindes gemeint, was Förderung und hohe Unterrichtsqualität voraussetzt (ebd., S. 51). Wenn es offiziell heißt, das Erziehungssystem müsse seinen Fokus in den Bedürfnissen (needs) des Kindes haben (ebd.), dann ist nicht beabsichtigt, zum offenen Unterricht zurückzukehren. Das Problem ist, wie jedes Kind die „basics“ (ebd., S. 52) lernen kann, keines zurückbleibt und das „tailored teaching in class“ (ebd., S. 59) allen Begabungen gerecht wird, ohne die Chancengleichheit

zu verletzen (ebd., S. 50). Vorausgesetzt ist dabei, dass die Erfindungen der Konservativen, also das National Curriculum, die regelmäßigen externen Evaluationen, das Testsystem und die Rankings, erhalten bleiben. Nach harter Kritik bei ihrer Einführung gelten sie inzwischen als bewährt und stehen jedenfalls bildungspolitisch nicht zur Disposition, was auch damit zu tun hat, dass die Regierung deutliche Leistungsverbesserungen auf Seiten der Schülerinnen und Schüler nachweisen kann (ebd., S. 14ff.).

3.4.2 Verständnis und Bedeutung von Standards

Das Nationale Curriculum (NC) in England, das seit 2004 in der vierten Fassung vorliegt, wird heute offiziell so definiert:

„The National Curriculum sets out a clear, full and statutory entitlement to learning for all pupils up the age of 16. It determines the content of what will be taught and sets attainment targets for learning. It also determines how performance will be assessed and reported. An effective National Curriculum gives teachers, pupils, parents, employers and the wider community a clear and shared understanding of the skills and knowledge that young people will gain at school.“¹¹⁵

Das National Curriculum, heißt es bei *National Curriculum online*,¹¹⁶ sei „an important element of the school curriculum“. Aber es ist deutlich mehr als das, nämlich eine weitgehend verbindliche Anleitung dessen, was gelernt werden muss, die so die Basis für alle Tests darstellt. Bei Einführung des Curriculums 1989 waren 80% der Lehrinhalte vorgegeben, was einen gravierenden Eingriff in die starke Autonomie der Unterrichtsgestaltung der Schulen und Lehrkräfte darstellte. Entsprechend kritisch wurde die Implementierung des National Curriculum von der Lehrerschaft begleitet.

Die Anforderungen des National Curriculum stellten insbesondere für die Primarschullehrkräfte eine fundamentale und abrupte Umorientierung dar, die von den Maximen des *Plowden-Reports* hin zur Output-Steuerung erfolgen musste. Allein mental waren damit erhebliche Stresserfahrungen verbunden. 1995 erfolgte eine erste Revision des NC. Die verbindlichen Anteile wurden eingeschränkt, indem ein größerer Teil der Gesamtunterrichtszeit in das Ermessen

¹¹⁵ http://www.nc.uk.net/nc_resources/html/about_NC.shtml (Stand 09.08.06).

¹¹⁶ <http://www.nc.uk.net> (Stand 09.08.06). Die nachfolgenden Zitate entstammen dieser Quelle. (Stand 09.08.06).

der Einzelschule gestellt wurde, der Umfang der Lehrpläne und die Anzahl der Lernziele drastisch reduziert wurden und eine Herabsetzung des verbindlichen Fächerkanons in der Altersstufe 14-16 vorgenommen wurde. Dies änderte jedoch nichts am starken Einfluss der zentralen Ebene auf die Gestaltung des Unterrichts – auf Kosten der Einzelschulen –, insbesondere über die gesetzlich verankerten regelmäßigen Überprüfungen der Schülerleistungen (*national assessment*).

„Standards“ im Sinne von normativen Kompetenz- bzw. Leistungserwartungen gibt es in England schon seit dem Education Act von 1870, mit dem staatliche Primarschulen eingeführt wurden. Damit erhielt das elementare Schulwesen gewisse Standards, die auch geprüft wurden. Allerdings gab es nie einen nationalen Lehrplan, und die Kompetenzerwartungen waren örtlich sehr unterschiedlich. Erst die Gründung der LEAs sorgte vorerst für eine gewisse Einheitlichkeit. Anders als in Kontinentaleuropa wurden aber keine übergreifenden Lehrpläne entwickelt. Die nationalen Gesetze gaben Fächer vor, nicht jedoch eine Arrangierung der Inhalte, die die Schulen in irgendeiner Form festgelegt hätten.

Der Ausdruck „standard“ wurde in der bildungspolitischen Diskussion seit den siebziger Jahren vor allem im Zusammenhang mit der Klage über die sinkenden Leistungen der Schülerinnen und Schüler verwendet. Gefordert wurde, dass die Standards erhöht werden müssten, womit eine Anhebung der Schülerleistungen sowie eine Verbesserung der Schulqualität gemeint war. Der Bezugspunkt war ein fiktiver Durchschnitt und es dominierte die Rhetorik des „Sinkens“ oder des „Immer-schlechter-Werdens“; verlässliche Daten lagen kaum vor. Das gilt auch umgekehrt für die Forderungen nach „Verbesserung“. Standards, die den gemeinsamen Ausgangspunkt bestimmen, gab es nicht. Dafür sorgte erst das National Curriculum.

Das National Curriculum verfolgt zwei übergeordnete Ziele, die offiziell so gefasst sind:

- Aim 1: The school curriculum should aim to provide opportunities for all pupils to learn and achieve.
- Aim 2: The school curriculum should aim to promote pupils' spiritual, moral, social and cultural development and prepare all pupils for the opportunities, responsibilities and experiences of life.

Angestrebt wird ein breites, aber balanciertes Curriculum, das vier hauptsächliche Zwecke erfüllen soll. Diese Zwecke werden wie folgt gefasst:

(1) Establish an entitlement

The National Curriculum secures for all pupils, irrespective of social background, culture, race, gender, differences in ability and disabilities, an entitlement to a number of areas of learning and to develop knowledge, understanding, skills and attitudes necessary for their self-fulfilment and development as active and responsible citizens.

(2) Establish Standards

The National Curriculum makes expectations for learning and attainment explicit to pupils, parents, teachers, governors, employers and the public, and establishes national standards for the performance of all pupils in the subjects it includes. These standards can be used to set targets for improvement, measure progress towards those targets, and monitor and compare performance between individuals, groups and schools.

(3) Promote continuity and coherence

The National Curriculum contributes to a coherent national framework that promotes curriculum continuity and is sufficiently flexible to ensure progression in pupils' learning. It facilitates the transition of pupils between schools and phases of education and provides a foundation for life-long learning.

(4) Promote public understanding

The National Curriculum increases public understanding of, and confidence in, the work of schools and in the learning and achievements resulting from compulsory education. It provides a common basis for discussion of educational issues among lay and professional groups, including pupils, parents, teachers, governors and employers.

Das Curriculum selbst wird nicht statisch verstanden. Vielmehr wird ausdrücklich festgehalten, dass es imstande sein muss, auf die Veränderungen in Wirtschaft und Gesellschaft, aber auch in der Schule selbst, angemessen zu reagieren. Die Lehrkräfte, individuell und als Profession, müssen ihren Unterricht auf die veränderten Bedürfnisse ihrer Schülerinnen und Schüler hin anpassen (reappraise) können, und die Schule muss sich auf die Auswirkungen des ökonomischen, sozialen und kulturellen Wandels einstellen können.¹¹⁷

Das National Curriculum enthält keine zeitlichen Vorgaben und legt auch nicht fest, wie die einzelnen Fächer im Stundenplan gewichtet werden müssen.

¹¹⁷ http://www.nc.uk.net/nc_resources/html/valuesAimsPurposes.shtml
(Stand 09.08.06).

Dies ist den einzelnen Schulen überlassen, die die konkrete Unterrichtsplanung vornehmen. Das National Curriculum¹¹⁸ ist, wie es der gesetzlichen Vorgabe entspricht, in vier Schlüsselstufen (Key Stages) gegliedert.

- Das *Key Stage 1* umfasst die ersten beiden Klassen und die Alter zwischen fünf und sieben.
- Das *Key Stage 2* umfasst die Klassen drei bis sechs und die Alter von sieben bis elf.¹¹⁹
- Das *Key Stage 3* umfasst die Klassen sieben bis neun und die Alter von elf bis vierzehn.
- Das *Key Stage 4* umfasst die Klassen zehn und elf sowie die Alter von vierzehn bis sechzehn.

Das National Curriculum gilt somit für die gesamte Pflichtschulzeit vom fünften bis zum sechzehnten Lebensjahr. Es besteht aus drei Kernfächern (*NC core subjects*), die auf allen Key-Stages unterrichtet werden, nämlich Mathematik, Englisch und Naturwissenschaften. Hinzu kommen neun weitere Fächer (*NC non-core subjects*), die nicht alle durchgehend angeboten werden. Die Fächer sind: Design and technology, Information and communication technology, History, Geography, Modern foreign languages, Art and design, Music, Physical education und Citizenship. Die ersten beiden Fächer sowie Physical education werden durchgehend unterrichtet, Geschichte und Geographie sowie Kunst und Musik nur bis Key Stage 3, moderne Fremdsprachen und Citizenship ab Key Stage 3. Religion wird ebenfalls auf allen Stufen angeboten, das Fach gehört zum Basiscurriculum, aber es wird neben dem NC angeboten. Für Key-Stage 4 gelten dann noch einige besondere Angebote wie Career oder Sex education.

Für jedes Fach und für jede Schlüsselstufe sind Unterrichtsprogramme (*programmes of study*) und Leistungsziele (*attainment targets*) festgelegt, welche beschreiben, was den Schülerinnen und Schülern vermittelt werden soll. Die Leistungsziele geben an, welches Wissen, welche Fähigkeiten und welches Verstehen Lernende mit unterschiedlichen Fähigkeiten und Reifegraden am Ende jeder Schlüsselstufe erreicht haben sollen. Die Ziele umfassen acht Niveaus, die sich durch zunehmende Schwierigkeitsgrade unterscheiden; hinzu kommt eine Stufe 9 für außergewöhnliche Leistungen. „Each level description describes the

¹¹⁸ http://www.nc.uk.net/nc_resources/html/ks3and4.shtml (Stand 09.08.06).

¹¹⁹ Mit elf Jahren findet in England der Übergang von der Primarschule auf die Secondary School statt.

types and range of performance that pupils working at that level should characteristically demonstrate.“

Daneben sind sechs Schlüsselqualifikationen (key skills) und fünf kognitive Kompetenzen (thinking skills) festgelegt, die die Schülerinnen und Schüler im Lauf ihrer Schulzeit erwerben sollen. Der Fokus liegt aber eindeutig auf den akademischen Standards der Schulfächer.

Die Beschreibungen der Kompetenzstufen sind so angelegt, dass sie die Basis der Leistungsbeurteilung durch die Lehrkräfte¹²⁰ für die ersten drei der Schlüsselstufen darstellen. Im Blick auf die vierte Stufe stellen nationale Qualifikationen das Hauptmittel der Leistungsbeurteilung dar. Für die ersten drei Stufen sind Durchschnittswerte festgelegt, die die Mehrzahl der Schülerinnen und Schüler bis zum Ende einer Stufe erreicht haben sollten. Das ist eine Art Normalerwartung, die danach unterschieden wird, in welcher Bandbreite der Niveaus (range of levels) die Schülerinnen und Schüler mehrheitlich arbeiten und welche Niveauerwartung für die „majority of pupils“ am Ende festgesetzt ist:

- Key Stage 1: Range of level 1-3, expected attainment at age 7: level 2;
- Key Stage 2: Range of level: 2-5, expected attainment at age 11: level 4;
- Key Stage 3: Range of level: 3-7, expected attainment at age 14: level 5/6.¹²¹

Die Beschreibung der acht bzw. neun Ebenen (level descriptions) sind für jedes Fach und jede Stufe national vorgegeben. Sie stellen eine verbindliche Sprachregelung für Kompetenzstufen dar. Die Lehrkräfte beurteilen bis zum Key Stage 3, auf welcher Stufe sich die Lernenden nach Absolvierung eines Unterrichtsprogramms befinden. Durch die Einführung der Stufe für „exceptional performance“ kann die Nivellierung der Erwartungen verhindert werden, denn so sind in jedem Fall Kompetenzstufen erreichbar und messbar, die über die Mindestanforderungen hinausreichen.

Die drei Hauptmerkmale des Systems lassen sich so zusammenfassen:

- Die Unterrichtsprogramme des National Curriculum können im Sinne von *Content Standards* interpretiert werden;

¹²⁰ „The level descriptions are not designed to assess individual pieces of work. They list aspects of attainment, based on the programmes of study, which teachers need to assess to build up a picture of a pupil’s performance over time in a range of contexts.“ URL: http://www.nc.uk.net/about/about_ks1_ks2.html (Stand 09.08.06).

¹²¹ Das schließt moderne Fremdsprachen mit ein.

- die Leistungsziele und Kompetenzstufen dagegen sind als kompetenzorientierte *Performance Standards* zu verstehen.
- Für jede Altersstufe werden Mindestanforderungen im Sinne von *Minimum standards* festgelegt.

Grundsätzlich hat die Einführung des National Curriculum zu einer Veränderung der Kultur vieler Schulen geführt, gerade in Schulen, in denen Leistungsüberprüfungen vorher kaum eine Rolle spielten. Die Implementation des National Curriculum fordert von den Lehrkräften eine detaillierte und systematische Unterrichtsplanung, orientiert an deutlich formulierten Zielvorgaben. Neben den Vorteilen bezüglich der Qualität des Unterrichts, die daraus resultieren können, wird der Zwang zur präzisen Unterrichtsplanung angesichts detaillierter Vorgaben auch als Belastung empfunden. Das National Curriculum und Standards haben in England damit auch Kritik erfahren.

Eine Kritik geht dahin, dass mit dem Fokus auf *academic excellence*, also der Zentrierung auf bestimmte schulische Fächer, der gesellschaftliche Nutzwert gering sei. Mit der Zentrierung des Wissens würden große Teile der Bevölkerung marginalisiert, die mit dieser Art Bildung wenig anfangen könnten (Coulby 2001). Am National Curriculum wurde weiter kritisiert, dass es keine Aussagen über das Lernen und Lehren enthalte, zu stark Fachwissen fokussiere und zu leistungsorientiert sei. Eltern und Lehrpersonen empfanden die Anforderungen als zu abstrakt und hoch. Befürchtet wird auch von Erziehungswissenschaftlern, dass mit der Leistungs- und Zielbetonung Kreativität und Professionalität der Lehrerinnen und Lehrer eine Einschränkung erfahren (Huber/Späni/Schmellentin/Cribblez 2006, S. 62f.). Schließlich wird am NC kritisiert, dass es der „thought world“ des pädagogischen Establishments entstamme und keinen öffentlichen Konsens voraussetze. Das NC sei eine reine Expertenkonstruktion (Woodhead 2004, S. 55).

Auch die Folgen sind kritisch betrachtet worden. Mit der Betonung der akademischen Fächer geht die Befürchtung einher, dass den Bedürfnissen von Schülerinnen und Schülern, die in diesen Fächern schwache Leistungen zeigen, nicht mehr Rechnung getragen werden könne. Es stellt sich die Frage, ob diese besondere Gruppe wirklich so gefördert werde, dass sie die geforderten Standards auch tatsächlich erreiche, oder ob die Lehrpersonen gezielt auf die Förderung der besseren Schüler setzten, weil es beim Leistungsvergleich auch um das Rating der Schule insgesamt gehe. Kritisch hinterfragt wird in diesem Zusammenhang, ob die beiden Ziele Inklusion und Wettbewerb der Schulen vereinbar seien, da doch der Wettbewerb exkludierend wirken müsse. Auch das Konzept der normativen Performance Standards ist der Kritik ausgesetzt. Bezweifelt

wird wiederum, dass die Setzung solcher Leistungsniveaus den unterschiedlichen Bedürfnissen der Schülerinnen und Schüler, insbesondere den sehr schwachen und sehr starken, gerecht werde. Zudem sei nicht eindeutig, ob es sich bei den Standards um normative Festlegungen handle oder um empirisch überprüfte Leistungserwartungen.

Im April 2006 veröffentlichte die Gewerkschaft ATL¹²² eine Untersuchung über die Auswirkungen der Reformen von New Labour auf die Arbeit der Lehrerinnen und Lehrer in den englischen Primarschulen (Association of Teachers and Lecturers (ATL): *Coming Full Circle* 2006). Die Studie beschreibt Effekte, die sich auf den Unterricht in *Key Stage 2* beziehen, also die Klassenstufen 3 bis 6. Die Studie ist auch in der Hinsicht bemerkenswert, als sie Longitudinaleffekte erfasst. 50 Primarschulen in England und Wales, die zwischen 1992 und 1994 befragt wurden, konnten bis auf zwei Ausnahmen erneut untersucht werden. Die Studie basiert auf Schulbesuchen, Unterrichtsbeobachtungen, Dokumentenanalysen und 188 Interviews mit Lehrkräften. Im Kern kritisieren die Lehrkräfte bestimmte Aspekte der Reform, heben andere positiv hervor, aber wollen nicht zum alten Zustand zurück. Das gilt allerdings nicht für die Leistungstests und das Schulranking.

Im Einzelnen wird von den beiden nationalen Strategien zur Literacy (NLS) und zur Numeracy (NNS) besonders die zweite als Erfolg bewertet. Beide Strategien¹²³ hätten den Unterricht bereichert und es sei wertvoll gewesen, dass sie an allen englischen Schulen realisiert worden seien. Die engen Vorgaben der Regierung wurden allerdings als Misstrauensvotum bewertet. Die entsprechenden Implementationsstrategien durch das Ofsted-Inspektorat und die LEAs wurden unterlaufen, wann immer das gerechtfertigt erschien. Die Schulen gaben an, dass sie die Programme adaptiert hätten, indem sie auf die Bedürfnisse der Kinder eingingen, dabei Stärken und Grenzen ihrer Lehrkräfte in Rechnung stellten, nicht notwendige oder nicht erfolgreiche Elemente aufgaben und auf bewährte Formen der Praxis zurückgriffen, die durch die neuen Programme eigentlich ausgeschlossen werden sollten. Aus den Unterrichtsbeobachtungen ergab sich, dass die Lernziele für den Unterricht spezifiziert und mit den Schülerinnen und Schülern geteilt wurden. Unterrichtet wurde überwiegend die ganze Klasse mit mündlichen Einführungen und Zusammenfassungen.

¹²² *Association of Teachers and Lecturers*. Die ATL ist Mitglied im nationalen Gewerkschaftsverband und umfasst nach eigenen Angaben über 160 000 Mitglieder.

¹²³ Umgesetzt mit literacy hours und daily maths.

Die Lehrkräfte werden von den Schulleitungen gemäß den Leistungszielen der LEAs beurteilt, was Leistungsdruck erzeugt, den die Lehrkräfte an die Schülerinnen und Schüler weitergeben. Die Lehrkräfte in allen Schulen fühlen sich „over-tested“, die Vorbereitung auf die Tests zerstörten das Curriculum besonders auf der Jahrgangsstufe 6; die Schulen, die mit „booster classes“ ausgestattet sind, führen jedes Jahr, also nicht nur am Ende der Stufe, optionale Tests durch, auf die sie die Schülerinnen und Schüler durch besondere Hausaufgabenprogramme oder Nachmittagsclubs vorbereiten.¹²⁴ Für viele schwächere Schulen seien die Tests demotivierend. Die Schulleitungen sind prinzipiell für Tests, kritisieren aber die Leistungstabellen und die technischen Probleme der Value-added-Versionen dieser Tabellen.

Weitere Ergebnisse betreffen die Auswirkungen von ICT im Unterricht, die neuen Schulassistenten (teaching assistants) und die neue nationale Strategie für den Primarbereich (*Primary National Strategy*, PNS). ICT-Medien werden pragmatisch genutzt und ohne rhetorische Überhöhung als zusätzliche Lernmöglichkeit für die Schülerinnen und Schüler betrachtet, die Assistenten werden eher als Bereicherung denn als Konkurrenz erfahren und die PNS wird als „freeing up“ des Curriculums und so als Bestärkung des professionellen Urteils der Lehrkräfte erlebt. Zudem erlaube die PNS mehr Spielraum für alternative Unterrichtsmethoden, Projektarbeit und vertiefendes Lernen durch Konzentration auf bestimmte Fächer.

Die 2003 lancierte Initiative *Excellence and Enjoyment – A Strategy for Primary Schools* erlaubt den Schulen eine stärkere Kontrolle über das eigene Curriculum. Die Schulen dürfen für die Klassen 4 und 5 des Key Stage 2 eigene Ziele setzen, die auf den Stand der Lernenden Rücksicht nehmen und dann auch die LEA-Ziele bestimmen. Die Beurteilungen der Lehrkräfte auf Key Stage 1 werden verändert. Verschiedene Schulen können Netzwerke bilden und zusammen Good Practice entwickeln. Die Curricula können gegenüber dem Erfahrungslernen geöffnet werden, und es ist erlaubt, fächerübergreifend zu unterrichten, ohne das Fachprinzip zurückzunehmen.

Die Veränderungen auf Klassenebene werden so beschrieben: Grundsätzlich wird ein Wandel vom früheren themenzentrierten Unterricht, der am Konzept des aktiven Lernens orientiert war, zum zielgesteuerten Fachunterricht festgestellt. Der Unterricht der gesamten Klasse im Unterschied zur Gruppen-, Partner- oder Einzelarbeit habe dramatisch zugenommen. Die Lektionen wer-

¹²⁴ Es gibt inzwischen auch Schulen der KS2-Stufe, die ihre Schülerinnen und Schüler in Feriencamps auf den Abschlusstest vorbereiten. BBC News 5. April 2006.

den von den Lehrkräften eingeleitet und zusammengefasst. Die Lehrerinnen und Lehrer führen den Unterricht und kontrollieren seinen Fortgang weit stärker als früher. Offene Formen des Unterrichts werden von den Lehrkräften mit dem *Plowden-Report* assoziiert. Die Schülerinnen und Schüler sitzen weit mehr in Reihen hintereinander als in Gruppen an Tischen. ICT-Lernformen werden sowohl für den Klassenunterricht als auch für andere Formen genutzt. Die Teaching Assistents werden zunehmend im Unterricht eingesetzt.

Wer diese Formen des Unterrichtens mit denjenigen zu Beginn der neunziger Jahre vergleichen konnte, wertete die früheren Formen keineswegs als die positiveren. Viele der interviewten Lehrkräfte standen ihrer früheren Praxis sehr kritisch gegenüber. Allerdings werden der zusätzliche Planungsaufwand und das viele „paperwork for accountability“ sehr negativ gesehen. Der Wechsel vom Themen- zum Fachunterricht in den neunziger Jahren war mit einem hohen Aufwand für die Anpassung des eigenen Repertoires verbunden. Nun befürchten manche Lehrkräfte, dass es einen erneuten Politikwechsel zurück zum themenzentrierten Unterricht geben könnte. Sie sprechen in Interviews vom „coming full circle“, geben aber zugleich an, dass in ihrem Erfahrungsbe- reich der Fachunterricht nach wie vor dominiere.

Noch ein Ergebnis ist relevant. Entgegen den Befürchtungen vieler Erziehungswissenschaftler (educationalists), dass die zentralen Vorschriften, besonders im Blick auf die Methoden des Unterrichts, zu einer Deprofessionalisierung und zu einem „deskilling“ der Lehrkräfte führen würden, erleben die Lehrerinnen und Lehrer selbst eher einen Zuwachs ihrer Professionalität.

„The strategies were seen as contributing to teachers’ professionalism by increasing their effectiveness and giving them the awareness and confidence to explain precisely what they are doing and why. In relation to their evaluation of the strategies, as with other aspects of government reform, practitioners’ perspectives on professionalism were strongly filtered through their core value of putting the child first“ (Association of Teachers and Lecturers (ATL) 2006, S.12).

Zusammengefasst werden die Auswirkungen der New Labour Reformen als „profund“ bezeichnet, obwohl sie in der Lehrerschaft umstritten waren und sind. Das professionelle Selbstbewusstsein der Lehrerschaft habe sich erholt, die PNS erlaube eine neue Mischung aus alten und neuen Methoden; wenn allerdings der durch Leistungstests, externe Evaluationen und Schulrankings erzeugte Druck erhalten bleibt, werden negative Auswirkungen befürchtet.

3.4.3 Überprüfung von Standards: Tests und Prüfungen

Nationale Assessments wurden 1992 zur Implementation der Bildungsziele des nationalen Curriculums eingeführt. Sie dienen der regelmäßigen Überprüfung der Schulen sowie der Leistungen der Schülerinnen und Schüler. Wie die nationalen Abschlussprüfungen stehen die Leistungsüberprüfungen in enger Verbindung mit dem Nationalen Curriculum. Sie wurden verstanden als „key policy tools“ (Sammons/Elliott/Welcomme et al. 2004, S. 97): „They are closely aligned to the national Curriculum and can be seen to be encapsulated stated goals of the system (the expectation is that the average child will achieve certain levels at given ages). They are also the chosen means to monitor schools and more recently 1998 onwards have been used to set national targets for given age groups“ (ebd.).

Internationale Vergleiche zeigen, dass England unter den westlichen Gesellschaften dasjenige Land mit den häufigsten Tests zu den unterschiedlichsten Zwecken ist (Whetton/Twist/Sainsbury 2000). Die Koordinierung der Formulierung der Lernanforderungen und der darauf bezogenen Tests und Prüfungen liegt in der Hand der nationalen *Qualifications and Curriculum Authority* (QCA), einer Nicht-Regierungsorganisation,¹²⁵ die staatlich finanziert wird, aber nicht der Behördenhierarchie untergeordnet und insofern quasi-autonom ist. Geldgeber ist das Departement for Education and Skills, geleitet wird die Test- und Entwicklungsagentur durch ein eigenes Board,¹²⁶ dessen Mitglieder durch den Erziehungsminister berufen werden. QCA ist zuständig für das öffentliche Prüfungswesen, die Weiterentwicklung des Nationalen Curriculums und für die nationalen Rahmenbedingungen der Berufsbildung. Das Board legt alle drei Jahre die Strategie für die weitere Entwicklung fest und muss dem jährlichen Operationsplan zustimmen.

2004 ist die für das National Curriculum zuständige Abteilung der QCA in die *National Assessment Agency* (NAA) überführt worden. Diese Tochteragentur der QCA ist verantwortlich für die nationalen Tests und die Weiterentwicklung des

¹²⁵ A non-departmental public body. Die Daten im Folgenden sind der folgenden Seite entnommen: <http://www.qca.org.uk> (Stand 06.08.06).

¹²⁶ Das dreizehnköpfige Gremium wird derzeit geleitet von Anthony Greener, der im Oktober 2002 berufen wurde. Greener war zuvor nach einer Managementkarriere Chairman des Getränkekonzerns Diageo plc. Das Board ist mehrheitlich aus Vertretern der Wirtschaft zusammengesetzt. Hinzu kommen Consultants aus dem Personal- und Erziehungsbereich sowie Spezialisten für Bildungssteuerung wie Mike Tomlinson. Die Praxis der öffentlichen Schulen ist mit einer Schulleiterin vertreten. Psychologen oder Erziehungswissenschaftler haben keinen Einsitz.

Prüfungswesens der englischen Schulen. Damit soll eine Qualitätsverbesserung der Testpraxis erreicht werden; zudem zielt die Reorganisation der Behörde auf eine Reduktion der bürokratischen Lasten, die die Schulen seit der Einführung der Tests zu tragen haben. Die Tests selbst (National Curriculum Assessments) finden am Ende der Key Stages 1, 2 und 3 in den Fächern Englisch, Mathematik und ab Key Stage 2 auch in den Naturwissenschaften statt.

In der Praxis heißt dies, dass die QCA standardisierte Tests für die drei Kernfächer entwickelt, die alle Schülerinnen und Schüler an den im National Curriculum festgelegten *attainment targets* messen. Die Tests ab Key Stage 2 werden extern korrigiert und benotet. 2000 ist noch das sogenannte *foundation stage* hinzugefügt worden, das Kinder im Alter zwischen drei und fünf Jahren umfasst. Nursery Schools und Reception Classes haben Curricula entwickelt, die verschiedene Lernbereiche umfassen und die ebenfalls zielgesteuert sind. Unterschieden werden sechs „areas of learning“,¹²⁷ die mit dem Education Act von 2002 verbindlichen Charakter angenommen haben und die von der QCA überprüft werden. Hier müssen für jedes Kind von den Lehrkräften dreizehn *Profile Summaries* verfasst werden, die den LEAs zugeschickt und dann national ausgewertet werden. Die Schule erhält die Ergebnisse, und es gibt auch einen Bericht für die Eltern (National Curriculum 2004, S.7ff.).

Neben den nationalen Tests sind auch die kontinuierlichen Leistungsbeurteilungen der Schülerinnen und Schüler durch die Lehrpersonen (*statutory teacher assessment*) Bestandteil des National Curriculum Assessment. Auch diese beziehen sich auf die im nationalen Curriculum niedergelegten Levels oder Kenntnisstufen, sie beinhalten jedoch zusätzlich zur obligatorischen Beurteilung der Leistungsniveaus eines Schülers am Ende eines Key Stage auch Leistungsinformationen, die nicht durch nationale Tests erhoben werden können. Die Leistungsbeurteilungen der Lehrkräfte und die extern ausgewerteten Ergebnisse werden an die QCA übermittelt, die das erhobene Datenmaterial auf nationaler Ebene auswertet. Nach Abschluss des vierten Key Stage, im Alter von sechzehn Jahren, findet die GCSE-Prüfung statt (*General Certificate of Secondary Education*).¹²⁸ Das frühere *baseline assessment* bei Schulbeginn ist ersetzt worden durch das Assessment der Basisstufe. Die Kosten beliefen sich im Jahr 2000 für die

¹²⁷ Die sechs Bereiche werden so bezeichnet: Personal, social and emotional development, communication, language and literacy, mathematical development, knowledge and understanding of the world, physical development and creative development.

¹²⁸ Es existiert auch eine zentrale Abschlussprüfung am Ende der Sekundarstufe II: General Certificate of Education (GCE).

Tests am Ende der Key Stages 1 bis 3 auf 24,3 Mio. £ bzw. 5,3 £ pro Schüler und Fach (Ackeren v. 2003a, S. 78).¹²⁹

Die QCA kontrolliert und entwickelt überdies die Qualität und die Einheitlichkeit der Beurteilung der Schülerleistungen seitens der Lehrpersonen durch die nationale Verbreitung von beispielhaften Beurteilungsverfahren und geeigneter Tests zur Feststellung des tatsächlichen individuellen Leistungsstands der Schülerinnen und Schüler. Es existieren nationale Standards für die einheitliche Beurteilung der Schülerleistungen in Bezug auf die acht Kompetenzstufen der *attainment targets*. Größter Wert wird darauf gelegt, dass die Anspruchsniveaus der Aufgaben möglichst weitgehend mit dem Anforderungsniveau der vorhergehenden Jahre übereinstimmen, damit die Stabilität des Systems gewährleistet ist und die Ergebnisse auch für Längsschnittuntersuchungen verwendet werden können.

Die Test- bzw. Abschlussprüfungsergebnisse der Schulen zu den Key Stages 2 und 4 werden jährlich in den Zeitungen in sogenannten „league tables“ publiziert, in denen die Schulen einen bestimmten Rang einnehmen. Grundlage sind die *School and College Achievement and Attainment Tables* des Department for Education and Skills, die zuvor als *Performance Tables* bekannt waren.¹³⁰ Kritische Stimmen bezweifeln, dass aufgrund der Daten zu den Schülerleistungen ein Urteil über die Qualität einer Schule möglich ist (z.B. Goldstein 2001). Auch in Reaktion darauf steht in England seit dem Regierungswechsel die Berechnung des individuellen Mehrwerts (value added) im Vordergrund. Damit wird der Leistungsfortschritt einzelner Schüler in der Primarschule vom KS1- zum KS2-Test bezeichnet, der die Bewertungsgerechtigkeit angesichts ungleicher Ausgangslagen erhöhen soll.¹³¹ Es soll nur das gemessen werden, was der Unterricht dem Eingangsniveau hinzufügt, das heißt, die Schulen sollen nicht für das belohnt oder bestraft werden, für das sie nichts können, nämlich den Input der Schülerinnen und Schüler, die sie aufnehmen müssen. Das statistische Verfahren wird auch im Sekundarbereich verwendet.

¹²⁹ Zum Vergleich: Die dreijährige Produktion des Plowden-Reports kostete seinerzeit die einmalige Summe von £ 120'699 (The Plowden Report 1967, S. 3).

¹³⁰ <http://www.dfes.gov.uk/performance/tables/> (Stand 11.08.06).

¹³¹ „For example, school A might show high percentages of pupils achieving at level 4 and above, while school B shows lower percentages. But in value added terms, the pupils at school B may have made more progress than other pupils who were performing at the same level at KS1, and therefore have a higher value added ‚score‘ than school A.“ URL:

http://www.dfes.gov.uk/performance/tables/primary_04/p3.shtml
(Stand 11.08.06).

Darüber hinaus werden zunehmend Hintergrund- und Kontextmerkmale in die Berechnungen einbezogen und Benchmark-Daten für Schulen mit ähnlicher Ausgangslage mitgeteilt. Mit den nationalen Tests sollen damit neben den bereits genannten noch weitere Ziele verfolgt werden: Sie bieten den Eltern eine Entscheidungsgrundlage für die Schulwahl¹³² und sie geben ihnen Informationen über das Leistungsniveau ihres Kindes. Obwohl die hochgesteckten Ziele der Labour Regierung im Blick auf die Steigerung der Schülerleistungen bis 2002 nicht voll erfüllt werden konnten, zeigt der Vergleich der Statistiken der letzten Jahre, dass die Zahl der Kinder zunimmt, die die von ihnen erwarteten Kompetenzniveaus erreichen.

In diesem Sinne kann die Regierung von einer Leistungssteigerung durch Standardsteuerung sprechen. Dabei spielen etwa folgende Zahlen eine Rolle: 1997, beim Amtsantritt der Labour-Regierung, verließ ein Drittel der Kinder die Primarschule ohne die Basisfertigkeiten für die Sekundarschule. 2005 beherrschten 79% der Primarschülerinnen und Primarschüler am Ende die Grundkompetenzen für Englisch und 75% für Mathematik. 2005 zeigte den höchsten je gemessenen Leistungsstand der Vierzehnjährigen in Englisch, Mathematik und ICT. Ebenfalls in diesem Jahr wurde eine Rekordrate von guten GSCE-Prüfungsergebnissen registriert. 56% der Sechzehnjährigen gegenüber 45% im Jahre 1997 erreichten ein Niveau von 5 und mehr „good GCSEs“ (Department for Education and Skills (DfES) 2005, S. 14; 2006b, S. 10f.).¹³³ Die Zahl der Schülerinnen und Schüler, die fünf und mehr A- bis C-Grade in ihren Fächern einschließlich Englisch und Mathematik erreichten, ist zwischen 1997 und 2005 um 8,5% gestiegen und liegt nunmehr bei 44,1% der Gesamtschülerschaft.¹³⁴

In der Literatur werden besonders die nationalen Tests hochgradig kontrovers diskutiert. Kritiker des englischen Testsystems betonen, dass objektive Messungen bzw. die Vergleichbarkeit von Schulleistungen über die Zeit nicht möglich seien, weil Standards letztlich immer sozial konstruiert seien und sich im Zuge

¹³² Die Befunde verschiedener Untersuchungen zur Frage, ob die freie Schulwahl zu einer verstärkten sozialen Segregation führt, sind widersprüchlich (vgl. Ackeren v. 2003a, S. 229f.).

¹³³ Am Ende des zweijährigen GCSE-Kurses erhält jeder Schüler eine Note (grade) für jedes Fach. Die Notenskala reicht von A (best) bis G (worst). Zudem gibt es noch die Klassifizierung U, die das Nichtbestehen des Kurses beschreibt. Das Niveau 5 ist gleichzusetzen mit einem C. Die meisten Universitäten verlangen ein C oder eine bessere Note in den Fächern Englisch und Mathematik.

¹³⁴ Education Guardian October 20, 2005.

des gesellschaftlichen Wandels verändern würden. Ein Problem wird auch in der Häufigkeit der flächendeckenden Tests gesehen, die in England tatsächlich vergleichsweise sehr hoch ist und entsprechende Kosten verursacht. Lehrpersonen verweisen auf die Stressbelastung der Schülerinnen und Schüler, aber auch auf den zeitlichen Aufwand, den die Vorbereitung auf die Tests in Anspruch nehme, worunter andere Fächer, aber auch die persönlichen Beziehungen zu den Schülerinnen und Schülern, leiden würden.

Die *National Union of Teachers* (NUT) zieht ein vernichtendes Urteil über die negativen Auswirkungen der Tests auf das Lernen der Schüler und schließlich auf deren Leistungen. Dieses Urteil führte auch vor dem Hintergrund der Schulranglisten sogar zu einem nationalen Boykottaufruf der Lehrgewerkschaften gegen die Teilnahme an den Tests. Im Gegensatz zu deutschen Autoren (Langen/Driscoll 2004; Rüdell 2005), die den Vorteil der englischen Testkultur darin sehen, dass diese nicht allein der Überprüfung und Selektion diene, sondern insbesondere auch der Förderung der Lernenden, kritisieren die englischen Lehrkräfte, dass die Testverfahren zu keiner ganzheitlichen und insofern auch keiner fairen Beurteilung führen, und favorisieren deshalb ihre eigenen Assessments. Kritische Studien zum Verfahren der Value-added-Messung stellen für den Sekundarbereich fest, dass Schulen mit einem hohen Leistungswert auch einen hohen Value-added-Wert haben, aber dass ein solcher Zusammenhang für Schulen mit mittlerem und niedrigem Leistungsniveau nicht gegeben sei. Zur Erklärung wäre man auf Inputfaktoren verwiesen, nicht auf den Unterricht (Gorard 2006).

Das englische Test- und Evaluationssystem soll zahlreiche Funktionen erfüllen. Der Lernstand der Schülerinnen und Schüler wird an vier Zeitpunkten engmaschig überprüft, die Qualität jeder einzelnen Schule wird regelmäßig extern evaluiert, die Entwicklung der professionellen Kompetenz der Lehrkräfte wird angeleitet, und statistisch beschreibbare nationale Trends werden in die Bildungspolitik eingegeben. Das wirft die Frage auf, wie die dabei erzeugten Datenströme koordiniert werden können und was mit der Inflation der Informationen angefangen werden soll. Quinlan/Scharaschkin (1999) schlagen vor, die Beurteilung der einzelnen Schülerinnen und Schüler wieder verstärkt in die Hände der Lehrpersonen zu geben und nationale Entwicklungen anhand stichprobenartig angelegter *large scale assessments* zu identifizieren. Der ehemalige Schulinspektor Woodhead (2004, S. 55f.) bezweifelt die Unabhängigkeit der Testbehörde und spricht von „inflated results“. In vergleichenden Studien wird darauf verwiesen, dass Finnland ohne externe Inspektionen und ganz auf interne Schulentwicklung bauend bessere Resultate erziele (Webb/Vulliamy/Hakkinen et al. 1998).

3.4.4 Qualitätssicherung und Evaluation

a) Die Akteure und Evaluationskonzepte der Zentralebene

Seit Beginn der neunziger Jahre kommt der Zentralebene im englischen Bildungssystem eine überragende Bedeutung für die Evaluation des Schulwesens zu. Der hauptsächliche Akteur ist das *Department for Education and Skills* (DfES), das auch die lokale Ebene zunehmend für Evaluationen und Kontrolle der Zielerreichung einsetzt. Das DfES hat seine Kompetenzbereiche nicht zuletzt auch durch die Etablierung der bereits erwähnten *Quasi Autonomous Non-Governmental Organizations* oder QUANGOs entscheidend ausgeweitet. Dazu gehören:

- das Office for Standards in Education (Ofsted),
- die Qualifications and Curriculum Authority (QCA) und
- die Teacher Training Agency (TTA).

Innerhalb des Department for Education and Skills wurden zudem zwei öffentlichkeitswirksame Spezialeinheiten gegründet, die *Standards Task Force* (STF) und der *Standards and Effectiveness Unit* (SEU). Die STF ist inzwischen aufgelöst und umorganisiert worden. Sie bestand aus hochkarätigen Vertretern des englischen Bildungswesens, Schulpraktikern und Repräsentanten der Wirtschaft. Die Mitglieder dieser Expertengruppe haben das Ministerium in Fragen der schulischen Evaluation beraten und verstanden sich auch selbst als Botschafter für herausragende Beispiele erfolgreicher Schulentwicklungsprozesse.

Die SEU¹³⁵ übernimmt die führende Rolle in der Koordination aller Maßnahmen und Strategien des Department for Education and Skills in der Evaluation des Schulwesens. Es gibt drei Kernbereiche im Aufgabenprofil der Einheit, nämlich:

- improving and sustaining standards of attainment by all pupils,
- promoting diversity and innovation and
- monitoring performance in education and intervening where necessary.

Eine wesentliche Aufgabe besteht darin, die Bemühungen der LEAs und der Schulen um Qualitätsentwicklung systematisch zu analysieren, Schulentwicklungsprozesse durch die LEAs anzuregen und kontinuierlich zu überprüfen.

¹³⁵ <http://www.standards.dfes.gov.uk/seu/> (Stand 05.08.06); alle folgenden Zitate stammen aus dieser Quelle.

Zudem ist die Einheit verantwortlich für eine Reihe von innovativen Politikstrategien, etwa die beiden erwähnten Programme zur Leseförderung und Verbesserung der mathematischen Grundkompetenz (NLS, NNS). Insgesamt lässt sich auch bei diesem Gremium die Strategie einer engeren Kooperation bzw. Kontrolle der LEAs durch das Ministerium erkennen.

Mit direktem Bezug auf die Einzelschule kann der Secretary of State, wenn eine Schule von Ofsted in die Kategorie „school requiring special measures“ eingestuft wird, eine unbegrenzte Zahl von Schulverwaltungsmitgliedern (*governors*) selbst ernennen sowie den Vorsitzenden des *governing body* bestimmen. Falls diese Maßnahmen nicht greifen, kann der Secretary of State die LEA auch anweisen, eine Schule zu schließen. Der zentrale Aspekt aber ist die Entwicklung auch von schlechten Schulen. Im Sommer 1998 wurden 515 „failing schools“ registriert, die Zahl ging auf 242 zu Beginn des Schuljahres 2005 zurück (Department for Education and Skills (DfES) 2005, S. 15).

Die vermutlich einflussreichste der drei unabhängigen Behörden ist die *Qualifications and Curriculum Authority* (QCA). Im Blick auf schulische Evaluation bestehen ihre Hauptaufgaben nach dem Education Act von 1997 in

- der permanenten Prüfung und Revision des National Curriculum,
- der systematischen Messung von Schülerleistungen an den Vorgaben des National Curriculum (National Curriculum Assessment) sowie der nationalen Auswertung dieser Ergebnisse und
- der Kooperation mit den Examination Boards in der Durchführung, Kontrolle und Regulierung der zu staatlichen Schulabschlüssen führenden staatlichen Prüfungen (public examinations).

Die Durchführung und Auswertung der National Curriculum Assessments ist das mit Abstand umfangreichste Unternehmen der QCA. Im Herbst erhalten die Schulen im sogenannten *Autumn Package* den umfassenden Bericht zum Leistungsstand sämtlicher Schülerinnen und Schüler. Der Bericht enthält die nationalen Ergebnisse und statistisches Datenmaterial zu Benchmarks und Value added, damit die Schulen ihre eigenen Ergebnisse in den School and College Achievement and Attainment Tables vor dem Hintergrund nationaler Trends und den Ergebnissen vergleichbarer Schulen interpretieren sowie Lernzuwächse der eigenen Schülerinnen und Schüler mit dem Durchschnitt vergleichen können. Neben den Autumn Packages und den Performance Tables gibt die QCA umfangreiche kommentierte *Standards Reports* heraus, die sich an die Lehrpersonen, die Schulleitung, das Ministerium und die interessierte Öffentlichkeit richten. Sie enthalten vornehmlich qualitative Analysen der Assessments auf den

Key Stages 1 bis 3 und geben auch Empfehlungen ab, wie die Ergebnisse in den Unterricht einbezogen werden können. Die *Evaluation Reports* schließlich enthalten im Wesentlichen die Wahrnehmung der Tests durch die Lehrkräfte und geben der Bildungsadministration Auskunft über deren Verlauf sowie etwaige Kritik.

Die externen staatlichen Abschlussprüfungen im Alter von 16 und 18 Jahren werden von sogenannten *Examination Boards* entwickelt und korrigiert. Um die Einheitlichkeit der Leistungsanforderungen unterschiedlicher Boards zu garantieren, hat die QCA die Aufgabe, die von den Boards vergebenen staatlichen Abschlüsse sowie die Prüfungen selbst kontinuierlich zu überprüfen und zu akkreditieren. Auch diese Kontrolle hat in den letzten Jahren deutlich zugenommen.

Die Ausbildung der Lehrkräfte soll ebenfalls zur schulischen Qualitätssicherung und -entwicklung beitragen. Die *Teacher Training Agency* (TTA) wurde 1994 als eine mit dem Department for Education and Skills kooperierende, ihr aber nicht direkt unterstellte Behörde gegründet. An erster Stelle stehen die Bemühungen der TTA um eine Verbesserung der Qualität in der Grundausbildung der Lehrkräfte (Initial Teacher Training, ITT). Die Agentur übernimmt die Akkreditierung der verschiedenen ITT-Programme, die an Universitäten und anderen Institutionen der höheren Bildung angeboten werden.¹³⁶ Die fortlaufende Qualitätsüberprüfung wird von Ofsted übernommen. Die Induktionsphase ist vom Ministerium per Gesetz geregelt. Auch dafür hat die TTA Standards definiert, die die Novizen im Beruf erfüllen müssen. Schließlich hat die TTA sogenannte Berufseingangsprofile entwickelt, die den Ausgangspunkt für die zielgerichtete Planung der professionellen Fortbildung der individuellen Lehrkräfte bilden sollen. Begleitet werden diese Maßnahmen zur Standardisierung der Lehrerausbildung durch nationale Kampagnen zur Rekrutierung hoch qualifizierter Kandidatinnen und Kandidaten für den Lehrberuf.

Die Lehrerinnen und Lehrer werden jährlich gemäß klar definierten Qualitätsstandards durch die Schulleitung evaluiert. Die Ergebnisse werden verbunden mit höheren, leistungsbezogenen Gehältern sowie der Möglichkeit einer schnelleren Beförderung. Für die Ausbildung der Schulleiterinnen und Schulleiter wurde ein National College for School Leadership gegründet. Schließlich ist auch die Idee umgesetzt worden, dass Schulleiterinnen und Schulleiter für die erfolgreiche befristete Übernahme der Leitung einer problematischen Schule ein höheres Gehalt bekommen. Nationale Standards für Headteachers sind

¹³⁶ Der größte Anbieter ist die Open University.

2004 veröffentlicht worden. Sie gehen von einem Kerngeschäft aus, das beschrieben wird als „to provide professional leadership and management for a school“ (Department for Education and Skills (DfES) 2004a, S. 3). Der Schulleiter oder die Schulleiterin ist der oder die führende Professionelle der betreffenden Schule. Ihre/Seine Aufgaben verlangen eine besondere Schulung. Die Schulleitung ist verantwortlich für

- evaluation of the school's performance to identify the priorities for continuous improvement and raising standards,
- ensuring equality of opportunity for all,
- developing policies and practices,
- ensuring that resources are efficiently and effectively used to achieve the school's aims and objectives, and
- the day-to-day management, organisation and administration of the school (ebd.).

Dass es selbst für Schulleitungen extensive Standards und „key-areas“ ihrer Aufgabenbereiche gibt (ebd., S. 4), hängt mit der Macht und dem Einfluss der dritten unabhängigen Behörde zusammen. Das *Office for Standards in Education* (Ofsted) wurde 1992 gegründet¹³⁷ und muss vor dem Hintergrund der wechselvollen Geschichte der Schulinspektion in England gesehen werden (Brighouse 1995). Verglichen mit dem früheren Regime des Her Majesty's Inspectorate (HMI) hat erst Ofsted Einfluss auf das Klassenzimmer nehmen können (Lee/Fitz 1997).

An der Spitze von Ofsted steht Her Majesty's Chief Inspector (HMCI). Die allgemeine Aufgabe von Ofsted besteht darin, die Standards und die Qualität des Schulwesens durch regelmäßige, unabhängige Inspektionen, durch öffentliche Berichterstattung und die Beratung des Bildungsministeriums zu verbessern. Seine spezielle Aufgabe besteht darin, ein effizientes Schulinspektionswesen aufzubauen und dafür zu sorgen, dass sämtliche öffentliche Schulen in England mindestens alle sechs Jahre inspiziert werden. Ofsted übernahm damit einen Teil der Aufgaben der ehemaligen LEA-Inspektorate sowie Teile von Her Majesty's Inspectorate, das seit 1886 existiert. Inspektionssysteme haben in England damit Tradition und sie sind grundsätzlich akzeptiert.

¹³⁷ <http://www.ofsted.gov.uk/> (Stand 05.08.06). Die folgenden Daten entstammen dieser Seite, wenn nicht anders vermerkt.

Die Behörde ist derzeit unterteilt in fünf Direktorate, die jeweils unabhängig voneinander geleitet werden. Zwei der Direktorate beziehen sich auf die innere Organisation, die Finanzen einerseits, die Expansion der Behörde andererseits. Vom 1. April 2007 an übernahm Ofsted weitere Inspektionsaufgaben im Bereich sozialer Dienste sowie in der Erwachsenenbildung.¹³⁸ Zu diesem Zweck ist ein fünftes Direktorat zur Reform der Inspektorate gegründet worden. Die drei hauptsächlichen Direktorate *Early Years*, *Education* und *Corporate Services* haben jeweils Außenstellen und/oder Unterabteilungen für verschiedene Aufgabenbereiche. Die Größe lässt sich an einem Beispiel zeigen: Die Direktion *Early Years* unterhält ein kleines Büro in London und beschäftigt rund 600 Mitarbeitende in den Außenstellen. Dazu kommen etwa 1100 Inspektoren, die von zu Hause aus arbeiten. Zu den Funktionsbereichen von Ofsted gehören auch die der Rekrutierung, Aus- und Weiterbildung von Inspektoren sowie die Qualitätskontrolle von privaten Agenturen (*contractors*), bei denen die Inspektoren unter Vertrag stehen.

Das englische Schulinspektorat ist in seiner Größe, Macht und den Ansprüchen einzigartig in ganz Europa. Ein derartig engmaschiges Netz von Kontrollen und Audits gibt es in keinem anderen Land. Dieser Service wird noch ausgebaut, wie der „strategische Entwicklungsplan“ zeigt, der im November 2004 veröffentlicht wurde. Der Plan listet zehn „objectives“ auf, die Expansion und Verbesserung der Nutzung anzeigen sollen (Ofsted 2004, S. 8ff.). Mit dem Regierungsantritt von Labour sollte die Schulinspektion neben der Kontrolle der Qualität des Unterrichts erneut eine – wenn auch eingeschränkte – beratende Dimension erhalten.¹³⁹ Zudem werden auch sämtliche LEAs einer regelmäßigen Inspektion durch Ofsted unterzogen.

Ofsted informiert die zu inspizierenden Schulen vier bis acht Monate im Voraus. Als Vorinformation vor dem Schulbesuch dienen den Inspektoren, jeweils ein Team von sechs bis acht Mitgliedern, das von der Schule zu erstellende Selbstporträt sowie Informationen aus der Ofsted-Datenbank, insbesondere zu den Leistungen der Schülerinnen und Schüler sowie der Schule insgesamt im

¹³⁸ Die Behörde heißt nun *Office for Standards in Education, Children's Services and Skills*. Der Name Ofsted bleibt erhalten.

¹³⁹ In einer international vergleichenden Studie der OECD zur schulischen Evaluation in sieben Staaten wurde 1995 bezüglich England in Frage gestellt, ob das Verhältnis von Beratung und Kontrolle richtig ausbalanciert sei. Festgestellt wurden disproportionale Anstrengungen zur Identifikation von Problemen, und zwar gegenüber den vorhandenen Kapazitäten der Schulen, darauf zu reagieren und diese auch tatsächlich anzugehen (OECD/CERI 1995).

lokalen und nationalen Vergleich. Im Vorfeld der Inspektion fanden erste Kontaktaufnahmen mit der Schule sowie mit dem *governing body* statt. Die Inspektion selber dauerte zu Beginn meist etwa eine Woche. Mit den revidierten Inspektionsrichtlinien von 2000 wurde dann zwischen „vollständigen“ Inspektionen und „kurzen“ Inspektionen für effektive Schulen unterschieden, um so den Verwaltungs- und Arbeitsaufwand sowohl für das Inspektorat als auch für die Schule zu reduzieren. Seit 2005 kommt ein neues System von Kurz-Inspektionen zum Einsatz. Damit werden sämtliche Schulen zwar wieder mindestens alle drei Jahre inspiziert, die Dauer und Intensität variiert jetzt allerdings stark je nach Leistungen und Erfolg einer Schule. Ein einzelner Inspektor bzw. eine einzelne Inspektorin oder ein Team von bis zu fünf Inspektorinnen und Inspektoren besuchen die Schule an einem bis maximal zwei Tagen; die Ankündigung des Besuchs erfolgt lediglich zwei bis fünf Tage im Voraus. Dabei muss die Schulleitung im Vorfeld ein Self Evaluation Form (SEF) ausfüllen, das die Inspektoren bei ihrem Besuch gewissermaßen anleitet, indem es den Fokus der Inspektion vorgibt und insbesondere der Beurteilung der Qualität der Schulleitung dient. Neben Unterrichtsbeobachtungen im Licht des National Curriculum werden auch Schülerarbeiten eingesehen und die Sicht der Lehrkräfte, Schülerinnen und Schüler, der Eltern und weiterer „Stakeholders“ eingeholt. Der Ablauf und die Richtlinien der Inspektion sind von Ofsted genau vorgegeben und für jedermann einsehbar.¹⁴⁰ Am Ende der Inspektion erhält die Schulleitung jeweils ein erstes mündliches Feedback zu den Ergebnissen.

Falls die Inspektoren übereinstimmend der Meinung sind, dass die Schule den Ansprüchen nicht genügt, und diese Einschätzung durch den HMCI bestätigt wird, werden Sondermaßnahmen ergriffen. Von den rund 22 000 öffentlichen Schulen in England, zu denen noch einige hundert Nursery Schools sowie etwa 1500 Spezialschulen der verschiedensten Art zu rechnen sind (Statistics of Education 2004, S. 17), haben seit Beginn der Ofsted-Inspektionen etwa 1500 den Anforderungen nicht genügt (Woodhead 2004, S. 56). Viele schlechte Schulen bestanden zuvor über Jahre, ohne dass irgendjemand tätig wurde. Erst die externe Evaluation und die Veröffentlichung der Daten sorgte für Wandel. „The trauma of public humiliation, coupled with a change in senior staff and some extra resources, led most failing schools to improve“ (ebd.).

¹⁴⁰ Siehe unter URL:
<http://www.ofsted.gov.uk/portal/site/Internet/menuitem.455968b0530071c4828a0d8308c08a0c/?vgnnextoid=8c53f6a35bb8c010VgnVCM1000003507640aRCRD>
(Stand 30.04.07).

Konkret bedeutet dies, dass die Schule vom Ofsted School Improvement Team so lange betreut und kontrolliert wird, bis die Inspektoren bei einem erneuten Besuch der Meinung sind, dass die Spezialmassnahmen nicht mehr notwendig sind. Falls dies nicht der Fall ist, übernimmt eine sogenannte Education Association die Leitung der Schule, und wenn auch dies nicht den notwendigen Erfolg zeigt, kann die Schule geschlossen werden (Education Act 1993). Unter Labour wurde eine zweite Kategorie förderbedürftiger schwacher Schulen definiert (schools with serious weaknesses), die ebenfalls eine intensive Beratung durch Ofsteds School Improvement Teams erhalten. Etwa 14% der Schulen gehören dieser Kategorie an. Darin zeigt sich wiederum die erwähnte Tendenz, der Schulinspektion vermehrt auch beratende Funktion zukommen zu lassen.

Fünf Wochen nach der Inspektion muss der Inspektionsbericht vorliegen, der an die Schule, den *governing body* und die verantwortliche LEA geht. Die Schule ist verpflichtet, den Bericht den Eltern und der gesamten Öffentlichkeit zugänglich zu machen. Die Hauptergebnisse der Inspektion sowie die daraus resultierenden *key issues for action* bilden die Basis für den sogenannten *action plan*, den der *governing body* innerhalb von 40 Tagen nach Veröffentlichung des Berichts erstellen muss. Auch der *action plan* muss an alle Eltern, das Kollegium, Ofsted und die zuständige LEA verschickt werden.

Ein Beispiel für einen *inspection report* ist der Bericht vom Juli 2005 über die William Tyndale Community Primary School in Islington, der wie alle Berichte öffentlich zugänglich ist. Der Bericht vermerkt, dass die Schule seit der letzten Inspektion eine „troubled and turbulent time“ erlebt habe, verursacht vor allem durch den häufigen Wechsel der Lehrkräfte; inzwischen sei aber eine zufriedenstellende Periode „of staffing stability“ festgestellt worden. Das Umfeld der Schule – „a very wide mix of social and ethnic backgrounds“ – sei bei der jetzigen Inspektion berücksichtigt worden. Die Inspektion erteilt Noten (grades) auf einer Viererskala.¹⁴¹ Für die William Tyndale School sind sie nicht allzu gut. Die „overall effectiveness“ wird mit „befriedigend“ eingestuft, besonders in den unteren Klassen lassen die Leistungen der Schülerinnen und Schüler zu wünschen übrig. Die Standarderreicherung und vor allem die Lernfortschritte sind „inadäquat“. Teaching and Learning allgemein sind „befriedigend“ wie die meisten anderen Kategorien; „gut“ ist einzig „personal development and well-being“ (Ofsted 2005).

Die Leistungsdaten der Schule sind unabhängig vom Inspektionsbericht in den Medien öffentlich zugänglich. Dargestellt werden etwa von den BBC News

¹⁴¹ Grade 1: Outstanding. Grade 2: Good. Grade 3: Satisfactory. Grade 4: Inadequate.

die durchschnittlichen Ergebnisse der Leistungstests in den Kernfächern vor dem Hintergrund des nationalen Durchschnitts sowie des Durchschnitts der LEA, der die Schule angehört. Die Schule wird weiter mit Kennziffern beschrieben, zu denen auch die Zahl der Absenzen gehört. Im Weiteren kann der Platz im Ranking eingesehen werden. Und schließlich wird die Leistungsentwicklung der letzten vier Jahre präsentiert, indem die Veränderung der Gesamtpunktzahl in den Tests bezogen auf das Jahr der Erreichung dargestellt wird.¹⁴²

b) Die Akteure und Evaluationskonzepte der Mesoebene

Ein wesentlicher Teil der Verwaltung und praktischen Gestaltung des englischen Schulwesens findet in den LEAs (Local Education Authority) statt, die zur mittleren Verwaltungsebene gehören und Teil der Regional- und Kommunalverwaltung sind. Die Größe der 150 LEAs variiert zwischen ländlichen und städtischen Gebieten sehr stark, damit auch der Grad der Professionalisierung und das Ansehen vor Ort. Mit Ausnahme der Funktionen, die von den Einzelschulen im Rahmen des *local management of schools* (LMS) selbstverantwortlich ausgeführt werden, sind die LEAs für organisatorische, administrative Bereiche und die globalen Schulbudgets verantwortlich. Nach dem Regierungsantritt von Labour sind LEAs auch dafür verantwortlich, Schulentwicklungsprozesse in Gang zu setzen, und zwar auf der Grundlage der Leistungsergebnisse ihrer Schulen. Die Einzelschulen entwickeln in Kooperation mit der LEA Ziele (*school targets*), deren Implementierung jährlich durch die LEA überprüft wird.

Zur Diskussion steht die Frage, wie die LEAs dieser neuen konstruktiven Rolle im schulischen Evaluationsprozess gerecht werden können. Die LEAs fungieren wie die Schulen zunächst als Empfänger national produzierter Daten, filtern diese und fügen darüber hinaus Informationen aus eigenen Quellen hinzu. Durch die nicht standardisierte Arbeit der LEAs, die offenbar sehr unterschiedlich mit dieser Aufgabe umgehen, verfügen die Schulen über Daten unterschiedlicher Qualität. Auch die Unterstützung im Umgang mit den Daten als Teil der Rückmeldung differiert. Es sind besonders die Klassenlehrkräfte, weniger die Schulleitungen, die im Bereich der Datenanalyse und -interpretation einen zusätzlichen Unterstützungsbedarf äußern. Gemäß der Aussage eines Leiters der Ofsted-Abteilung für LEA-Inspektionen war die Leistungsfähigkeit der

¹⁴² Wieder ist das Beispiel die William Tyndale Primary School:
http://news.bbc.co.uk/1/shared/bsp/hi/education/04/school_tables/primary_schools/ht (Stand 11.08.06).

einzelnen LEAs noch zu Beginn der zweiten Legislatur von New Labour sehr unterschiedlich. Er beklagte den Mangel an Standards für die Evaluation durch die LEAs (Kotthoff 2003, S. 137f.), ein Mangel, der inzwischen behoben ist, was nicht heißt, dass alle LEAs gemäß den gleichen Standards arbeiten.

Als wichtigste evaluative Maßnahme fungiert der *Education Development Plan* (EDP) der LEA, der, wie gesagt, dem Ministerium vorgelegt werden muss. Die LEAs erhalten bei der Erarbeitung des Planes die Unterstützung des Standards and Effectiveness Unit (SEU). Im EDP ist das Schulentwicklungsprogramm (*school improvement programme*) enthalten. Darin muss die LEA konkrete Angaben und Zielsetzungen formulieren, die sich auf die Leistungsergebnisse und Lernfortschritte der Schülerinnen und Schüler gemessen an den Vorgaben des National Curriculum bzw. den nationalen Assessments beziehen, weiter zur Qualität der Lehr- und Lernprozesse sowie zur Qualität der Schulleitung bzw. des Managements.

Ähnlich wie die Einzelschulen durch die Formulierung und jährliche Überprüfung von individuellen und überprüfbaren Entwicklungszielen (*school targets*) ihrer zuständigen LEA gegenüber Rechenschaft ablegen müssen, sind auch die LEAs dazu angehalten, die Qualität ihrer Arbeit gegenüber dem Ministerium und der Öffentlichkeit durch Entwicklung und kontinuierliche Revision ihrer Education Development Plans unter Beweis zu stellen, was eine fortlaufend angereicherte Datenbasis voraussetzt. Seit 1998 werden die LEAs durch Ofsted inspiziert. Die 1999 überarbeiteten Richtlinien der Inspektion beziehen sich explizit auf die von den LEAs vorzulegenden Education Development Plans als dem zentralen Ausgangspunkt der Überprüfung.

Ein Problem ist, dass in bestimmten LEAs aufgrund der Einschätzung der Inspektoren effektive Konzepte für die Evaluation ihrer Schulen entwickelt wurden, aber dass keine Erhöhung der schulischen Standards festgestellt werden konnte. Im Anschluss daran wird die Frage gestellt, ob die regelmäßigen, aufwändigen und teuren Inspektionen der LEAs durch Ofsted ihr Ziel, die Steigerung der Schulqualität, tatsächlich erfüllen. Fraglich ist aber auch, ob die LEAs einen direkten Einfluss auf die Qualität der Leistungsergebnisse der Schulen haben können. In den Inspektionen der LEAs werden die Test- und Prüfungsergebnisse der Schülerinnen und Schüler sowie die Ofsted-Inspektionsergebnisse der Schulen ihres Zuständigkeitsbereichs als wichtigste Indikatoren für die Qualität der LEAs selber herangezogen.

Die Beurteilung der Lehrerbildung durch die Ofsted-Teams ist früh einer Kritik unterzogen worden, die vor allem die Konsistenz in der Anwendung der Kriterien in Frage stellte (Jones/Sinkinson 2000). Der hohe subjektive Anteil in den Beurteilungen ist auch im Blick auf die Schulinspektionen herausgestellt

worden (Richards 2001). Die subjektive Seite der Inspektionen bezieht sich auch auf die Erfahrungen der Lehrkräfte, die sich persönlich bedroht fühlten und oft mit Abwehr reagiert haben (Holmes 2005).

c) Schulische Evaluation auf der Ebene der Einzelschule

Die hohe Autonomie der Einzelschulen in England ist im Laufe der neunziger Jahre im Blick auf ihre Curricula sowie ihre Evaluation stark eingeschränkt worden. Dies ist auf die rasche und systematische Entwicklung eines zentralen nationalen Systems der Qualitätssicherung zurückzuführen, das in Europa ohne Beispiel ist. Während in Finnland etwa das langjährige, fächerzentrierte nationale Curriculum zugunsten schulischer Curricula abgeschafft wurde, ist in England der genau umgekehrte Prozess zu beobachten. Zudem ist die einzelne Schule zu einem Objekt komplizierter Evaluationsmaßnahmen geworden. Auf der anderen Seite ist die operative Verantwortung der Schulen gestärkt worden, bis hin zur jüngsten Einführung der Trust Schools.

Seit dem Antritt der Labour-Regierung ist eine Tendenz vorhanden, die Einzelschule wieder verstärkt als aktiven Partner in den Prozess der Schulevaluation mit einzubeziehen. Insgesamt hat in den letzten Jahren die schulinterne Evaluation erheblich an Bedeutung gewonnen. Angestrebt wird damit die Wiederherstellung einer Balance zwischen externer Kontrolle der Einzelschule durch Schulinspektionen und interner Evaluation durch die Einzelschule. Die konservative Politik bot den Schulen keine oder nur mangelhafte Unterstützung durch regionale Stellen bzw. Verwaltungseinheiten. Gleichzeitig war bei den Schulen ein Problem unübersehbar, sie hatten in der Vergangenheit die Leistungsanforderungen an die Schülerinnen und Schüler zu tief angesetzt.

Um die Qualitätsentwicklung durch die Schulen selber voranzutreiben, wurden die Schulen angehalten, in Zukunft anspruchsvolle schulinterne Entwicklungsziele (*school targets*) in Bezug auf die Schülerleistungen zu setzen. Dabei sollten die *governing bodies* eine entscheidende Rolle bei der Entwicklung und kontinuierlichen Evaluation dieser Ziele spielen. Die *school targets* können sich auf sämtliche Bereiche und Aspekte des Schullebens beziehen. Gesetzlich vorgeschrieben sind nur solche Ziele, die die Anforderungen des National Curriculum bzw. die nationalen Tests und Prüfungen betreffen. Die Entwicklung und Formulierung von *school targets* sollte mit dem bereits in allen Schulen stattfindenden Prozess der Erarbeitung/Revision von Schulentwicklungsplänen (School Development Plan) eng verknüpft werden.

Die schulische Selbstevaluation wurde 1998 unter dem programmatischen Titel „School Evaluation Matters“ von Ofsted in den Vordergrund gerückt.

Der Schwerpunkt der Selbstevaluation wurde dabei eindeutig auf die Leistungsergebnisse der Schülerinnen und Schüler gelegt. Das von Ofsted vorgelegte Konzept zur Selbstevaluation von Einzelschulen lehnt sich in seiner Anlage und bezogen auf die Qualitätsindikatoren sehr deutlich an die Strukturen der Ofsted-Inspektionen an. Die Behörde hat zusammen mit der Standards and Effectiveness Unit 1999 einen Trainingskurs für Schulleiter bzw. Schulmanager – „Helping schools to carry out self-evaluation“ – konzipiert, der sehr große Popularität erzielte. Dabei werden den Teilnehmenden in einem praktisch orientierten Seminar Kriterien und Konzepte sowie Methoden und Techniken der Selbstevaluation vermittelt, die sich weitestgehend an den Inspektionsrichtlinien von Ofsted orientieren. Schulen sollen dadurch befähigt werden, die erzielten Leistungsstandards, die Qualität des Unterrichts und des Schulmanagements aktiv zu interpretieren und anzunehmen. Auch soll der Kurs dazu beitragen, Ergebnisse und Empfehlungen aus Ofsted-Inspektionen produktiv in einen *action plan* umsetzen zu können.

Neben den zahlreichen qualitativen Studien zu den Wirkungen der Schulinspektion erscheint die Zahl der quantitativen Untersuchungen relativ gering. Diese zeigen, dass der überaus größte Teil der Schulleiterinnen und Schulleiter der Inspektion positiv gegenübersteht und sie als förderlich für die Schulentwicklung einschätzt. Für die Lehrerinnen und Lehrer bedeuten Inspektionen hingegen in erster Linie eine Stressbelastung, die mit abnehmendem Status in der Schulhierarchie zunimmt. Sie berichten, dass die Qualität ihres Unterrichts während der Inspektion gegenüber der Alltagssituation deutlich verschieden sei, und dies sowohl im positiven wie negativen Sinne. Kritik erzeugt zum Beispiel, dass eine persönliche Rückmeldung direkt nach der Unterrichtsbeobachtung durch die Inspektoren meistens fehlt. Feedbacks beschränken sich in vielen Fällen auf die Mitglieder der Schulleitung. Aber das heißt nicht, dass Inspektionen für die Lehrkräfte ausschließlich negative Effekte haben müssen.

Kogan/Maden (1999) stellen Ergebnisse einer von ihnen durchgeführten Evaluationsstudie zum Ofsted-Inspektionssystem dar. Folgende Effekte vor und nach der Inspektion werden beschrieben:

Vor der Inspektion:

- *Vorteile:* Einleitung von Selbstüberprüfungsprozessen zur Vorbereitung der Inspektionswoche; verstärkte gegenseitige Unterstützung im Kollegium.
- *Nachteile:* Das System wird als bestrafend angesehen und erzeugt ein ängstliches Klima mit enormem Stress für Lehrer und Lernende.

Nach der Inspektion:

- *Vorteile:* Änderung der Managementstrukturen, verbesserte Kommunikations- und Interaktionsstrukturen, Änderung der Unterrichtsmethoden und der curricularen Organisation; verstärkte Schulentwicklungsaktivitäten.
- *Nachteile:* Größere Fehlzeiten der Lehrkräfte aufgrund von Stress, vor allem in Schulen, die negativ eingestuft worden sind; höhere Arbeitszeitbelastung wiederum in Schulen, die schlecht abgeschnitten hatten. An Schulen, an denen spezielle Maßnahmen eingeleitet wurden, zeigt sich offenbar kein dauerhafter Effekt im Sinne eines verbesserten Self-Monitorings.

Van Ackeren (2003a, S. 228) notiert unter den positiven Effekten, die sich in England im Laufe der Zeit mit der Verfügbarkeit extern erhobener quantitativer Daten einstellen: einen vertrauteren Umgang mit den Daten, mehr Diskussionen über Fragen, die die Daten aufwerfen, den verstärkten Gebrauch der Ergebnisse für die curriculare Planung auf schulischer und Klassenebene sowie den Nutzen für die Schülerinnen und Schüler, indem die Daten der Einschätzung des eigenen Leistungsstandes dienen.

Von entscheidender Bedeutung für die Wirkung der Schulinspektion sind, so die Fallbeispiele in Kotthoff (2003), die Qualität der Schule und des Schulmanagements sowie die Schulkultur. Geringe Effekte sind für Schulen festzustellen, die bezüglich Leistung und sozio-ökonomischem Hintergrund der Schülerpopulation bereits über dem Durchschnitt liegen bzw. privilegiert sind. Andere Untersuchungen zeigen aber, dass der Einfluss auch auf durchschnittliche Schulen, die sich von der Inspektion aufgrund ihrer Leistungen nicht unmittelbar bedroht sehen, eher gering ist. Einerseits stehen die Eltern den Inspektionen und den daraus für sie erhältlichen Informationen positiv gegenüber. Gleichzeitig werden die hohen Kosten¹⁴³ angesichts der aus der Sicht der Eltern jeweils wenig überraschenden Ergebnisse kritisiert. Es gibt Bedenken

¹⁴³ 1998 lag der Etat von Ofsted bei £ 151 Mio. Die Inspektion einer mittelgroßen Grundschule kostet ca. 26 000 £, einer mittleren Sekundarschule ca. 65 000 £ (Ackeren v. 2003a, S. 168).

gegenüber einer zu weitgehenden Berücksichtigung der Elterninteressen, wenn dies die Professionalität der Lehrkräfte unterminieren würde.

Die Inspektoren selber äußern sich sehr positiv über die ihnen zur Verfügung gestellten Materialien (Inspektionshandbuch) und die Konkretion und Vielfalt der verwendeten Qualitätsindikatoren. Bedenken sind feststellbar gegenüber der Praktikabilität und dem Verwaltungsaufwand des Verfahrens. Die Schülerinnen und Schüler schließlich stehen gemäß einer Untersuchung den Inspektionen eher positiv gegenüber. Von Lehrerinnen und Lehrern werden deutliche Verhaltensänderungen ihrer Schülerinnen und Schüler während der Unterrichtsbeobachtung registriert. Die grundlegende Haltung der Abwehr nicht gegenüber der Rechenschaftslegung an sich, sondern gegenüber der Art und Weise, wie sie geschieht, bleibt bei vielen Lehrkräften bis heute erhalten.

Insgesamt wird von Kritikern am englischen Evaluationssystem hervorgehoben, dass eine Dominanz qualitätssichernder gegenüber qualitätsfördernder Maßnahmen herrsche. An der den Evaluationsmaßnahmen zugrunde liegenden Marktorientierung des Schulsystems ist früh kritisiert worden, dass diese die Bildungschancen benachteiligter Bevölkerungsgruppen weiter verschlechtere (Ball 1994). In Zweifel gezogen wurde auch die Annahme, dass der „educational quasi-market“ überhaupt zu einer Qualitätsoptimierung führe (z.B. Whitty/Power/Halpin 1998). Hargreaves (1995) zufolge verbinden sich mit der in England zu beobachtenden extensiven Formalisierung und Messung schulischer Qualität durch quantitative Leistungsindikatoren und die umfassende Inspektion aller Schulen massive Einschränkungen der zuvor weitreichenden Autonomie der englischen Lehrerschaft.

3.5 Auswertung des Ländervergleichs

Die vier Länder weisen unterschiedliche Bildungssysteme auf, deren Entwicklungen keineswegs konvergieren, zum Teil aber in identische Richtungen gehen. Das betrifft die Konstruktion der Curricula ebenso wie die Beurteilung und Bewertung der Schülerleistungen, die Maximen der Schulentwicklung, die Formen der Schulevaluation, die Stellung der Lehrkräfte, die Entwicklung der Schul- und Bildungsgesetze und auch die Kooperationen zwischen der Bildungspolitik, der Forschung und der Profession der Lehrerinnen und Lehrer. In allen diesen Bereichen finden sich Unterschiede, aber auch vergleichbare Tendenzen. Besonders im Bereich der curricularen Steuerung deuten die Entwicklungen in einer Studie zu zahlreichen Ländern auf „a repeating chain of action-reaction“ (Kuiper/Akker v.d./Hooghoff 2005, S. 76), wie etwa auch für

Schweden erkennbar ist; dabei trifft man verschiedentlich das Phänomen, dass mit abnehmender Zahl und Konkretion der (nationalen) Ziele zu deren Vermittlung und Umsetzung im Unterricht andere Curriculum-Dokumente allgemeinverbindlicher oder schulspezifischer Art zum Einsatz gelangen. Offenbar können auch Dezentralisierung und Deregulierung an ihre Grenzen stoßen, was zudem daran zu erkennen ist, dass „a too strong movement towards decentralization in the end results in a movement in the opposite direction (more top-down steering and control), and the other way around“ (ebd.).

Curricula können sich auf die Bestimmung der Fächer und Kernziele beschränken, aber sie liegen auch in Form von inhaltlichen Lernprogrammen vor. Die Schülerleistungen werden durchgehend getestet, aber die Art und Weise des Testens ist verschieden, ebenso die mit den Resultaten verbundenen Konsequenzen. Zudem unterscheiden sich der Anteil und die Bedeutung der Lehrerurteile. Auffällig ist, dass die Selbstbewertungen der Schülerinnen und Schüler wenn, dann eine nachgeordnete Rolle spielen. Die bildungspolitischen Programme unterscheiden sich in der Nähe zur Profession, entsprechend sind auch externe und interne Evaluationen unterschiedlich gewichtet. „Qualitätssicherung“ und „Steuerung durch Standards“ kann daher bei ähnlicher Rhetorik durchaus Verschiedenes heißen.

Die entwickeltste und stärkste Form der Steuerung unter den europäischen Ländern weist England auf. Zwischen 1988 und 2006 ist ein System entstanden, das nicht nur von einem in allen Fächern und auf allen Stufen verbindlichen nationalen Curriculum ausgeht, sondern das die Leistungen der Schülerinnen und Schüler an fünf Punkten misst, davon an vier mit Tests und/oder externen Prüfungen. Alle Teile des Systems sind standardisiert und überall ist die Entwicklung auf Zielsteuerung eingestellt. Die Schulen sind operativ autonom, aber zugleich einem engmaschigen Netz von Vorschriften und Kontrollen unterworfen. Auffällig sind auch die rasche Folge immer neuer Gesetze und die Lancierung immer neuer bildungspolitischer Kampagnen. Die Datenflut erscheint kaum beherrschbar.

Die wichtigsten Akteure auf Seiten der Behörden sind deutlich ökonomisch ausgerichtet. Anders als beim *Plowden-Report* von 1967 spielen Pädagogik und Psychologie kaum noch eine Rolle. Den Strategiewechsel in der Bildungspolitik haben in den achtziger Jahren Thinktanks wie das Adam-Smith-Institute ausgelöst. Eine marktwirtschaftliche Sicht war auch dem englischen Bildungssystem sehr fremd. Entsprechend groß waren die Schocks über den Education Reform Act von 1988, ohne dass dieser wirklich eine Marktöffnung bewirkt hätte. Er war der Einstieg für die staatlichen Schulen in Zielsteuerung und Qualitätssicherung mit Standards, nicht, wie beabsichtigt, in eine allmähliche Privatisie-

rung. Der Sektor der Privatschulen vergrößerte sich kaum und die GM-Schulen wurden durch New Labour wieder abgeschafft.

In der Literatur sprechen verschiedene Autoren von einem „quasi-market“. Die staatlichen Schulen wurden unter Effizienzdruck gesetzt, zum Teil wurde auch die lokale Aufsicht privatisiert, aber gleichzeitig entstand ein System von dichten Kontrollen durch Behörden, die im staatlichen Auftrag agieren. Die grundlegende Philosophie der rigorosen Output-Steuerung blieb auch beim Regierungswechsel erhalten, wenngleich im Blick auf die Profession nachgebessert wurde. Sprichwörtlich wurde „the iron hand of Ofsted“ (Davis 2002, S. 279), wobei eher die Lehrkräfte als die Schulleitungen oder die lokale Aufsicht gegen das System der Audits eingestellt sind. Und auch die Lehrkräfte wollen überwiegend nicht zurück zu einem Unterricht, der, wie früher, auf Standards verzichtet. Die Berufung auf Standards ist ein Gewinn in der Professionalisierung, kritisch beurteilt werden High-Stakes-Tests und das Schul-Ranking.

Die Entwicklung in den Vereinigten Staaten ist ebenfalls durch einen Politikwechsel gekennzeichnet. Nach dem Krisenreport *A Nation at Risk* (1983) setzte sich eine Politik durch, die auf der Ebene der Bundesstaaten auf curriculare Bildungsstandards setzt und die mit dem Bundesgesetz *No Child Left Behind* (2001) eine bislang nicht gekannte Steuerung durch den Bund erfahren hat. Die Bundesgelder werden nicht mehr, wie bislang, subsidiär eingesetzt, vielmehr folgt die Verteilung Anreizen, die einen standardbasierten Unterricht sowie Kontrolle durch High-Stakes-Tests durchsetzen sollen.

In den neunziger Jahren entwickelte sich eine regelrechte Standard-Bewegung, die von vielen Berufsverbänden unterstützt wurde. Ein nationales Curriculum, wie in England, konnte allerdings nicht eingeführt werden. Nach wie vor sind die Bundesstaaten für die öffentlichen Schulen zuständig, damit auch für die Curriculumentwicklung und die Strategien der Qualitätssicherung. Beides wird sehr unterschiedlich in Angriff genommen; während Staaten wie Texas ähnlich rigoros vorgehen wie England, sind andere sehr viel zurückhaltender. Es gibt nationale Leistungstests für drei Stufen, und auch ein nationales Reporting, aber keine nationale Schulentwicklung. Die Konzepte allerdings sind in zwischen sehr ähnlich.

Ein Kernkonflikt ist derjenige zwischen Politik und Profession. Die Bildungspolitik ist zunehmend geprägt von betriebswirtschaftlichen Konzepten der Accountability und setzt auf statistische Verfahren der Leistungsmessung. Der Konflikt entzündet sich vor allem an der Frage der *High-Stakes-Tests*, die mit positiven oder negativen Sanktionen verbunden sind. Die Lehrerverprofession bekämpft vor allem die Konsequenzen der Tests, nicht oder nicht primär diese selbst. Die Auswirkungen von High-Stakes-Tests auf Schulen und Lehrkräfte

sind in der Literatur oft negativ beschrieben worden, während die Vorbereitung auf Tests (Teaching to the Test) keineswegs nur negative Seiten hat. Gute Testaufgaben können den Unterricht befördern, ebenso gute Rückmeldesysteme. Es kommt darauf an, welchen Zweck die Tests erfüllen sollen.

Im Blick auf die Implementation von Bildungsstandards liegen relativ klare Befunde vor. Top-down-Strategien, die kaum Unterstützungssysteme kennen, sind wenig erfolgreich. Wer Standards allein mit Gesetzen und Verordnungen durchsetzen will, wird auch dann die eigenen Reformziele nicht erreichen, wenn, wie in England nach 1988, ein engmaschiges Kontrollsystem aufgebaut wird. Das von der konservativen Regierung aufgebaute System hatte die Lehrerschaft gegen sich, was die Versuche der Labour-Regierung nach 1997 erklärt, die Basis zurückzugewinnen und die lokale Aufsicht als Teil der neuen Steuerung zu verstehen.

In den beiden kontinentaleuropäischen Systemen, Niederlande und Schweden, sind vor allem Dezentralisierungstendenzen zu beobachten gewesen. Das Bildungssystem in den Niederlanden war traditionell immer schon vergleichsweise autonom; diese Tendenzen sind nochmals, etwa durch Pauschalbudgets der Einzelschulen, verstärkt worden. In Schweden bestand jahrzehntlang eine zentrale Schulsteuerung, die im Zuge einer allgemeinen Verwaltungsreform aufgelöst wurde. Heute gibt es eine lokale Steuerung der Schulentwicklung in den Gemeinden mit einigen zentralen Elementen wie der Abschlussprüfung oder allgemeinen Lernzielen.

In beiden Ländern sind Bestrebungen erkennbar, curriculare Standards zu entwickeln, nicht jedoch ein engmaschiges, zentrales System der Überwachung aufzubauen. Die Politikstrategien in den Niederlanden wie in Schweden sind vergleichsweise professionsnah, anders kann eine lokale Steuerung auf Gemeindeebene auch kaum sehr erfolgreich sein. Die *Local Education Authorities* (LEAs) in England sind damit nicht vergleichbar, weil sie Distrikte und nicht Gemeinden umfassen. Ein Indiz für die These der Qualitätssicherung durch vermehrte Schulautonomie und verbesserte lokale Kontrolle ergibt sich aus der Tatsache, dass auch die LEAs inzwischen mit der Implementation von Bildungsstandards befasst sind.

Zwischen allen vier Ländern gibt es insofern eine Gemeinsamkeit, als die Bildungspolitik inzwischen stark vom Konzept der *School Choice* bestimmt wird. In den Niederlanden ist die Schulwahl schon seit längerem freigestellt und stellt kein bildungspolitisches Programm mehr dar. Faktisch handelt es sich um ein Voucher-System; die Eltern und Schülerinnen und Schüler können zwischen privaten und öffentlichen Schulen wählen, die voll aus dem Steueraufkommen finanziert werden. In Schweden existiert inzwischen ein ähnliches System, das

Wahlen zwischen staatlichen und unabhängigen Schulen zulässt, die alle vom Staat unterhalten werden. In den Vereinigten Staaten nimmt die Schulwahl zu und in England steht sie seit 2005 als Priorität des Regierungshandelns auf der bildungspolitischen Agenda.

Die Erfahrungen in Schweden zeigen, dass mit der Freistellung der Schulwahl auch unter den Bedingungen der nicht-selektiven Gesamtschule die Segregation zunimmt. Allerdings sind auch Leistungsverbesserungen durch mehr Wettbewerb zwischen den Schulen zu verzeichnen. Dafür sprechen auch amerikanische Evidenzen (Hoxby 2002; Hoxby 2003; siehe aber Ladd 2003). Der gezielte Einsatz von Stipendien unter der Voraussetzung von Schulwahl und Wettbewerb steigert gemäß einigen Studien die Leistungen vor allem von Schülerinnen und Schülern aus unterprivilegierten Familien (Wolf/Petersen/ West 2001). Allerdings sind die Befunde nicht eindeutig, weil die Voucher- und Choice-Programme bislang recht klein und limitiert sind (Gill/Timpani/Ross et al. 2001).

In England ist das jetzige System nicht mehr einfach eines des „neighbourhood schooling“. Die Hälfte der Schülerinnen und Schüler besucht nicht die lokale Schule und 30% gehen nicht in die drei nächstgelegenen Schulen ihres Wohnortes. Für die meisten Schülerinnen und Schüler würden Wahlmöglichkeiten bestehen, allerdings ist die Kapazität zu wählen ganz unterschiedlich, vor allem wegen der Transport- und Wohnkosten. Nicht alle Eltern werden wählen, aber der Weggang eines Schülers hat Konsequenzen für alle anderen.

Nachbarschaftsverschulung produziert starke Effekte der Sortierung der Lernenden; dabei ist die Sortierung nach Ethnien viel stärker als die nach Fähigkeiten oder Einkommen. Die Niveaus der Sortierung liegen höher in selektiven LEAs. Gegenden mit größeren Wahlmöglichkeiten verzeichnen auch eine stärkere Sortierung der Schülerinnen und Schüler. Die Evidenz, dass durch mehr Wettbewerb zwischen den Schulen die Leistungen der Schülerinnen und Schüler verbessert werden, ist in England wesentlich weniger stark als in den Vereinigten Staaten. Wer Schülerinnen und Schüler verliert, die die Schule abwählen, erhält weniger Gelder, aber das wirkt sich in England bislang anders aus als in bestimmten amerikanischen Bundesstaaten (Burgess 2005).

Im deutschen Sprachraum gibt es keine vergleichbare Entwicklung. Es ist kein Zufall, dass Schulen aus den vier Auftragsländern im jüngsten Strategiepapier der englischen Regierung (Department for Education and Skills/Prime Minister's Strategy Unit 2006) keine Rolle spielen. *School Choice* ist in der Schweiz, Deutschland, Luxemburg und Österreich bislang kein bildungspolitisches Thema. In Schweden gingen im Schuljahr 2004/05 bereits 7% aller Schülerinnen und Schüler in unabhängige Schulen (ebd., S. 12). In Ländern wie Dä-

nemark ist die Schulwahl freigestellt, weil die Eltern über die Erziehung ihrer Kinder ultimativ entscheiden (ebd., S. 43). Und Länder wie Norwegen haben unter dem Druck von nicht zufriedenstellenden PISA-Ergebnissen die Gründung staatlich finanzierter, unabhängiger Schulen erleichtert (ebd., S. 50f.). Freie Schulwahl gibt es seit 1998 auch in Finnland, allerdings besuchen zurzeit nur 2% der Schülerinnen und Schüler unabhängige Schulen (ebd., 55ff.).

Die politische Kernfrage ist, ob die diversen Maßnahmen zur Qualitätssicherung zu einer empirisch fassbaren Verbesserung führen oder nicht. Die Frage ist wegen der hohen Zielsetzungen und auch wegen der Kosten der Reformen heikel. Damit verbunden ist ein Rechtfertigungsdruck, der zu selektiven Erfolgsmeldungen führen kann, die sich von einer anderen Seite aus und mit anderen Zahlen leicht bestreiten lassen. Die expansive Bildungsforschung führt zu einer gewissen Unübersichtlichkeit der Resultate. Oft sind die Resultate auch gegensätzlich, und es stellt sich die Frage, wie Daten und Ergebnisse politisch verwendet werden sollen. Wenn sich, wie in den Vereinigten Staaten, zwei Parteien gegenüberstehen, kann sich jede auf „ihre“ Forschung berufen.

Geht man von Merkmalen der Systementwicklung aus, dann gibt es besonders zwischen den drei europäischen Ländern weitere Gemeinsamkeiten. Alle verwenden in irgendeiner Form Tests oder externe Prüfungen. Es gibt überall interne Schulevaluationen und externe Inspektorate, die unterschiedlich organisiert sind, aber einen vergleichbaren Zweck erfüllen. Die curricularen Standards sind verschieden ausgerichtet, einmal als enge Vorgaben wie das nationale Curriculum in England, einmal als Kernziele wie in den Niederlanden, aber es sind immer Standards im Sinne verbindlicher Erwartungen. Die Art und Weise, wie die Standarderreicherung kontrolliert werden soll, ist wiederum unterschiedlich, aber *dass* im Sinne einer Rechenschaftspflicht kontrolliert werden muss, ist unstrittig.

Sehr verschieden sind die je gewählten Strategien, Standards verbindlich zu machen. Grob lassen sich zwei Muster unterscheiden, eines, das auf gouvernementalen Entscheidungen und die Tätigkeit von Behörden setzt, und ein zweites, das von den Akteuren ausgeht und deren Know-how zu verbessern sucht. Die beiden Muster spiegeln sich in der englischen Diskussion. Die Labour-Regierung hat wohl versucht, die Lehrkräfte einzubinden, aber sie hat die grundlegende Philosophie der Steuerung nicht verändert. Diese Philosophie ist gouvernemental, eher angelegt auf Kontrolle als auf Entwicklung. Dass beides vereint

werden müsse, fordert die Lehrgewerkschaft NUT:¹⁴⁴ „The wider public, as well as Government, have a right to know how effective the education service is. It is quite possible for accountability to be both rigorous and supportive of schools“ (NUT 2005, S. 4).

Ein zentrales Thema in allen vier Ländern ist auch das der Accountability, also der Verantwortung für den Bildungserfolg oder -misserfolg. Auch dieses Thema ist im deutschen Sprachraum kaum präsent. Wenn, dann wird es moralisch besetzt, nicht jedoch auf konkrete Strategien bezogen. Amerikanische Untersuchungen zeigen, dass in den Bundesstaaten starke Accountability-Strategien (policies) dort implementiert werden, wo die lokale Kontrolle schwach ist. Starke Top-down-Strategien gefährden aber den Grad der Flexibilität und der Kontrolle der Aktivitäten vor Ort, also in den Schulen und Distrikten. Maßnahmen zur Accountability haben einen Einfluss auf das Leistungsverhalten, allerdings ist der Effekt größer, wenn die lokalen Zuständigkeiten stärker sind. Das spricht zum Beispiel gegen die mit dem NCLB-Act verstärkte Zentralisierung der Bildungsfinanzierung, die die örtliche Budgetierung und Ausgabenkontrolle geschwächt hat (Loeb/Strunk 2005).

Was Standards letztlich sind, kann nur an den Produkten abgelesen werden. Curriculare Standards wie in England sind nicht nur Lernprogramme, die sich auf Schulfächer und Schlüsselstufen beziehen, sondern zugleich Leistungserwartungen, die Kompetenzstufen voraussetzen. Hier kann beschrieben werden, auf welchem Niveau die Schülerinnen und Schüler zum Zeitpunkt der Beurteilung arbeiten. Alle Fächer (subjects) sind für alle fünf Stufen nach dem gleichen Muster durchkonstruiert worden. Kein anderes europäisches System verfügt bislang über ein vergleichbares Produkt.

Auffällig ist, dass alle vier Länder Bestrebungen zeigen oder gezeigt haben, die Curricula stärker zu nationalisieren. Ohne ein vergleichbares Angebot lassen sich die Ergebnisse schwerlich testen, oder die Tests setzen eigene curriculare Normen, bei denen nicht sicher ist, ob der Unterricht sie vorausgesetzt hat. Das Problem ist, wie präzise, wie umfangreich und wie verbindlich die curricularen Standards gefasst sein sollen. Amerikanische Lehrpläne unterscheiden sich in dieser Hinsicht sehr stark von schwedischen Lernzielen.

Auch da, wo stark auf schuleigene Curricula gesetzt wird wie in Schweden oder in den Niederlanden, sind „zentralistische“ Tendenzen unübersehbar. Ziele oder Standards werden vorgegeben, allerdings nicht in Form von additiven

¹⁴⁴ National Union of Teachers. Es handelt sich um die Antwort der Gewerkschaft auf das Weißbuch *High Standards, Better Schools for All*.

Beispielsammlungen wie zurzeit noch in deutschen Lehrplänen. In England lässt sich eine generelle Tendenz zur Zentralisierung feststellen, einhergehend mit Teilprivatisierungen, Schulautonomie, lokalem Wettbewerb und gleichzeitig starker Steuerung durch den Staat. „Public schools find themselves located within a matrix of accountability relationships; upward to government and outward to communities and private organizations involved directly in their governance“ (Fitz 2003, S. 240).

Aus dem Vergleich ergeben sich folgende Befunde:

- Alle Länder verfahren nach dem Prinzip der Output-Steuerung, wenngleich in unterschiedlicher Weise. Alle verfügen auch in der einen oder anderen Form über Rückmeldesysteme.
- Content und Performance Standards sind quer zu den verschiedenen Systemen international üblich. Das Gleiche gilt für Formen der Überprüfung, ob und wie weit die Ziele erreicht wurden, sowie für die Rechenschaftslegung.
- Die Implementation sollte systemverträglich erfolgen, verlangt aber eigenen Aufwand, der erheblich ist. Top-down-Steuerungen sind ohne weitere Maßnahmen wenig wirksam. Die lokale Ebene ist für die Implementation ausschlaggebend.
- Wesentlich für das Gelingen der Reform ist eine verlässliche, kohärente und zeitlich unbefristete Politikstrategie. Zwischen Politik und Profession sollte es keine unnötigen Konflikte geben.
- Auch in liberalen Systemen nimmt der Grad an Verbindlichkeit zu. Gleichzeitig steigen die Anforderungen, transparente Leistungsnachweise vorzulegen. In den europäischen Ländern gibt es Inspektorate und/oder externe Evaluationen. Ebenfalls üblich sind nationale Tests und externe Prüfungen. Zu den nationalen Tests kommen weitere, die meistens freiwillig sind.
- Die Unterscheidung von Kompetenzniveaus bei den Performance Standards ist unterschiedlich weit entwickelt, wird aber überall vorangetrieben. Bessere Produkte unterstützen die Qualitätssicherung.
- School Choice, Voucher-Finanzierung und Accountability sind zentrale Themen der gegenwärtigen Bildungspolitik, ausgenommen in Zentraleuropa.
- In allen Ländern gibt es einen engen Zusammenhang zwischen Bildungsforschung und Systementwicklung. Allerdings gibt es eine Tendenz, Forschungsergebnisse selektiv zu verwenden.
- Steuerungsimperative orientieren sich an betriebswirtschaftlichen oder Konzepten des New Public Managements. Der Einfluss verschiedener Wissenschaften ist zwischen Politik und Profession ungleich verteilt, die Bildungs-

ökonomie gewinnt in Skandinavien und den angelsächsischen Ländern an Bedeutung.

- Vor allem in den Vereinigten Staaten stehen sich „progressive“ und „konservative“ Ansätze unvermittelt gegenüber. Die Konfrontation ist angesichts der Datenlage kein Vorteil.

4 Verständnis, Strategien, Akteure, Ebenen und Kontexte der Implementation von Bildungsstandards

Die deutsche Expertise von 2003 weist ausdrücklich auf das Problem der *Implementation* von Bildungsstandards hin: „Die Entscheidung, Bildungsstandards zu entwickeln und im deutschen Schulsystem zur Wirkung zu bringen, setzt ein Konzept zur Implementation voraus“ (Klieme/Avenarius/Blum et al. 2003, S. 111). Die Grundzüge der Implementation werden darin angedeutet. Sie zielen darauf ab,

- das Verständnis und die Akzeptanz von Bildungsstandards zu sichern,
- in die Arbeit mit Bildungsstandards einzuführen und
- mit den Möglichkeiten der professionellen Nutzung und Auswertung von Bildungsstandards vertraut zu werden (ebd.).

Gemäß dieser drei Aufgaben müssen für die Schulen *Information, Anleitung, Fortbildung* und *Beratung* angeboten werden. Insbesondere müssen die Schulen Hilfen erhalten, „wenn die Auswertung von Tests Probleme und Handlungsbedarf erkennen lassen“ (ebd.). Generell müssen frühzeitig und im nötigen Umfang „vielfältige Unterstützungen“ bereitstehen, „damit die Standards von den Schulen angenommen und wie geplant umgesetzt werden können“ (ebd.). Einrichtungen der Lehrerbildung, die deutschen Landesinstitute sowie die Schulaufsicht werden im Prozess der Implementation zu „wichtigen Unterstützungssystemen“ für die Einführung von Bildungsstandards und deren Umsetzung im schulischen Alltag (ebd.).

Dabei werden die nötigen Rahmenbedingungen angesprochen, also Ressourcen, Qualifikationen oder Vereinbarungen, aber auch die bildungspolitischen Begründungen für die Einführung von Bildungsstandards (ebd., S. 111f.).

„Für die Akzeptanz entscheidend werden die offiziellen Botschaften sein, die nicht nur den nachdrücklichen politischen Willen und die Zwecksetzung erkennen lassen, sondern den Schulen und Lehrkräften signalisieren, wie sie auf lange Sicht in ihrer täglichen Arbeit (professionell und persönlich) von Bildungsstandards profitieren können. Die Interpretation des gesamten Vorhabens und aller angebotenen Unterstützungsleistungen hängt davon ab, ob die Schulen und Lehrkräfte die angestrebte Veränderung der Steuerung von einer Input- zu einer Outputorientierung nachvollziehen können“ (ebd., S. 112).

Offizielle Botschaften sind die eine Seite, Ressourcen und konkrete Instrumente zur Umsetzung die andere. Kompetenzmodelle und Tests sowie Rückmeldesysteme und kollegiale Kooperationen sind neue Instrumente. Darauf bezogen sollen Schulen Informationen, Handreichungen und Fortbildungen erhalten, die zu verstehen sind als „Wegbeschreibungen, als Orientierungshilfen, als Werkzeuge und Handlungsgerüste“. Unterstützungssysteme und Instrumente dieser Art sollen dabei nicht als „neuerliche Vorgaben oder verordnete Aufträge“ erscheinen, die lediglich „abgearbeitet“ werden müssen (ebd.).

„Funktion und Nutzen der Standards können die Schulen dann nachvollziehen, wenn sie den Rahmen und die langfristige Absicht kennen, nämlich den Schulen mehr Gestaltungsspielräume und Verantwortung zu übertragen“ (ebd.).

Damit ist die Ebene der Akteure und ihres direkten Umfeldes als zentrale Größe der Implementation erfasst. Die Akteure – Lehrkräfte, Schülerinnen und Schüler, Eltern – müssen die neuen Möglichkeiten der Outputsteuerung erkennen und daraus Vorteile für ihr Verhalten ziehen können. Der Aufbau von Akzeptanz muss unterstützt werden; je distanzierter das Szenario der Bildungsstandards vor Ort wahrgenommen wird, desto geringer ist die Erfolgswahrscheinlichkeit. Die lokalen Theorien sind dann stärker als die behördlichen Weisungen, die tatsächlich nur „abgearbeitet“ werden, ohne in der inneren Organisation der Schulen Wirkungen zu zeigen.

Im Folgenden wird im Anschluss an Erkenntnisse aus der Forschung und an internationale Erfahrungen (vgl. Kapitel 3 dieser Expertise) dargelegt,

- welche inhaltlichen, technischen und nicht-technischen Anforderungen sich bei der Einführung von Bildungsstandards stellen bzw. von welchem *Grundverständnis* ihrer Umsetzung die vorliegende Expertise ausgeht;
- dass es bei der evaluativen Anwendung von Bildungsstandards nicht um eine formale Form der Rechenschaftslegung und damit um einen Verwaltungsakt, sondern um die *Einführung eines Referenzsystems für professionelles schulisches Handeln und die Vergewisserung seiner Qualität* geht;
- dass ohne *breitenwirksame Strategien und Verfahren der Schulentwicklung*, die den in Gang zu setzenden Prozessen angemessen sind und die über die evaluative Funktion einer Einführung von Standards hinausgehen, mit keiner Nachhaltigkeit in Bezug auf Qualitätsverbesserungen im Bildungssystem zu rechnen ist;

- dass die Entwicklung und Einführung von Bildungsstandards *keine Monostrategie* der Weiterentwicklung von Schule und Unterricht, und die Ertrags- oder Outputsteuerung künftig nicht die einzige Säule der Qualitätssicherung des Bildungssystems darstellen kann;
- dass nicht nur die Lehrpersonen und deren schulisches Umfeld, sondern ebenso die *Unterstützungssysteme* im Bildungswesen in die sukzessive Umsetzung von Bildungsstandards einzubeziehen sind;
- dass die Implementation von Bildungsstandards nicht ohne „Tools“, d.h. praxistaugliche *Verfahren und Werkzeuge* auskommt, die auf den verschiedenen Ebenen des Bildungssystems, vorzugsweise in jenen Segmenten, deren Funktion die Unterstützung des professionellen Lernens von Lehrpersonen darstellt, angesiedelt sind.

Abbildung 3 fasst die für eine nachhaltige Implementation wesentlichen und im Folgenden weiter zu thematisierenden Aspekte zusammen.

Abbildung 3: Aspekte und Handlungsfelder einer Implementation von Bildungsstandards

Begriff und theoretisches Verständnis einer Implementation von Bildungsstandards (Kap. 4.1)	
<p>Leitideen und Strategien (Kap. 4.2)</p> <ul style="list-style-type: none"> • Anzustrebende Akzeptanz als Mittel zur Zielfokussierung, Evaluation und Qualitätsentwicklung auf System- und Schulebene; • Nutzung in der unterrichtlichen Arbeit; • Verbesserung der Leistungs- und Lernqualität. 	<p>Ebenen und Akteure (Kap. 4.3)</p> <ul style="list-style-type: none"> • Nationale Ebene; • Bundesländer-/Kantoneebene, regionale Netzwerke; • Schulebene: Gemeinde- und Einzelschulen; • Ebene der Lehrkräfte, Fachgruppen, Schülerinnen und Schüler und Eltern.
<p>Kontext- und Unterstützungssysteme (Kap. 4.4)</p> <ul style="list-style-type: none"> • Stützende politische Langfriststrategie; • Lehrerbildungs- und Fortbildungseinrichtungen; • Landesinstitute, pädagogische Abteilungen, zentrale Einrichtungen zur Qualitätsentwicklung und Beratung; 	<p>Werkzeuge und Verfahren (Kap. 5)</p> <ul style="list-style-type: none"> • Ausgearbeitete Bildungsstandards im Kontext von Lehrplänen und Curricula; • Valide Tests und Monitorings im Kontext eines mehrperspektivischen Assessment-Systems; • Externe Evaluationen und Rückmeldesysteme;

<ul style="list-style-type: none"> • Schulaufsicht, Schulinspektion; • Lehrerverbände und -gewerkschaften; • Lehrmittel und Schulbuchproduzenten; • Bildungsforschung. 	<ul style="list-style-type: none"> • Verfahren der internen Schulevaluation; • Schulleitung und Schulleitbild; • Referenzrahmen und Werkzeuge zur Schul- und Unterrichtsentwicklung.
--	---

4.1 Begriff und theoretisches Verständnis einer Implementation von Bildungsstandards

Unter dem Begriff der „Implementation von Bildungsstandards“ verstehen wir die Einführung und handlungswirksame Umsetzung eines neuen Referenzsystems für die fachlichen Kernziele schulischen Handelns in die pädagogische Alltagsarbeit an den Schulen und in der Lehrerbildung. Ein Kernziel ist dabei, dass die Aneignung des Referenzrahmens mit Tiefenwirkung erfolgt und das System als Ganzes erfasst („scaling up“, vgl. Coburn 2003).

Bei der Implementation bildungspolitischer Vorgaben, die sich auf eine systemweite Verbesserung der Schulqualität richten, handelt es sich um einen in seiner Komplexität meist verkannten, unter mannigfaltigen Bedingungen und Wechselwirkungen stehenden sozialen *Anpassungs-, Entwicklungs- und Lernprozess*, welcher sowohl normative, kulturelle, inhaltliche, methodisch-messtheoretische, organisational-strukturelle als auch personelle und sozial-kommunikative Komponenten und Bedingungen umfasst. Die erfolgreiche Realisierung komplexer schulischer Neuerungen verlangt deshalb über ihre Qualität hinaus nach einer kohärenten und langfristig angelegten, institutionellen *Qualitätspolitik* auf allen Ebenen des Bildungssystems.

Grundsätzlich lassen sich zwei allgemeine Governance- bzw. Regelungs- oder Steuerungsstrategien bei der Implementation schulischer Neuerungen unterscheiden: Eine bürokratisch-hierarchische, *autoritativ-interventionistische* („*top-down*“) *Strategie* und eine an einer „Rekontextualisierung“ externer Ziele und Strukturvorgaben (Fend 1987, 2006) sich orientierende *horizontale* („*bottom-up*“) *Strategie der partizipationsbezogenen Kontextsteuerung* (vgl. Kap. 3.1.5). Während die erste Strategie meist nach Maßgabe eines *Fidelity-Ansatzes* (vgl. Ruiz-Primo 2006) sowie der Annahme einer unilateralen Wirkungsrichtung davon ausgeht, dass ein als fertiges Programm ausgearbeitetes Paket von institutionellen Regelungen, Aufgabenbestimmungen und Praxisänderungen von den angezielten Akteuren übernommen und in deren operatives Handlungsrepertoire als Standardpraktik integriert wird (Knapp (1995, 2002) spricht von einem „Outside-

in“-Ansatz), geht die zweite Strategie davon aus, dass es sich bei der Aneignung einer Innovation um einen in wechselseitigem Austausch zwischen den Systemebenen und nach Maßgabe vielfältiger Voraussetzungen erfolgenden individuellen und kollektiven Anpassungsprozess bzw. um einen *mehrstufigen, adaptiven Lern- und Entwicklungsprozess* handelt (Knapp ebd.: bezeichnet dieses Vorgehen als „Inside-out“-Ansatz).

Aus vielfältigen Erfahrungen mit komplexen Implementationen im Bildungsbereich (und darüber hinaus in den Bereichen des Politik-, Verwaltungs- und Wirtschaftshandelns) und ihren Wirkungsverläufen ist bekannt (vgl. v.a. Kap. 3.1.4), dass die Basisakteure und Verantwortungsträger eines Systems (z.B. Lehrkräfte) politisch intendierte Neuerungen kaum je als fertige Programme einfach „übernehmen“ und auch ihre eingeschliffenen Praktiken kaum ändern. Zumindest nicht nachhaltig und bezogen auf die Tiefenstruktur einer Innovation. Neuerungen werden oft nur selektiv, partiell, oberflächlich, und nicht selten verzerrt übernommen und bleiben ohne wirkliche „depth“, „sustainability“, „spread“ und „shift in ownership“, wie dies Coburn (2003) hierarchisch-verordnenden Steuerungsansätzen gegenüber kritisch festgehalten hat. Zu einer ähnlichen Einschätzung kommen ebenfalls viele Autoren, die sich in der Perspektive einer interdisziplinären sozialwissenschaftlichen Innovationsforschung mit Prozessen des Wissenstransfers und der Wissenstranformation beschäftigt haben (vgl. Berman/McLaughlin 1977; Huberman/Miles 1984; Fullan 1994, 2001; Reinmann-Rothmeier/Mandl 1998; Hall/Hord 2001; Altrichter/Wiesinger 2004; Gräsel/Parchmann, 2004).

Rein klassische „Transfer“- „Verordnungs“- „Übernahme“- bzw. „Governance“-Modelle¹⁴⁵ im Sinne einer bürokratischen „Top-down“-Regulierung komplexer Reformvorhaben, aber auch rein evolutionäre „Selbstregulations“- und selbstreferenzielle „Basisentwicklungs“-Ansätze („Bottom-up“-Modelle), werden zugunsten von *Mischformen der Regulierung* (Sabatier 1986) sowohl in der Theorie als auch in der Praxis kaum mehr vertreten. Jenseits vereinfachender Reformprozessvorstellungen geht es darum, schulische Neuerungen unter Beteiligung wechselseitig voneinander abhängiger Akteure auf allen Systemebenen in einem ko-konstruktiven, mehrstufigen Entwicklungsprozess zu rekontextua-

¹⁴⁵ „Governance“ verstanden in seiner traditionellen (engen) Form des regierungsmäßigen, unilateralen Steuerungshandelns (im Gegensatz zu einem erweiterten, modernen Governancebegriff, wonach problemlöseffizientes Steuerungshandeln als Wechselwirkung von politisch-administrativen Vorgaben und partizipationsbezogenem Aushandeln („Sense-Making“) an der Basis zu verstehen ist).

lisieren und daraus resultierende neue Kompetenzen, Routinen, Einstellungen und Identitäten aufzubauen.

Zu beachten ist, dass zur Frage der Effektivität von Implementationen – etwa unter welchen Bedingungen sehr komplexe Reformvorhaben ihre Adressaten erreichen und welches die Erfolgsdeterminanten ihrer Wirksamkeit sind – erstaunlicherweise nach wie vor nur wenig Forschungswissen im engeren Sinne existiert (vgl. für einen Überblick über einige Probleme z.B. Fullan/Pomfret 1977; Berman 1980; Ruiz-Primo 2006).

Inwieweit die von einem (bildungspolitischen) Auftraggeber intendierte Realisierung einer komplexen Neuerung, wie sie die Einführung von Bildungsstandards als einem neuen Referenzrahmen für die Kommunikation überprüfbarer Basisziele im Bildungsbereich darstellt, gelingt, dürfte (bei aller Begrenztheit des Wissens) insbesondere von folgenden personen-, gegenstands- und kontextbezogenen Faktoren abhängen:

- der subjektiven Beteiligung der Akteure (Commitment, Engagement, Partizipation an Entscheidungsfindungen);
- ihrem Wissen, ihren Überzeugungen, dem wahrgenommenen Problemdruck sowie der Wahrnehmung und Interpretation des Gegenstandes (strategische Ausrichtung, Qualität und Praktikabilität der einzelnen Reformkomponenten);
- der Dauer ihrer Beschäftigung mit den neuen Ideen sowie den (positiven) Anreizen, die damit verbunden sind;
- ihren Vorerfahrungen mit früheren und analogen Problemlösungen und den dabei erworbenen Kompetenzen;
- ihren individuellen und kollektiven Ressourcen (Unterstützungssysteme wie professionelle Beratung, Coaching) zur Aufgabenerfüllung;
- einem koordinierten und verlässlichen Handlungsplan, der alle Ebenen und wichtigen Teilschritte umfasst und auf Kontinuität und Stabilität angelegt ist;
- von mannigfaltigen kontextuellen Konstellationen wie aktiver Unterstützung durch das lokale Umfeld, der existierenden Schulkultur, einem kompetenten innerschulischen Management sowie schulübergreifenden Netzwerken.

Auch bei der Einführung und testbasierten Überprüfung von Bildungsstandards ist davon auszugehen, dass die neuen Vorgaben ihre Adressaten nicht von selbst erreichen werden, sondern deren Verbreitung und Kommunikation von mannigfaltigen Anreizen und „Change-Bedingungen“ abhängt und ebenso vielfältige Prozesse der subjektiven Interpretation und Assimilation – auch ge-

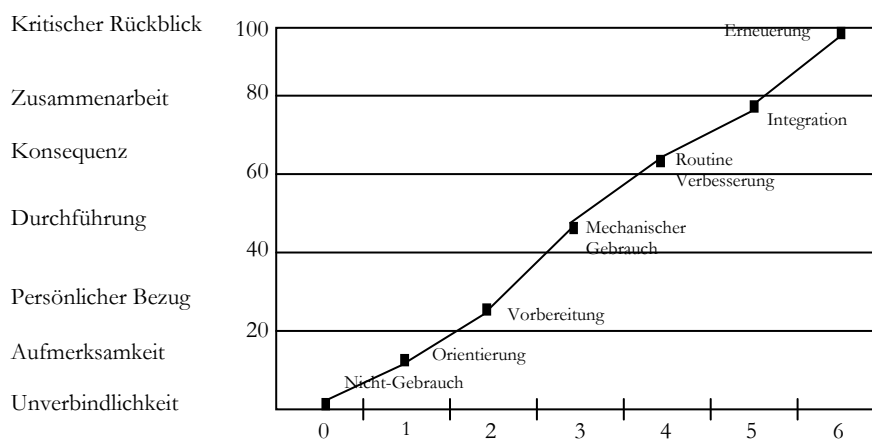
gen individuelle und kollektive Widerstände – auslöst. Diese Anreize (wozu nicht einfach nur Geld, sondern z.B. eine erhöhte Schulautonomie gehört) und Prozesse lassen sich positiv bewerten, wenn sie – bei grundsätzlicher Akzeptanz der Innovation – zur Verinnerlichung und Verwirklichung der Impulse führen, oder aber negativ, wenn die Umsetzung der Vorgaben durch wenig gewogene Adressaten mehr oder weniger subtil abgewehrt wird bzw. selektiv und verzerrt – ohne Commitment und Engagement – erfolgt. Da die Innovation nicht nur die Einzelschulen und die Lehrkräfte, sondern das Bildungswesen als kollektiven Akteur tangiert, ist für eine adäquate Beschreibung des Gesamtzusammenhangs eine systemische Perspektive unverzichtbar.

Aus der neueren akteurtheoretischen *Innovations- und Transferforschung* (Rogers 2003; vgl. Nickolaus/Graesel 2006) sowie aus jüngeren Arbeiten zu einem erweiterten Verständnis von Governance in Bildungssystemen (Educational Governance; vgl. z.B. Fend 2006; Altrichter/Brüsemeister/Wissinger 2007) wird zudem deutlich, dass es sich bei großen und komplexen Innovationsvorhaben kaum je um einphasige Prozesse handelt, die das System in einem einzigen Schritt in der ganzen Breite erfassen. Vielmehr entfalten Reformen, die die normativen und operativen Tiefenstrukturen, Einstellungen und Identitäten eines kollektiven Akteurs betreffen, ihre Wirkung in mehreren, zeitlich sich erstreckenden Phasen. In einer vom BMBF in Auftrag gegebenen Positionsbestimmung zur Transfer- und Innovationsforschung beschreibt Holtappels (2005) zwei Stadienmodelle des Verlaufs von Innovationen:

- Nach Giaquinta (1973) erfolgt eine Innovation häufig in drei sich überlappenden Phasen der (1) *Initiation*, der (2) *Implementation* und der (3) *Institutionalisierung*. „Nach der Phase der Initiation, in der Information, Motivation, Problemanalyse und Überzeugung geschieht, folgt die Phase der Implementation, in der die Innovation umgesetzt, erprobt und angewendet und dabei in das vorhandene System integriert wird. In der Phase der Institutionalisierung (oder der Inkorporation) wird die Innovation zur Alltagspraxis und zur Routine. Alle drei Phasen können innerhalb eines Schulsystems gleichzeitig ablaufen, weil die Innovation noch nicht sofort alle erreicht oder Schulen unterschiedlich entwickelt sind; sogar innerhalb einer Einzelschule wären diese Phasen eventuell gleichzeitig beobachtbar“ (Holtappels 2005, S. 4).
- Unter Bezugnahme auf Hall (1979), Loucks/Hall (1979) sowie Huberman/Miles (1984) umreißt Holtappels sodann ein „Aneignungs-Nutzungs-Modell“ der schrittweisen Aneignung neuer Ideen durch Lehrpersonen (vgl.

Abb. 4): „In der Phase der Unverbindlichkeit dominiert Nicht-Gebrauch, erlangt die Neuerung aber Aufmerksamkeit kommt es zu Aktivitäten der Orientierung, erst mit persönlichem Bezug wird ein erster Durchbruch erzielt und wird Vorbereitung erreicht; die Durchführung zeigt jedoch zunächst eher mechanischen Gebrauch, bevor es zu Routine und Verbesserungen kommt. Integration in institutionelle Praxis erfordert Zusammenarbeit, stabile Erneuerung benötigt kritischen Rückblick und Evaluation. Solche Phasen sind bei Innovationen in der Implementationsstrategie nicht nur zeitlich zu veranschlagen. Individuelle Aneignungs- und Veränderungsprozesse sind allerdings nicht bruchlos auf die Selbsterneuerungsfähigkeit eines Systems zu übertragen, weil die kollektive Veränderung eine andere Dynamik besitzt und weitere Faktoren Einfluss auf Organisationswandel nehmen“ (Holtappels 2005, IFS, S. 5).

Abbildung 4: Stadien und Niveaus der Aneignung und Einführung von Innovationen bei Lehrkräften (nach Loucks/Hall 1979; Hall 1979; Huberman/Miles 1984)



Quelle: Holtappels 2005, IFS, S. 5

Die beiden Modelle machen deutlich, dass es sich bei einer Innovation um einen mehrstufigen individuellen und kollektiven Aneignungsprozess handelt, wobei der Weg des individuellen und organisationalen Lernens – von der erstmaligen Kenntnisnahme über Stufen der Verinnerlichung bis hin zur dauerhaften Integration ins personale und kollektive Handlungsrepertoire – sehr lang

sein kann. Meist dauert es eine Zeit, bis die Akzeptanz einer Reform die Gruppe der eigentlichen Innovationsträger, die an ihrer Entwicklung oder an einem initialen Modellversuch teilgenommen haben (die „early adopters“), überschreitet und eine deutliche Breitenwirkung einsetzt. Bekannt ist auch, dass der in der Literatur als s-förmige Kurve beschriebene Verlauf der Verbreitung einer Innovation (Rogers 2003) später wiederum abflacht und mit einer nicht unbedeutenden Menge von „late adopters“ zu rechnen ist. Das Problem lässt sich auch dadurch verdeutlichen, dass es sich bei einem Projekt wie der Einführung von Bildungsstandards nicht bloß um wenige Akteure handelt, deren Einstellungen, Wissen und Praktiken verändert werden sollen, sondern dass die gesamte Lehrerschaft angesprochen ist, d.h. allein in Deutschland gegen 800.000 Akteure an der Basis des Bildungswesens. Das heißt, es stellt sich, wie bei allen bedeutenden Innovationen, ein *Scale-up-Problem*, zu dessen Lösung es keine einfachen und empirisch erhärteten und schon gar keine linearen Strategien gibt.

Kein Zweifel besteht zudem darüber, dass es sich bei der Einführung von Bildungsstandards und darauf bezogenen Tests nicht um eine von den pädagogischen Akteuren an der Basis gewünschte Initiative, sondern um ein durch eine unter dem Druck von TIMSS und PISA stehende Bildungspolitik initiiertes Projekt handelt; um ein Projekt, bei dem es von der inhaltlichen Substanz her betrachtet – Übergang von einer stoff- zu einer kompetenzorientierten Sichtweise schulischer Bildungsziele, Wirksamkeitsüberprüfung mittels Tests und Ergebnisrückmeldungen – um nichts Geringeres als einen Paradigmenwechsel im Bereich der Qualitätspolitik und Bildungssteuerung geht. Es wäre naiv, davon auszugehen, dass sich eine Innovation solchen Ausmaßes ohne weitreichende Maßnahmen, die über die Konstruktion von Standards und die Anwendung von Tests hinausgehen, erfolgreich wird umsetzen lassen.

Weder kann somit das Reformprojekt, in dem neue Formen der Definition von Bildungszielen, Tests und Evaluation von Schulen wichtige Pfeiler darstellen, von schulstrukturellen, curricularen und unterrichtlichen Zusammenhängen abstrahieren, noch können die Probleme und Prozesse bei der Aneignung der neuen Ideen und Ziele durch die verantwortlichen Akteure und ihr Umfeld außer Acht gelassen werden. Bildungsstandards, die für sich genommen lediglich eine neue Form der Inhalts- und Zielformulierung darstellen, sowie Tests und daran geknüpfte Ergebnisrückmeldungen können erst dann einen Beitrag zur Qualitätsentwicklung leisten, wenn ihre zielklare politische und administrative Kommunikation mit pädagogischen Initiativen zur Schulentwicklung und zum professionellen Lernen von Lehrkräften verbunden wird. Dafür – und nicht ausschließlich für die Entwicklung der Standards und die Produktion der darauf

bezogenen Informationsmaterialien – müssen bedeutende Ressourcen bereitgestellt werden.

4.2 Leitideen und Strategien einer neuen Qualitätspolitik

Die Einführung und evaluative Anwendung von Bildungsstandards stellt weder die Dissemination eines fertigen Programms noch einen bürokratischen Verwaltungsakt dar. Jenseits der politischen Verordnung und schulrechtlichen „Akkreditierung“ von ausgearbeiteten Standard-Dokumenten bezieht sich deren Umsetzung auf einen komplexen Lernvorgang, der sowohl einen Wandel von Überzeugungen als auch individuelles und institutionelles, organisationales Lernen einschließt. Dies entsprechend einem Verständnis des Bildungssystems als eines komplexen humanen Regelwerks von Handlungsrouninen, Ressourcen und Technologien (vgl. Fend 2006), dessen Funktionieren auf einer Aneignung seiner Werte und Normen, konsensualer Veränderung und permanenter Lernbereitschaft seiner voneinander interdependenten Akteure beruht.

Eine solche Auffassung setzt Strategien sowohl auf der Systemebene wie auch in Bezug auf Prozesse des produktiven professionellen Lernens an der Basis voraus. Auf der Systemebene stellt sich die Frage nach einer geeigneten Qualitätspolitik, in der Balancen gesucht werden zwischen Top-down-Strategien der Kommunikation entwickelter Standards und Tests und der Partizipation der Basis an deren praxisbezogener Prüfung, Validierung und inhaltlichen Weiterentwicklung und Ausgestaltung. Auf der Ebene der Personal- und Schulentwicklung sind einerseits die institutionellen und materiellen Bedingungen zu schaffen, andererseits ist nach mikropolitischen Strategien zu suchen, welche die eingeleiteten Prozesse wirksam unterstützen.

4.2.1 Anforderungen und Leitideen

Nimmt man das dargelegte Verständnis von Innovation und Implementation zur Grundlage, so lassen sich für die Einführung von Bildungsstandards folgende Leitvorstellungen und Anforderungen formulieren.

(1) Inhaltliche Qualität der Bildungsstandards und Tests. Das Unternehmen „Bildungsstandards“ steht und fällt mit der inhaltlichen Validität und methodischen Ausarbeitung sowohl der Standards als auch der darauf bezogenen Aufgaben und Tests.

Nur Bildungsstandards, die sich auf begründbare und ausgehandelte Kernbereiche und Kernkompetenzen eines Fachs beziehen, und nur Tests, die tatsächlich

jene Leistungen erfassen, welche mit den Standards gemeint sind, werden langfristig Akzeptanz erlangen und – falls weitere notwendige Bedingungen zu ihrer erfolgreichen Implementation erfüllt sind – positive Wirkungen entfalten. Aus der Sicht der Adressaten sollen gute Standards in erster Linie klar verständlich, knapp und nachvollziehbar formuliert sein, zudem fokussiert auf Grundprinzipien und Grundkompetenzen, d.h. auf den Kernbereich eines Fachs. Idealerweise sollen Standards dabei nicht einfach eine allgemeine „Messlatte“ anlegen, sondern zwischen Dimensionen und Stufen des Wissens und Könnens differenzieren und so die Unterscheidung von Abstufungen sowie Profilbildungen ermöglichen (je nach gewählter Konzeption zwischen Mindest- und Maximalniveaus). Weiter sollen diese Kompetenzstufungen und dimensional Profile an guten Beispielen ablesbar und illustrierbar sein.

Dass solche Qualität real existierender Standards heute in keinem der Auftragsländer bereits gegeben ist, sollte man sich ebenso vor Augen halten wie die Feststellung, dass das Erreichen dieser ambitionierten Ziele aus wissenschaftlicher Sicht keineswegs selbstverständlich oder gesichert ist. Einen messtheoretisch und pädagogisch wertvollen Test zu entwickeln ist ein hoch anspruchsvoller wissenschaftlicher Prozess, der neben substanziellen fachdidaktischen und kognitionspsychologischen auch methodologische Kompetenzen erfordert. Keineswegs trivial ist es beispielsweise, mit Tests tiefes Verstehen und Problemlösen zu erfassen. Es ist jedoch zu fordern, dass Tests gerade auch dies nachweisbar gut zu bewerkstelligen vermögen. Als unsicher gelten muss bis heute, ob und inwieweit es (in unterschiedlichen Fächern) gelingen wird, dem viel zitierten Anspruch der Kompetenzmodellierung – im Sinne der Bildung kognitionspsychologisch und fachdidaktisch fundierter, dimensionaler und niveaugestufte Kompetenzmodelle – in einem strengen Sinne gerecht zu werden. Auch in Bezug auf Tests, soweit sie heute überhaupt vorliegen, muss erwähnt werden, dass auch deren Qualität nicht unumstritten ist (vgl. etwa Bartnitzky 2006). Ehmke/Leiss/Blum/Prenzel (2006) beschreiben den „state of the art“ der Testentwicklung in der als Referenzfach dienenden Mathematik und geben einen Ausblick auf noch ausstehende Arbeitsschritte und die damit verbundenen Anforderungen.

Damit hochwertige Standards und Tests entwickelt werden können, braucht es die entsprechende Infrastruktur – eine weitere zentrale Anforderung. Mit dem bundesweit tätigen IQB (Institut zur Qualitätsentwicklung im Bildungswesen – Wissenschaftliche Einrichtung der Länder an der Humboldt-Universität zu Berlin) wurde in Deutschland ein wesentlicher Schritt in diese Richtung getan. In den anderen Ländern besteht in dieser Hinsicht noch deutlicher Handlungsbedarf, da entsprechende Einrichtungen fehlen.

Der wichtigste Grund, warum Standards und Tests selbst höchsten Anforderungen genügen müssen, ist deren absehbarer und nicht zu unterschätzender Einfluss auf das Unterrichtsgeschehen. Da Standards (auch dort, wo sie in bisherigen Fassungen als umgeschriebene Lehrpläne daherkommen) in den Auftragsländern definitiv mit Tests gekoppelt werden sollen, ist von einem bedeutenden Einfluss auf den *Unterricht* auszugehen; dies durchaus auch im Sinne eines „teaching to the test“ – sowohl im Guten (als durchaus erwünschte „intelligente“ Ausrichtung des Unterrichts auf überprüfbare Kernziele) wie auch im Schlechten (als unerwünschte, auf kurzfristige Wirkung bedachte „reduktionistische“ Praktik). Einen nicht minder wichtigen, jedoch eher indirekten Einfluss auf den Unterricht nehmen Standards sodann über die Lehrmittel. Umgekehrt formuliert: Welche Lehrmittel verwendet oder zugelassen werden, wird künftig zu einem guten Teil davon abhängen, inwiefern diese zur Erreichung offizieller Standardkataloge beitragen.

(2) Öffentliche Sichtbarkeit einer zusammenhängenden Strategie und Verlässlichkeit der politischen Unterstützung bei der Umsetzung. *Mit der erfolgreichen Einführung und Verbreitung nationaler Bildungsstandards kann nur gerechnet werden, wenn die politischen und pädagogischen Ziele in einer Langzeitstrategie verankert sind und die politische Unterstützung nicht nur auf der rhetorischen Ebene, sondern auch auf der Ebene der Bereitstellung von Ressourcen verlässlich und koordiniert gesichert ist.*

Ohne bindenden Konsens über eine Reihe von Grundfragen und ohne nachhaltige politische Unterstützung auf allen Ebenen der Bildungsadministration ist hinsichtlich eines Projektes, das auf die Kernstruktur des Bildungs- und Unterrichtssystems abzielt, nur schwerlich mit einem Erfolg zu rechnen. Zu diesen Fragen, die mit Vorteil für das Gesamtsystem verbindlich und möglichst einheitlich geregelt werden sollten, gehört, ob Mindest- oder Regelstandards generell bindend sein sollen, wer und mit welcher (wissenschaftlichen) Infrastruktur Standards und Tests entwickelt, in welchem Verhältnis Bildungsstandards und Lehrpläne stehen, wie häufig welche nationalen oder subnationalen Tests eingesetzt werden und welchen Zwecken (Förderung, Selektion etc.) sie dienen. Aber auch das Verhältnis von Tests zu anderen Formen der (Selbst-)Evaluation muss geklärt werden, sodann die zentrale Frage, wie mit den Testergebnissen umzugehen ist und wer zu den Daten mit welchen Nutzungszwecken Zugang hat. Schließlich sollte ein verbindlicher Konsens darüber getroffen werden, welche Folgen die Ergebnisse aus Leistungstests für die Allokation von Ressourcen für die Qualitätssicherung und -weiterentwicklung in den Schulen haben.

Die Auftragsländer sind mit diesen Entscheiden zurzeit unterschiedlich weit, und es zeichnen sich auch klare nationale Unterschiede ab. Nach Auffassung dieser Expertise sind solche nationalen Differenzen wichtig, widerspiegeln sie doch unterschiedliche – föderalistisch oder zentralistisch geprägte – bildungspolitische Kulturen. Gemeinsam dürfte den Strategien aller Länder jedoch die Einsicht sein (vgl. Kapitel 6), dass sich (in Übereinstimmung mit dem wissenschaftlichen Kenntnisstand) die Probleme der Implementation von Bildungsstandards durch Strategien des bloßen Verwaltungshandelns und der behördlichen Dienstwegkommunikation nicht werden lösen lassen.

(3) Sicherung von Akzeptanz in der Profession. *Wie die Expertise von Klieme et al. (2003, S. 112f.) sehen auch wir im Erzeugen von Akzeptanz eine der wichtigsten Herausforderungen. Akzeptanz in der Profession hängt – zusätzlich zur objektiven Qualität der Bildungsstandards und Tests – einerseits von der Zielklarheit einer adressatenbezogenen politisch-administrativen Kommunikation der Reform, andererseits von ihrem erfahrbaren Nutzen in der täglichen Unterrichtsarbeit der Lehrpersonen ab.*

Damit Bildungsstandards den Weg von der bildungspolitischen Systemebene zu den Akteuren finden und dort zur Etablierung eines neuen Standards des Lehrerhandelns beitragen, muss die neue bildungspolitische Strategie und müssen die Eigenschaften der Neuerung überzeugend und mit Vorteilen für die Akteure begründet und *kommuniziert* werden. Das heißt, die Lehrkräfte und die weiteren Akteure müssen dem Kernanliegen positiv gegenüberstehen und in der Arbeit mit Standards einen *Nutzen* für ihre Arbeit sehen. Dies ist dann der Fall, wenn Bildungsstandards nicht einfach als neuartige curriculare Zwangsjacke, mit der sich bloß ein erhöhter Zielerreichungsdruck verbindet, wahrgenommen werden, sondern wenn sie die Gestaltungsräume und Möglichkeiten erkennbar erweitern, Schule und Unterricht nach eigenem Verständnis zu planen und durchzuführen – dies mit einem wahrnehmbaren Nutzen für die Schülerinnen und Schüler. Sodann müssen die mit dem Projekt „Bildungsstandards“ verbundenen Entwicklungsaufgaben mit den gegebenen Mitteln lösbar erscheinen, ohne eine lästige Überforderung darzustellen. Und schließlich muss die Einführung von Standards als gemeinsame professionelle Aufgabe wahrgenommen werden, die nicht nur die Lehrkräfte betrifft, sondern auch die Arbeit der Unterstützungssysteme einbezieht.

Insgesamt lässt sich das Ziel der Akzeptanz auch so umreißen: Es muss gelingen, die Lehrpersonen als Hauptakteure im Bildungswesen von der Wichtigkeit solider inhaltsbezogener Referenzpunkte für die Ziele des fachlichen Lernens zu überzeugen. Bildungsstandards, verstanden als gestufte Kompetenzzan-

förderungen, stellen einen Bezugsrahmen dar, der es über die Rückmeldung von Testergebnissen ermöglichen soll, Distanz – Nähe und Abstand – zu einem angestrebten Lernziel einzuschätzen und damit sich selbst bzw. die eigene Klasse in einem gemeinsamen Bezugssystem zu sehen.

Bisherige Studien aus den Auftragsländern zu den Einstellungen und den Bedingungen der Akzeptanz von Bildungsstandards bei Lehrkräften verweisen auf die Notwendigkeit nicht nur einer dezidierten Überzeugungsarbeit, sondern auch einer gezielten Vorbereitung. Die Ergebnisse einer österreichischen Studie (Freudenthaler/Specht 2005) zeigen einerseits, dass die meisten der befragten Lehrkräfte (mit Abweichungen zwischen Schulformen und Fächern) eine kritisch-offene Haltung gegenüber Bildungsstandards haben. Die Haltung lautet: Mal sehen, was kommt. Andererseits wird erstens sichtbar, dass der propagierte Nutzen und die Notwendigkeit dieses neuen Steuerungsmittels zur Lösung vorherrschender Probleme des Schulsystems mehr oder weniger stark bezweifelt werden. Zweitens zeigt sich auch, dass diese Reform verbunden ist mit zahlreichen Unsicherheiten, Ängsten und Befürchtungen, die die Effektivität der Bildungsstandards von vornherein unterlaufen, wenn sie bestehen bleiben.

Diese beiden letzten negativen Befunde widerspiegelt auch die Auswertung einer Befragung des Verbands Bildung und Erziehung (VBE) für Deutschland.¹⁴⁶ Generell kommen in der Einleitung von Verbandseite Befürchtungen im Sinne eines Technologieverdachts zum Ausdruck, die „unser gewachsenes Bildungsdenken, das sich auf die Tradition der Aufklärung und der pädagogischen Reformbewegung stützt, in Frage stellt“ (2005, S. 1). Gegenüber vermehrten Außenkontrollen durch Schulinspektionen, Zielvereinbarungen, Punktbewertungen oder Leistungskontrollen präferiert der Verband *Selbst*bewertungen, Portfolio-Modelle, *interne* Qualitätsüberprüfungen und *freiwillige* Weiterbildung (ebd.), wodurch sich eine deutliche Ablehnung von externer Kontrolle sowie der Wille zur Bewahrung der Autonomie der Lehrkräfte zeigen.

Als Fazit ist festzuhalten, dass unter den befragten deutschen Lehrkräften eine große Skepsis gegenüber Bildungsstandards und Vergleichsuntersuchungen vorherrscht. Die Vorbehalte kommen von der Basis, also von den Lehrkräften, die fürchten, überwacht zu werden. Die Befürchtungen nehmen innerhalb der Schulhierarchie nach oben hin ab. Deutsche Grundschullehrkräfte hegen die meisten Zweifel an den vorgesehenen Schulreformen. Schulleiter begrüßen Evaluationsmaßnahmen und Standards schon viel eher, Schulaufsichtsbeamte,

¹⁴⁶ VBE-Mitgliederbefragung von 2005. Für diese Information danken wir Kurt Czerwenka.

Bildungspolitiker und bestimmte Hochschullehrer sind als die Motoren der Reformen zu sehen. Ähnliches gilt übrigens für die Akzeptanz von Schulrückmeldungen, die von Schulleiterinnen und Schulleitern stärker begrüßt werden als von Lehrkräften.

Aus Studien wie diesen wird deutlich, dass das Prinzip der Rechenschaftslegung akzeptiert wird, wenn die Autonomie des Lehrkörpers und der Schulen gewährleistet bleibt. Das erklärt etwa die angesichts des Arbeitsaufwandes an sich erstaunliche Bevorzugung interner Evaluationsmaßnahmen. Sie gelten als wenig bedrohlich, während externe Evaluation und Unterricht unter der Voraussetzung von Bildungsstandards als die Autonomie gefährdend eingestuft werden. Das verweist auf einen erheblichen Informationsbedarf. Wenn die Implementation von Bildungsstandards den Lehrkräften nicht kommuniziert wird, wenn sie ihr Know-how nicht entsprechend entwickeln können und nicht hinter dem Projekt stehen, werden sie die Reform unterlaufen.

Ditton/Merz/Edelhäuter (2002), die eine Einstellungsbefragung bei Lehrpersonen und Schulleitungen durchgeführt haben, berichten von differenzierter gewordenen Meinungen. Nach Auffassung der Autoren lässt sich gegenüber dem Stand von vor einigen Jahren eine Tendenz hin zu einer positiveren Aufnahme insbesondere externer Evaluationsverfahren bei Lehrpersonen erkennen. Voraussetzung dafür ist eine genauere Kenntnis der politischen Strategie und der damit verbundenen Vorteile für die eigene Praxis. Abwehr entsteht oft aus Unkenntnis oder Analogieschlüssen; Zustimmung verlangt eine positive Erwartung des Nutzens.

Für die *Schweiz* gibt es keine entsprechenden Daten. Bekannt ist, dass der LCH (Dachverband Schweizer Lehrerinnen und Lehrer) eine pragmatische Haltung einnimmt und dem Anliegen Bildungsstandards positiv-offen gegenübersteht. Bedingung dafür ist, dass es gelingt, Chancen und Mehrwert des Projekts zu kommunizieren und „Bildungsstandards (...) verträglich in das bestehende System [zu integrieren], ohne übertriebene Heilserwartungen zu wecken oder Ängste zu schüren.“¹⁴⁷

(4) Verankerung der Reform in den regionalen und lokalen Kontexten einer auf den Unterricht zielenden pädagogischen Qualitätsentwicklung.

Neben tauglichen Instrumenten und Strategien auf der Systemebene, d.h. einer Makropolitik der Verbreitung und Etablierung von Standards und Tests, erfordert die Implementation des bildungspolitischen Programms „Bildungsstandards“ eine überzeugende Anbindung an eine

¹⁴⁷ So der Zentralpräsident LCH, Beat Zemp, am 26. Januar 2006 in Murten.

Mikropolitik breitenwirksamer Qualitätsentwicklung mit der Zielebene des Unterrichts und damit der Lernprozesse von Schülerinnen und Schülern.

Während die Probleme der Implementation aus politisch-administrativer Sicht, neben der kommunikativen Verbreitung der neuen Steuerungsphilosophie, vor allem in der Entwicklung und Testmäßigen Überprüfung qualitätsvoller Standards sowie in der Rückmeldung der Resultate gesehen werden können, handelt es sich unterhalb der politischen Ebene um die produktive, schul- und unterrichtsbezogene pädagogische Nutzung der neuen Instrumente. Damit ist die Ebene des professionellen Lernens und der Schulentwicklung als weitere entscheidende Ebene der Implementation von Bildungsstandards angesprochen (vgl. z.B. Altrichter/Posch 1996). Entsprechend einer von der Theorie der Schulentwicklung postulierten (und durch den Forschungsüberblick in Kap. 3 bestätigten) Abkehr von Top-down-Strategien des verordneten Wandels hin zu Strategien, die sich am professionellen Lernen teilautonomer Schulen und Lehrerkollegien orientieren, bildet die Qualität einer auf die neue Steuerungsphilosophie ausgerichteten Schul- und Unterrichtsentwicklung den zweiten Pfeiler, mit dessen Stabilität das Unternehmen „Bildungsstandards“ steht und fällt. Damit wird auch deutlich, dass die Einführung von Bildungsstandards nur über eine Langzeitstrategie zu erreichen ist, die über eine reine Outputsteuerung hinausgeht, den Prozesscharakter der Reform anerkennt und über eine einzelne Legislaturperiode hinausreicht.

(5) Ausrichtung von Unterstützungssystemen, Werkzeugen und Verfahren des professionellen Lernens und der Schulentwicklung auf Bildungsstandards. *Ergänzend zur Schaffung geeigneter materieller Bedingungen und Freiräume für eine gedeihliche Qualitätsentwicklung besteht eine vorrangige Aufgabe der schulbezogenen Unterstützungssysteme darin, durch taugliche Verfahren und Werkzeuge für Prozessqualität im Lernen von Lehrpersonen, Kollegien und Schulen zu sorgen.*

Letztlich wird sich die Wirkung der Reform daran bemessen, ob das neue Referenzsystem für Ziele und Inhalte und eine darauf bezogene Testkultur als Qualitätsrahmen Eingang in den Unterricht finden und Einfluss darauf nehmen, was Schülerinnen und Schüler tatsächlich lernen, d.h. welche zum Teil messbaren Kompetenzen sie in welchen Domänen aufbauen. Für die Lehrpersonen heißt dies, dass sie ein tiefes Verständnis der Inhalte und Ziele der Innovation erwerben, was Wege des Informationsflusses erfordert, die über amtliche Verlautbarungen und Dokumente hinausgehen. Zentrales Mittel dieses Lern- und

Wissenstransferprozesses sind verschiedene Formen und Verfahren der individuellen und kollektiv-kooperativen Fortbildung. Dies setzt auf den regionalen und lokalen Ebenen des Bildungssystems neben Autonomie- und Gestaltungsspielräumen die Verfügbarkeit von Werkzeugen und Verfahren des professionellen Lernens voraus. Während einige Wege des Wissenstransfers und seiner Instrumente sich seit langem bewährt haben, müssen auch neue Verfahren der Fortbildung entwickelt werden. Zum Teil geht es darum, sich mit Werkzeugen und Methoden, die sich in einem anderen Schulsystem bewährt haben, unter Berücksichtigung von Passungsbedingungen, die sich auf den eigenen Kontext beziehen, vertraut zu machen. Solches geschieht gegenwärtig mancherorts z.B. mit der (probeweisen) Einführung von externen (Schulentwicklungs-, Evaluations-)Beratungsdiensten.

(6) Fachliche Bildungsstandards im Rahmen einer breiteren pädagogischen Zielsetzung von Schule. *Bildungsstandards verstanden als überprüfbare fachliche Kompetenzen beziehen sich weder auf alle Fächer noch erfassen sie die gesamte Breite curricularer Ziele bezogen auf einen multikriterialen Bildungsbegriff. Eine leistungsfachliche Verengung des schulischen Bildungsgedankens durch evaluativ gebrauchte Standards muss vermieden werden. Neben und jenseits von normierten Kerninhalten und Grundkompetenzen, die von Bildungsstandards angesprochen und durch Tests geprüft werden, soll genügend Raum bleiben für Lernprozesse, Stoffe und Kompetenzen, die sich an einem breiten Begriff fachlicher Bildung sowie an überfachlichen Bildungs- und Schulstandards orientieren.*

In der postulierten Form überprüfbarer Basiskompetenzen, die sich an elementaren disziplinären Konzepten von „literacy“ orientieren, erfassen Bildungsstandards nicht das ganze Spektrum dessen, worauf sich Schule und Unterricht auf allen Stufen fachlich und pädagogisch ausrichten sollen. Anspruchsvolle Prozessstandards wie etwa der (mündlichen) Kommunikation, der fachlichen Diskursivität sowie schwer objektiv messbare höhere Fachstandards des Verstehens, des kreativen Problemlösens oder des ästhetischen Empfindens u.a.m. (vor allem in höheren Schulen) müssen ihren Stellenwert als ebenfalls zentrale Bildungsziele behalten. Aufgabe der in Diskussion stehenden Bildungsstandards im engen Sinne und der darauf bezogenen Tests ist es, einen wirksamen Beitrag dazu zu leisten, dass minimale Sockelziele von allen Schülerinnen und Schülern in einem Schulsystem erreicht werden. Mehr nicht, aber auch nicht weniger. Der curriculare Gestaltungsraum von Schulen und Lehrpersonen darf durch Standards und Tests, die das „Gut Bildung“ nur in begrenztem Maße messbar machen, nicht eingeschränkt werden. Sodann ist darauf zu achten, dass


sich Bildungsstandards (als Minimal- oder Regelstandards) nicht einseitig an den oberen Leistungsspektren orientieren. Gerade die elementaren Kompetenzstufen sollen so ausgestaltet werden, dass insbesondere leistungsschwache Schülerinnen und Schüler nicht unter die Räder fachdidaktisch überrissener Ansprüche kommen. Auch zu diesen Problemen hat eine balancierte Strategie der Implementation der neuen Instrumente einer elementaren Qualitätssicherung Rechnung zu tragen.

(7) „Input-“, „Prozess-“- und „Ergebnisqualität“ als die drei Säulen schulischer Wirksamkeit. *Auch bei der uneingeschränkten Bejahung der Entwicklung, Einführung und evaluativen Anwendung von Bildungsstandards darf die damit angestrebte Outputsteuerung des Bildungssystems nicht zur Monostrategie und einzigen Säule der Qualitätssicherung und Weiterentwicklung von Schule und Unterricht werden.*

In Übereinstimmung mit einem von der Forschung geteilten Verständnis von Qualitätssicherung und -entwicklung bedeutet die Einführung und evaluative Anwendung von kriteriumsorientierten Bildungsstandards keinen Bruch mit bisherigen Verfahren der Qualitätssicherung und auch keinen Strategiewechsel im Sinne einer radikalen Neuorientierung, sondern eine notwendige und sinnvolle *Ergänzung* eines von den Kriterien, Perspektiven und Blickwinkeln her breit gefassten Systems der Qualitätssteuerung im unter öffentlicher Verantwortung und Kontrolle stehenden Bildungswesen. Abbildung 5 illustriert die drei Säulen eines integralen Konzepts der Qualitätssicherung und -steuerung im Bildungswesen, wie sie auch nach einer Einführung von Bildungsstandards Gültigkeit haben werden. Qualität im Bildungswesen kann nur durch das Zusammenführen von *Input* (Qualität von Ausgangs- und Rahmenbedingungen), *Prozess*(merkmalen auf Schul- und Unterrichtsebene) und *Output* (Lernerträgen gemessen an multikriterialen Leistungsmaßen), und durch wechselseitige Rückkoppelungen und Informationsrückflüsse zwischen diesen Säulen sowie das Ziehen der Konsequenzen in Bezug auf Entwicklungsaufgaben und -ziele nachhaltig gesichert und gesteigert werden (vgl. auch Klieme 2005). Der Knackpunkt einer outputorientierten Steuerung besteht dabei in der Rückkoppelung von durch Tests und andere Evaluationsverfahren ermittelten Lernergebnissen in schulbezogene Entwicklungsmaßnahmen sowie in verändertes Unterrichtshandeln – was einen weit komplexeren Auswertungs- und Interpretationsvorgang darstellt als eine mit einer noch so ausgefeilten Leseanleitung versehene und bezüglich einer Vielzahl von Kontrollvariablen „faire“ Ergebnisrückmeldung.

Abbildung 5: Die drei Säulen der Steuerung und Qualitätssicherung des Bildungssystems

INPUT <i>(gesteuert über Ziele und Ressourcen)</i>	PROZESS <i>(gesteuert über die Qualität der Lehr-Lernprozesse)</i>	OUTPUT <i>(gesteuert über Resultate und Rückmeldungen)</i>
<ul style="list-style-type: none"> ● Kohärente Lehrpläne und Lehrmittel, welche an Standards ausgerichtet sind, ● gut ausgebildete und engagierte Lehrpersonen, ● Struktur des Bildungswesens und Ausstattung von Schulen, ● effektive Szenarien und Werkzeuge für kontinuierliche Lehrer(fort)bildung und Unterrichtsentwicklung. 	<ul style="list-style-type: none"> ● internalisierte und klare Erwartungshaltungen von Lehrpersonen und Schülerinnen und Schülern, ● Schul- und Fachinteresse, positive Lerneinstellung, ● Standards für Unterrichtsqualität und Schulklima (z.B.: kognitive Aktivierung, Klassenführung, Schülerorientierung, Qualität des Fachwissens sowie der pädagogischen Interaktion). 	<ul style="list-style-type: none"> ● auf Bildungs- und Schulstandards basierende multi-dimensionale und niveaugestufte Leistungstests, ● auf Verfahren der Fremd- und Selbstevaluation basierende Qualitätseinschätzungen der schulischen Arbeit, ● angemessene Rückmeldung in Bezug auf die Realisierung von Standards und Entwicklungsfortschritten an Schülerinnen und Schülern, Eltern und Schulträger.



Bei allen diesen Leitvorstellungen muss der Blickwinkel von und auf Schulen und Lehrpersonen zentral bleiben. Dieser Blickwinkel manifestiert sich z.B. in konkreten Fragen, die sich die Akteure an der Basis des Bildungssystems stellen (und stellen müssen) und auf die sie eine befriedigende Antwort erwarten sowie verlässliche Unterstützungen, was sie im Blick auf die Antwort tun können. Wenn z.B. Gewerkschaften (Gewerkschaft Erziehung und Wissenschaft (GEW) 2004b, S. 214f.) dies einklagen, so hat das seinen guten Sinn:

- Können Standards helfen, meinen Unterricht zu verbessern und junge Menschen wirkungsvoller zu fördern?
- Wie kann ich meine diagnostischen Kompetenzen verbessern?
- Gestufter Kompetenzaufbau – Kompetenzmodelle – wo kann ich mich fortbilden?

- Was geschieht mit Ergebnissen von Lernstandserhebungen? Wem „gehören“ die Daten?
- Bekomme ich Unterstützung, wenn meine Klasse oder einzelne Schülerinnen und Schüler schlecht abschneiden oder werde ich angeprangert? Muss ich gar mit Sanktionen rechnen?
- Wie kann man Kompetenzstandards überprüfen, ohne dass zuvor ein systematischer Kompetenzaufbau stattgefunden hat?
- Wie wird der Tatsache Rechnung getragen, dass Schulen und einzelne Schulklassen sehr unterschiedliche Bedingungen haben? Kann man sich darauf verlassen, dass es kein Schulranking gibt?
- Die Standards sollen zu mehr Bildungsgerechtigkeit und Vergleichbarkeit von Schulnoten und Schulabschlüssen beitragen. Ist auch an eine Angleichung der Rahmenbedingungen wie Jahreswochenstunden und Stundenumfang sowie der fachlichen Qualifikation der Lehrkräfte gedacht?

Solche Fragen muss eine politische Konzeption der Einführung von Bildungsstandards beantworten können. Die Antworten werden nicht in allen Auftragsländern gleich ausfallen, aber sie müssen erarbeitet und kommuniziert werden.

4.2.2 Strategien der Einführung und Sicherung von Bildungsstandards

Eine gelingende Implementation setzt Strategien auf der Systemebene und in Bezug auf die pädagogische Qualitätsentwicklung an der Basis voraus. Auf der Grundlage der Leitideen lassen sich komplementäre Strategien unterscheiden, welche in kultureller Passung in allen Auftragsländern entwickelt werden müssen: eine systemische und politikakzentuierte und eine schulnahe, am Kerngeschäft Unterricht orientierte. Während sich die erste Strategie um eine kohärente Politik der Einführung und Umsetzung auf der Makro- und Mesoebene des Bildungssystems bemüht, richtet sich die zweite auf die Mikropolitik der Stimulation und Unterstützung des professionellen Lernens und der schulischen Qualitätsentwicklung, das heißt, auf Wissenstransfer durch moderne Formen der Fort- und Weiterbildung.

Politische Makrosteuerung in kultureller Passung – oder „Accountability“ in der Perspektive der Bildungsadministration

Erfahrungen mit der Implementation von standardbasierten Schulreformkonzepten, vor allem in den USA, aber auch in anderen von uns näher untersuchten Ländern, sowie die dazu verfügbare Literatur machen deutlich, dass es sich

dabei um einen komplexen Prozess handelt, der die Lern- und Umsetzungsprozesse auf den verschiedenen Ebenen des Bildungssystems betrifft. Qualität kann sich dabei nur in einem Wechselspiel von äusserer Impulsgebung und produktivem individuellem und kollektivem Lernen an der Systembasis entfalten. Das heisst, dass Initiierung und Steuerung eines umfassenden Projekts des Wissenstransfers, wie es die Einführung von Bildungsstandards darstellt, einer in kultureller Passung erfolgenden Verbindung von Top-down- und Bottom-up-Strategien bedarf. „Kulturelle Passung“ meint dabei, dass die Lösungen in jedem Land anders aussehen und Problemlösungen nicht einfach auf einen anderen Kontext übertragen werden können. Weiter bringt es die politische Dynamik mit sich, dass auch immer wieder Strategiewechsel zu beobachten sind; so zum Beispiel in den traditionell föderativen Vereinigten Staaten, wo sich nach dem Krisenreport *A Nation at Risk* (National Commission on Excellence in Education 1983) mit dem *No Child Left Behind Act* (2001) erstmals ein bislang nicht gekanntes Mass an Bundessteuerung durchgesetzt hat. Oder in England, wo 1988 ein engmaschiges zentrales Kontrollsystem aufgebaut wurde, das nach 1997 zugunsten stärkerer lokaler Kontrollen wiederum etwas zurückgenommen wurde.

Die Tatsache kultureller Differenz bedeutet, dass Reformen an das bestehende System und damit an eine historisch gewachsene „grammar of schooling“ anschlussfähig sein und mit der jeweiligen politischen Systemdynamik leben müssen. Das gilt auch für die Auftragsländer. Während die Diskussion um eine wirksamere Bildungssteuerung z.B. in Deutschland auf der Basis eines Systems geführt wird, in dem eine gleichheitsorientierte hohe Regelungsdichte gepaart mit einer zentralen Dienstaufsicht durch die Bundesländer (deren bevölkerungsreichste grösser sind als z.B. die Schweiz) bzw. den Staat Tradition hat und stark ausgeprägt ist (auf der gymnasialen Stufe kommen zusätzlich noch die zentralen Abiturprüfungen dazu), ist das Schweizer System wesentlich schwächer zentral gesteuert. Obgleich die Kantone (nicht der Bund) ebenfalls eine mittlerweile bedeutende Regelungsdichte aufweisen, erfolgt die Standardsicherung (inkl. Lehrerwahlen) im Wesentlichen immer noch lokal über meist gemeindeeigene Schulaufsichtsbehörden (Schulkommissionen, Schulpflegen) sowie ergänzend über die (inter)kantonale Zulassung von Lehrmitteln.

Bildungssysteme sind kollektive Akteure, die aus funktional ausdifferenzierten Sektoren und Teilgruppen bestehen, in denen individuelle Personen auf verschiedenen Ebenen zusammenwirken, idealerweise um gemeinsam geteilte Ziele und Ergebnisse zu erreichen. Um Systeme solcher Komplexität zu steuern, müsste man deren Dynamik kennen. Gerade diese Transparenz ist aber kaum je gegeben. Bei der Systemsteuerung des Bildungswesens handelt es sich

im Sinne von Dörner (1983) um ein Problem des Umgangs mit einem hohen Maß an Komplexität, Eigendynamik und Intransparenz von Seiten der steuernden Instanzen – um ein Problemlösefeld mithin, in dem zwangsläufig immer auch unerwünschte Nebenwirkungen auftreten; dies umso mehr, je ausgeprägter eine wenig Feedback-sensible Top-down-Strategie der Steuerung befolgt wird.

Zieht man Bilanz über die in Kapitel 3 dieser Expertise vorgestellte Analyse des Forschungs- und Entwicklungsstandes in vier ausgewählten Ländern, so zeigt sich im Einklang mit der neueren Transfer- und Innovationsforschung und in Übereinstimmung mit einem ebenfalls von der Schulentwicklungsforschung seit längerem vollzogenen Perspektivenwechsel, dass dominante Top-down-Strategien bei der Steuerung komplexer Systeme wenig erfolgreich sind. Wer Standards ohne Berücksichtigung von auf der intermediären Ebene der Schule auftretenden Bedürfnissen und Problemen vor allem mit behördlicher Verwaltungskommunikation, Gesetzen und Verordnungen gleichsam aus der Ferne durchsetzen will, wird auch dann die eigenen Reformziele nicht erreichen, wenn die ausgearbeiteten Bildungsstandards inhaltlich Sinn ergeben und die Testaufgaben, in die sie münden, von hoher Qualität sind.

Der Ländervergleich zeigt jedoch ebenso deutlich, dass auf verbindliche Vorgaben und klare Impulse „von oben“, d.h. auf Top-down-Elemente und Anreize der Steuerung, nicht verzichtet werden kann. Allen vier bezüglich ihrer Strategien von Qualitätssicherung und „Accountability“ näher analysierten Ländern ist gemeinsam, dass sie in irgendeiner Form Tests oder externe Prüfungen verwenden, ergänzt durch interne und externe Evaluationen oder Inspektionen. Auch wenn die curricularen Standards in den Ländern verschieden ausgerichtet sind (als enge Vorgaben in England oder als Kernziele in den Niederlanden) und die systemischen Strategien, diese verbindlich zu machen, verschieden sind (zwischen gouvernemental bzw. auf die Tätigkeit von Behörden setzend und eher entwicklungs- und professionsorientiert), so besteht kein Unterschied in Hinsicht auf die Verbindlichkeit einer Rechenschaftspflicht, die nicht nur aus internen Rückmeldemechanismen besteht, sondern in irgendeiner Weise auf überprüfbare Ergebnisse ausgerichtet ist.

Mit Bezug auf eine *Innovationssteuerung auf der Ebene des Gesamtsystems* heißt dies, dass es auf eine kontextuell verträgliche Balancierung von Anreizen und Kontrollen, die einerseits von oben, andererseits von unten bzw. von intermediären Ebenen ausgehen, ankommt. „Accountability“, also die Verantwortung oder Haftbarkeit für den Bildungserfolg oder -misserfolg, verweist zudem auf eine Aufgabe, die sowohl der Systemspitze oder den Bildungsbehörden als auch den intermediären Systemgliedern und den individuellen Personen an der Sys-

tembasis zukommt. Dabei gilt es zu beachten, dass sich das Verhältnis zwischen Steuerung und Selbstbestimmung im Verlauf der Implementation ändert: Während in der Startphase eine Top-down-Strategie notwendig erscheint, soll es im weiteren Verlauf zu einer Rücknahme der Lenkungs Vorgaben kommen, welche mit einer Zunahme der Eigenverantwortlichkeit der einzelnen Schulen einhergeht.

Folgende Verantwortungen sollen nach Meinung dieser Expertise – als Elemente einer Makrosteuerung – vorzugsweise bei der *Systemspitze* und damit bei den politisch-administrativen Behörden des Bildungssystems angesiedelt bzw. *belassen* werden:

- *Steuerung von Standardentwicklung und Testinstrumenten als nationale Aufgabe.* Der Systemspitze obliegt die wertbezogene und theoriegeleitete Entwicklung und Operationalisierung qualitätvoller Standards und der zugehörigen Testaufgaben auf verschiedenen Anspruchsniveaus auf der Basis einer gesellschaftlichen Auseinandersetzung mit den inhaltlichen Aufgaben von Schule und Bildung sowie eines anzustrebenden nationalen Konsenses über das kulturelle Programm des Bildungssystems und seiner einzelnen Sektoren. Dabei ist eine Abstimmung zu suchen zwischen den stoff- bzw. lehrplanorientierten und den im engeren Sinne standard- bzw. kompetenzorientierten Teilen eines solchen Programms. Die wissenschaftliche und technische Infrastruktur im Dienste der durch Bildungsstandards angestrebten Beobachtung und ergebnisorientierten Steuerung des Bildungssystems muss bereitgestellt werden.
- *Festlegung der operativen Elemente einer standardbezogenen Leistungsevaluation.* Zur Makrosteuerung gehört weiter die Festlegung der Periodizität, der Fächerbreite und der Stichprobenbildung im Hinblick auf den Einsatz von Testinstrumenten in den Schulklassen sowie die Verantwortung für eine qualifizierte Rückmeldung der Ergebnisse unter Ansehung relevanter Testeingangsvoraussetzungen. Weiter gehört dazu die Entwicklung von Konzepten der Förderung und der Zuweisung von Mitteln und Maßnahmen an jene Schulen, deren Schülerinnen und Schüler die Mindestanforderungen nicht erfüllen. All dies soll unter der Leitidee größtmöglicher kommunikativer Offenheit und Transparenz erfolgen. Übereinstimmend mit Erfahrungen anderer Länder sowie auf der Basis von Überlegungen aus der pädagogischen Lern- und Motivationspsychologie plädieren wir jedoch nicht für öffentliche Schulrankings.
- *Etablierung neuer Balancen zwischen zielbezogener Ergebnisverantwortung und Autonomie in der Prozessgestaltung.* Eine auf eine erhöhte Qualitätsverantwortung von

Einzel Schulen zielende externe Evaluation verlangt nach einer Stärkung der Autonomie der Schulen als professionelle Handlungseinheiten. Abhängig von den bestehenden Konfigurationen der Bildungssysteme in den Auftragsländern erfordert dies teils tiefgreifende Maßnahmen im Bereich einer Reform des Verwaltungshandelns. Schulen, deren Leistungen auf Basis kriteriumsorientierter Tests periodisch überprüft werden, sollen erhöhte Freiräume in der Gestaltung ihrer Stundentafeln, der Detaillierung ihrer Lehrpläne und ihres Schulalltags erhalten. Auch sollten sie Professionalisierungs- und Unterstützungsangebote im Sinne von Serviceleistungen vermehrt selber *anfordern* können. Solche zu Leistungstests und Ertragssteuerung komplementären Maßnahmen erfordern nicht nur einen sich wandelnden Verantwortungssinn der Einzel Schulen, sondern auch ein verändertes Rollenverständnis der Bildungsadministration und der Unterstützungssysteme. Konkrete Maßnahmen, welche gleichzeitig die Autonomie und Selbstverantwortung von Einzel Schulen als auch deren Verpflichtung zur Erreichung bestimmter fachpädagogischer Leistungsziele im Fokus haben, könnten etwa durch die Administration gesetzte verbindliche Zielvereinbarungen sein. Eine Strategie der Bildungsadministration sollte darin bestehen, solche und andere rahmengebende Vereinbarungen mit Schulen betreffend die Verknüpfung von Zielverbindlichkeit und Ergebnisverantwortung nach außen und erweiterten Gestaltungsfreiräumen nach innen in rechtsverbindlicher Form zu etablieren.

- *Bereitstellung einer strukturell-organisatorischen und personellen Infrastruktur im Rezeptionsfeld von Bildungsstandards einschließlich der Verpflichtung der Kontext- und Unterstützungssysteme (Lehrerbildung und -fortbildung, Pädagogische Arbeitsstellen und Landesinstitute, Schulaufsichtsgremien, Beratungsdienste usw.) als Mitverantwortungsträger für die Einführung in die neue Philosophie der Rechenschaftslegung.* Den Kontextsystemen kommt dabei neben der Vermittlung des notwendigen Wissens über Standards, Kompetenzen und Tests sowie neben dem Umgang mit Leistungsrückmeldungen die Aufgabe zu, Schulen und Lehrpersonen ebenfalls mit komplementären, multiple Kriterien einbeziehenden Formen der Fremd- und Selbstevaluation und mit Verfahren und Instrumenten der pädagogischen Qualitätsentwicklung vertraut zu machen. Aufgabe der Makrosteuerung müsste dabei sein, die Abstimmung der Aufgaben unter den Unterstützungssystemen zu fördern und mit diesen ebenfalls zu – gegebenenfalls veränderten – Zielvereinbarungen zu kommen. Das heißt, es geht darum, dass die bestehenden Unterstützungssysteme ihre Aufgaben für und Dienstleistungen an die Schulen ebenfalls überdenken und allenfalls neu ausrichten, zumal neue Bedürfnisse nach Dienstleistungen und Unterstüt-

zungsformen hinzukommen werden. Dabei soll auch die Option nicht ausgeschlossen werden, Mittel aus bestehenden Diensten in neue Aufgabenfelder, vorzugsweise in unterrichtsnahe Serviceleistungen, die den Lehrkräften helfen, die Bedingungen des Schülerlernens zu verbessern, umzulagern. Solche Aufgaben sind z.B. die Ausbildung von fachpädagogischen Coaches und Lehrberatern oder die Unterstützung des Aufbaus kooperativer professioneller Netzwerke.

- *Wahl einer optimalen Steuerungs- und Unterstützungsphilosophie.* Welches Modell der strategischen Makrosteuerung in einem Bildungssystem praktiziert wird, und damit die Konfiguration der praktizierten Maßnahmen, hängt letztlich davon ab, welchem Zweck die Implementation der Bildungsstandards dienen soll: Steht primär die Kontrollfunktion, die Rechenschaftsfunktion oder die Qualitätsentwicklungsfunktion im Vordergrund? Insgesamt lassen sich die Strategien der Standardsicherung und -unterstützung zwischen Polen wie Accountability vs. Entwicklung, Kontrolle vs. Autonomie, Förderung/Unterstützung vs. Zielerreichungsdruck – allgemein: zwischen Angebotsorientierung und Nachfrageorientierung, Zug- und Druckkomponenten – anordnen. Fend (2003) macht in seiner Darstellung vier Strategien prototypisch sichtbar:

Abbildung 6: Typologien von Reformmaßnahmen nach Fend 2003



Quelle: Fend 2003, S. 36.

Im Hinblick auf die konkrete Einführung von Bildungsstandards in den Auftragsländern gilt es, kontextuell passungsfähige Balancen zwischen diesen Polen zu finden, also zwischen der Eigenverantwortung von Schulen, Massnahmen der externen Kontrolle und Evaluation, Unterstützungsangeboten und der Setzung administrativ-rechtlicher und organisatorischer Rahmenbedingungen. Wichtig ist die Feststellung, dass jedes Bildungssystem seine eigene Konfiguration aufweist aufgrund historischer und kultureller Voraussetzungen, die nicht austauschbar sind. Welches System der Qualitätssicherung gewählt wird, hängt vom bestehenden Bildungssystem und seinen spezifischen Bedürfnissen ab und kann nur im entsprechenden kulturellen Kontext entschieden werden. Die von uns untersuchten Länder unterscheiden sich zum Teil recht deutlich in den Konfigurationen der Qualitätssicherung und den verwendeten Verfahren sowie in der Geschichte der Bildungsreformen und Innovationen, die den Ist-Zustand hervorgebracht haben. Auch in den Auftragsländern sind die Traditionen verschieden. Was mit einiger Sicherheit jedoch gesagt werden kann ist, dass von Qualitätssicherungssystemen, die vor allem Top-down-Strategien einsetzen und „Accountability“ für den Bildungserfolg mit „high-stakes tests“ verbinden, nur sehr begrenzt positive Impulse für die Schulentwicklung ausgehen, nicht zuletzt, weil diese demotivierend auf die Selbstverantwortung und Selbstwirksamkeit der Akteure wirken und den Grad der Flexibilität und Eigenkontrolle der

Aktivitäten vor Ort gefährden. Kein wirksames System der Qualitätssicherung wird deshalb nach heutigem Wissen ohne (Beibehaltung einer) substanzielle(n) Verantwortungsdelegation an die Basis bzw. starke lokale – planerisch-gestalterische und budgetäre – Zuständigkeiten auskommen.

Eine systembezogene Makropolitik muss sich um den Aufbau von kohärenten Strukturen und Kapazitäten (Ressourcen, Dienstleistungen) auf allen voneinander abhängigen und interagierenden Ebenen des Systems bemühen. Es geht darum, dass die Einführung und Anwendung von Standards und der Umgang mit den damit gekoppelten Evaluationsverfahren systemverträglich erfolgen und sowohl von Schulen als auch von deren Unterstützungssystemen als positive Aufgabe wahrgenommen werden.

Dazukommen muss, gesteuert bzw. unterstützt von der Systemspitze (Bildungs- und Schuladministration), eine kohärente Reformpolitik im Sinne einer gegenseitigen Abstimmung und gemeinsamen Ausrichtung („alignment“) von Strukturen, Strategien, Regelungen und Materialien mit Blick auf die Zielsetzung der Reform, und dies auf allen Ebenen des Systems. Neu an den im Zusammenhang mit der Einführung von Bildungsstandards favorisierten Steuerungsmodellen ist die Erhöhung der Flexibilität und Verantwortung der lokalen Einheiten. Die Aufgabe von Ministerien und Bildungsverwaltungen besteht nicht mehr darin, den Alltag von Schulen durch ein engmaschiges Netz von Eingriffen und Vorschriften zu steuern, sondern vielmehr darin, für geeignete Rahmenbedingungen und Instrumente sowie qualitativ hochstehende Unterstützungssysteme (Lehrerbildung, Fortbildung, Beratung, Instrumente der Selbst- und Fremdevaluation usw.) zu sorgen. Unter der Leitidee teilautonomer Schulen geht es darum, administrative Kontrolle abzugeben, um komplementär, durch Selbststeuerung, Partizipation und vermehrte Transparenz an der Basis einen Zugewinn an Effektivität und Qualität zu erzielen.

Ein zentraler Punkt, was die gemeinsame Ausrichtung von Reformzielsetzungen und Strukturen unter dem Blickwinkel von *Bildungsgerechtigkeit* angeht, dürfte die Frage nach einer Reform der Schulstrukturen bzw. der *Schulformen* sein. Bildungsstandards mit Bezug auf Kompetenzmodelle ergeben im Prinzip nur dann Sinn, wenn sie nicht mehr auf bestimmte Schultypen, sondern auf die entwicklungs- und altersbezogene Leistungsfähigkeit der einzelnen Schülerinnen und Schüler bezogen werden. Weitere Fragen, die in eine ähnliche Richtung weisen und keinesfalls unter den Tisch gewischt werden dürfen, betreffen die Angleichung von Rahmenbedingungen auf der Ebene von Stundentafeln (Unterrichtsumfang, Fächerspektrum), die Vergleichbarkeit von Abschlüssen sowie die fachliche Qualifikation der Lehrkräfte.

Unsere Länderanalyse in Kapitel 3 sowie weitere Forschungsdaten zeigen, dass unterhalb der Systemebene vor allem den das Klassenzimmer umgebenden Kontexten, Strukturen und Organisationen auf der lokalen bis zur regionalen Ebene eine herausragende Bedeutung für eine erfolgreiche Implementation standardbasierter Reformen zukommt. Im Zentrum systemischer Initiativen in Bezug auf die Umsetzung von Bildungsstandards muss die signifikante Verbesserung der Qualität von Curricula und Unterricht stehen, womit folgende Elemente zentral werden:

- (1) klare, abgestufte Leistungserwartungen an *alle* Schülerinnen und Schüler;
- (2) gegenseitige Abstimmung von Bildungsstandards, Lehrplänen, Kerncurricula, Unterrichtsmaterial/Schulbüchern, Lehreraus- und -weiterbildung, Schulevaluationen und Tests;
- (3) Leistungsmessungen bei den Schülerinnen und Schülern;
- (4) Professionalisierung der Lehrkräfte;
- (5) ein breites Standard- und Accountability-Verständnis, das sich nicht nur auf die Outputseite, sondern auch auf die Inputs (Standards für die Ressourcen und Rahmenbedingungen) und auf die Lehr-Lern-Prozesse (den Unterricht) bezieht.

Zu den Problemen einer systemischen Strategie gehört u.a., dass diese infolge der hohen Komplexität und Vernetztheit der Prozesse, der wahrgenommenen Top-down- oder Außensteuerung sowie aufgrund der Involvierung unterschiedlicher Interessengruppen und Kräfte bei den Akteuren an der Basis zu Überforderung, Überladung und Frustration führen kann. Mit Spillane/Reiser/Reimer (2002, S. 404) lässt sich sagen: „Implementing agents encounter policy in a complex web of organizational structure, professional affiliations, social networks, and traditions.“ Die Lehrpersonen tendieren dann dazu, auf der Basis ihres Selbstverständnisses als pädagogische Akteure ihr Kerngeschäft, den Unterricht, von formalen Strukturen und komplexen, diffus wahrgenommenen Ansprüchen von außen zu entkoppeln. Rein systemische Betrachtungsweisen neigen zu Disseminationsstrategien: Programme, Papiere, Verfahren und Regelungen werden aus der Sicht ihrer Absender scheinbar effizient kommuniziert und deren „Adaption“ wird in kurzen Zeiträumen erwartet. Solche Strategien vernachlässigen häufig die Innensichten der Akteure und kümmern sich nicht oder zu wenig um die Mikrobereiche einer Reform, also die ganz konkreten handlungsbezogenen Probleme und die Arbeitskontexte. Auch werden die Voraussetzungen der Lehrpersonen und der Zeitbedarf, den selbst kleine Schritte benötigen, häufig falsch eingeschätzt.

Zusammenfassend gesagt: Das *politische* Projekt „Bildungsstandards“ und seine strategische Makrosteuerung muss als Rahmen aufgefasst werden, welcher der pädagogischen Ausgestaltung durch die hauptverantwortlichen Akteure an der Basis bedarf. Dies sind die Lehrpersonen, die letztlich nicht durch Druck, sondern vor allem durch Motivbildung, Überzeugungsarbeit, Entgegenbringung von Vertrauen, positive Anreize, Förderung und Ressourceneinsatz gewonnen werden müssen – dies unter der Voraussetzung einer Pflicht zur Rechenschaftslegung, die sowohl Tests als auch externe Evaluationen umfasst.

Die Akteur-Perspektive von „Accountability“, Innovation, Steuerung: Professionelles Lernen, Personal- und Schulentwicklung

Ob eine Bildungsreform ihre beabsichtigte Wirksamkeit entfaltet, entscheiden letztendlich die Akteure an der Basis. Wie die bereits erwähnte Langzeitstudie von Ekholm (1997) zeigt, gehen von einer bloßen Verwaltungsreform, die auf eine Delegation der Verantwortung von staatlichen Instanzen auf die einzelnen Schulen abzielt, nur geringe Impulse für pädagogische Schulentwicklung aus. Damit Lehrpersonen Adressaten und Agenten von Veränderungen werden, wollen sie von der Fruchtbarkeit neuer Ansätze überzeugt werden. Ohne Erfahrung eines Mehrwerts der Arbeit mit Bildungsstandards für die tägliche Unterrichtspraxis wird sich die Profession kaum dauerhaft auf neue Verfahren und Formen der Qualitätssicherung einlassen. Zudem kann ein verändertes Lehrerhandeln nicht einfach verordnet werden; die Lehrpersonen selbst müssen den gewünschten Wechsel vollziehen. Eine Fehlannahme wäre, wenn davon ausgegangen würde, dass erwünschte Veränderungen sich vielleicht nicht sofort, aber doch mit der Zeit von selber einstellen, wenn die entsprechenden Regelungen akkreditiert und publiziert sind, allenfalls begleitet von kurzer – „effizienter“ – Einführungs- und Informationsarbeit. Veränderungen in der Schul- und Unterrichtspraxis setzen schulbezogen-organisatorische, letztlich jedoch immer individuelle Veränderungen – und damit Lernen – voraus.

Mit den Worten der *American Federation of Teachers* (1995) lässt sich die Optik der Profession so festhalten:

„The nation can adopt rigorous standards, set forth a visionary scenario, compile the best research about how students learn, change the nature of textbooks and assessment, promote teaching strategies that have been successful with a wide range of students, and change all the other elements involved in systemic

reform. But, unless the classroom teacher understands and is committed to the plan and knows how to make it happen, the dream will come naught“ (S. 1f).

Aus der *Innensicht der Akteure* (Lehrpersonen, Schulleitungen, lokale Schulaufsichtsbehörden) bedeutet die Aneignung neuer Ziele und die damit verbundene Veränderung von Einstellungen und des Handelns einen komplexen und subjektiv bedeutsamen Vorgang des professionellen *Neu- und Umlernens*. Dass der Fokus auf die lernenden Akteure zu richten ist, legen empirische Befunde aus verschiedenen Forschungsperspektiven nahe, zum Lehrerlernen, zur professionellen Entwicklung von Lehrpersonen, zu Schulen als lernenden Organisationen, zum „Change-Management“, zur Schulentwicklung etc. (vgl. etwa Fullan 1991; Desimone 2002; Spillane/Reiser/Reimer 2002; Huberman/Miles 1984; Rolff 1993). Nach diesen Befunden genügt es nicht, den Lehrkräften ein Verständnis für die pädagogische und steuerungspolitische Idee von Bildungsstandards zu vermitteln. Entscheidend ist, dass diese diese Veränderungen selber wollen, also ihre individuelle und kollektive Veränderungsbereitschaft, und dass sie in die Entwicklungs-, Entscheidungs- und Umsetzungsprozesse aktiv, unter Rekurs auf ihre Erfahrungen und ihr Vorwissen, einbezogen werden. Das heißt, weder dürfen die Lehrkräfte als selbstständig Denkende und verantwortlich Handelnde nicht genug ernst genommen noch ihr mitgebrachtes Wissen und ihre Überzeugungen übergangen werden.

Da von einer Profession individuell und kollektiv über lange Zeit aufgebaute Handlungs- und Denktraditionen sich nicht einfach abstreifen lassen, werden Reformimpulse immer vor dem Hintergrund von durch Biografie, Ausbildung und Erfahrungen geprägten pädagogischen Mentalitäten und Weltansichten assimiliert und bewertet. Eine wichtige Rolle kommt dabei, sowohl was die Problemwahrnehmungen als auch was die Bereitschaften und die Möglichkeiten zu Veränderungen anlangt, den *Einzelschulen* als pädagogischen Handlungseinheiten zu (Fend 1986; Steffens/Bargel 1993). Von der Wahrnehmung systemweit verbreiteter Innovationsimpulse bis zu deren Akzeptanz an lokalen Einzelschulen, ihrer Umsetzung im unterrichtlichen Handeln bis hin zur Entfaltung der Wirksamkeit auf der Ebene des Schülerlernens ist es in der Regel ein sehr weiter Weg. Eine mittlerweile gut abgestützte Forschung zum berufsbiografischen Lernen von Lehrpersonen (vgl. für einen Überblick etwa Richardson/Placier 2001; Berliner 1986; Bransford/Brown/Cocking 1999; Munby/Russell/Martin 2001), aber auch zum organisationalen Lernen von Schulen (vgl. etwa Rolff 1993; Altrichter/Schley/Schratz 1998) hat zusammen mit den Erfahrungen aus schulnahen Reformprojekten zur Identifikation einer Reihe von Prinzipien und Bedingungen geführt, die für eine nachhaltige, den Unterricht

und das Lernen von Schülerinnen und Schülern erreichende pädagogische Qualitätsentwicklung als zentral gelten können.

Ausgehend von Haenisch (1994, S. 4ff.) und unter Einbezug von Erkenntnissen aus der Innovations- und Schulentwicklungsforschung (vgl. Huberman/Miles 1984; Fullan 1991; Rolff 1993; Altrichter/Posch 1996; Holtappels 2005) lassen sich aus der Sicht der Profession folgende Gruppen von *Bedingungen und Schlüsselfaktoren für eine wirksame Innovationssteuerung* unterscheiden, welche wir auch für die Einführung von Bildungsstandards für zentral halten:

(1) Bedingungen, die die Innovation selbst betreffen

- Positiv wahrgenommene Ziele und Inhaltsqualität, Akzeptieren der Vision auf dem Hintergrund eines wahrgenommenen Bedürfnisses nach Veränderung;
- Wahrnehmung eines Mehrwerts gegenüber dem vorangehenden Zustand;
- keine unmöglich erscheinenden Umwälzungen;
- anknüpfen an bisherige Erfahrungen;
- Ziele und Hintergründe verständlich darstellen;
- Spielräume für eigene Ausgestaltungen lassen;
- Einflussnahme der Betroffenen schon während des Entwicklungsprozesses;
- Innovationen über die Fächer „transportieren“;
- günstige Zeitpunkte für Innovationen nutzen.

(2) Bedingungen, die sich auf die individuellen Lehrpersonen beziehen

- Gelegenheit, sich zureichend mit den Reformanliegen beschäftigen zu können;
- Beherrschung der Neuerungen;
- Nutzenwahrnehmung für den eigenen Unterricht und das Schülerlernen;
- wahrgenommene Praktikabilität, Qualität und kontextuelle Passung der Innovation;
- Erlebte Wertschätzung im Einbezug von bisheriger Erfahrung und Reflexion;
- Vermittlung von Erfolgserlebnissen, Vertrauen und Zuversicht; Erfahrung von Anerkennung und Ermutigung;
- Vermeidung von Druck und Anordnung, aber Schaffung von Verbindlichkeiten.

(3) Bedingungen, die im Zusammenhang mit dem Arbeitsfeld der einzelnen Schule stehen

- Partizipation an gemeinsamen Situationsanalysen und Mitsteuerung der Reform unter Bedingungen von Teilautonomie und Spielräumen der einzel-schulischen Gestaltung;

- Möglichkeiten zu schulinterner und schulübergreifender Entwicklungsarbeit;
- effektive Leitung, adäquate Impulsgebung und aktive Unterstützung durch die Schulleitung und/oder durch Steuergruppen;
- Vorhandensein von Kompetenzen für systematische Entwicklungsarbeit im Hinblick auf Konsensbildung;
- Stabile Infrastruktur an schulinterner und schulübergreifender Kooperation und Kommunikation;
- Entwicklung eines positiven professionellen Lern- Fortbildungsklimas;
- Teamkultur und -kompetenzen, Wahrnehmung von Toleranz im Kollegium, Verständnis für Opposition; Qualität der kollegialen Beziehungen.

(4) Bedingungen, die den regulativen Rahmen und die unterstützenden Kontexte der Reform betreffen

- Bisherige (positive/negative) Erfahrung mit Reformen;
- Qualität der innovationsbezogenen Angebote der Lehrerfortbildung;
- Verfügbarkeit/Abrufbarkeit von Werkzeugen und Methoden der Umsetzung und des individuellen und organisationalen Lernens;
- Bereitstellung ausreichender sozialer und materieller Ressourcen;
- Abstimmung des regulativen Rahmens für Schule vor dem Hintergrund der neuen Ziele und Anforderungen;
- Möglichkeit, offizielle Zeit in die neuen Aufgaben zu investieren, und andere Anreize;
- Unterstützung und Ermutigung durch Schulaufsicht und zentrale Reformträger;
- Qualität der Beziehungen zwischen lokalen und zentralen Akteuren der Innovation;
- Einbezug der Eltern, vor allem unter der Perspektive, dass diese durch ihre Einstellungen und ihre häusliche Unterstützung einen wesentlichen Beitrag leisten zur Fähigkeit ihrer Kinder, die schulischen Angebote optimal zu nutzen.

(5) Bedingungen, die sich auf die konkrete Umsetzung der Innovationen beziehen

- Für die Adressaten durchsichtige und verständlich kommunizierte Reformstrategie;
- nicht zu viel auf einmal, Overload durch zu viele parallele Reformprojekte vermeiden;
- ausreichend Zeit zum Ausprobieren und Reflektieren;
- Prozessevaluation, positive Fehler- und Feedbackkultur.

Sollen Schulen und Lehrkräfte an der Basis für die Einführung von Bildungsstandards und Tests, und insbesondere für einen produktiven Umgang mit diesen gewonnen werden, so halten wir unter den erwähnten Bedingungsfaktoren deren zwei für besonders wichtig: Die Einrichtung von *stärker autonomen Schulen* sowie Qualitätsverbesserungen bei den Verfahren und Werkzeugen der *Lehrerfortbildung*. Die Schaffung beider Voraussetzungen, deren zentrale Wichtigkeit nachfolgend nochmals herausgestellt wird, sollte schrittweise angestrebt werden und erfordert einen Strategiewechsel.

Schulautonomie

Damit sich Schulen verantwortlich fühlen können für Lernergebnisse, die ihnen auf der Basis von Tests sowie als Ergebnisse von Evaluationen rückgemeldet werden, müssen sie über entsprechende Rahmenbedingungen verfügen, innerhalb derer sie ihre Angebote adaptiv für die Lernenden und nach Maßgabe ihres eigenschulischen Profils und Kontextes gestalten können. Das heißt, eine an einem erweiterten Verständnis von Educational Governance sich orientierende Philosophie einer „neuen Rechenschaftslegung“, welche die Überprüfung des Erreichens von Standards durch ein Monitoringsystem und externe Schulevaluationen umfasst, verlangt die Koppelung von verbindlichen Zielerwartungen und Ertragsorientierung mit der Gewährung von schulorganisatorischen, curricularen und pädagogischen *Spielräumen von Autonomie*, verstanden als erweiterte Möglichkeiten, die angestrebten Ziele auf unterschiedlichen Wegen und im Rahmen eines eigenen pädagogischen Profils zu erreichen. Dazu gehören der flexible Einsatz der zeitlichen und materiellen Ressourcen, das eigene Curriculum oder auch die Personalentwicklung.

Eine dergestalt schulische (Teil-)Autonomie steht in keiner Weise im Widerspruch zu einer Zielsteuerung über effektive und transparente Formen der Leistungsüberprüfung. Andererseits bedeutet die Einführung von Bildungsstandards unter Bedingungen verstärkter Eigenverantwortung von Schulen *keine* Standardisierung des Unterrichts. Standards wollen „den Unterricht gerade nicht in ein enges Korsett zwingen, sondern den Schulen inhaltliche und pädagogische Freiräume geben“ (Klieme/Avenarius/Blum et al. 2003, S. 15). Fachstandards wollen Lehrpersonen mehr Möglichkeiten geben als bisher, ihr zielerreichendes Handeln auf unterschiedliche Lernvoraussetzungen von Schülerinnen und Schülern einzustellen.

Eine in einzelschulischer Mitverantwortung durchgeführte Standardsicherung soll den Schulen überdies dabei helfen, ihre seit einigen Jahren gefährdete, zum Teil bereits an die Wirtschaft verlorene Autonomie in Bezug auf Zertifizie-

rung, Selektion und allgemein die Beurteilung von Schülerleistungen zurückzuerlangen. Den Schulen dürfte es zum Vorteil gereichen, wenn es ihnen gelingt, die Diskurshoheit über Zielerreichung und Standards zu behalten und die Evaluation erlangter schulischer Leistungs- und Kompetenzniveaus nicht den „Multichecks“ der Wirtschaft bzw. dem außerschulischen Feld zu überlassen.

Die Schulen müssen echte Formen der Autonomie ausbilden können. Das bedeutet auf der Ebene der nationalen und regionalen Politik einen Rückbau staatlicher Detailsteuerung bzw. eine Lockerung der zum Teil hohen Regelungsdichte und somit eine *Reform des Verwaltungshandelns*, verbunden mit der Allokation von Ressourcen bei den Gemeinden und Einzelschulen. Pädagogische Selbstverantwortung steht auf schwankendem Grund, wenn sie weder rechtlich abgesichert noch durch finanzielle Eigenständigkeit untermauert ist. Vor allem in Deutschland, wo Schulen in ein dichtes Netz von zentral angesiedelten Verwaltungskontrollen eingebunden sind, müssen ihnen künftig in einem höheren Maße als bisher Kompetenzen zur Gestaltung von Stundentafeln, Stundenplänen und der Schulorganisation, zur Ausgestaltung der Schulleitung, zur Auswahl des Personals, Freiheiten in der Stoffwahl, der Feingliederung und Ausformung des Curriculums in Richtung eines schuleigenen curricularen Profils (bei gegebener Verbindlichkeit von Standards) und in der Gestaltung von individuellen Förderangeboten, aber auch des finanziellen Haushalts, übertragen werden. Wie Avenarius/Ditton/Döbert et al. (2003) im ersten Bildungsbericht für Deutschland in Bezug auf den Ist-Zustand festhalten, besteht dazu vor allem in Deutschland erheblicher Handlungsbedarf (bei gleichzeitiger Feststellung, dass gewisse Veränderungen sich in mehreren Bundesländern abzeichnen):

„Die große Bedeutung, die dem Thema ‚Schulautonomie‘ in der Schulgesetzgebung, in bildungspolitischen Debatten, aber auch in der Bildungsforschung beigemessen wird, darf nicht darüber hinwegtäuschen, dass deutsche Schulen im Vergleich zu Schulen in andern OECD-Mitgliedstaaten über wenig Selbständigkeit zu verfügen scheinen.[...] Nach dem in PISA 2000 entwickelten Index für ‚Schulautonomie‘ befindet sich Deutschland unter 26 Staaten weit abgeschlagen im unteren Drittel.[...] Besonders gering ist danach der Einfluss deutscher Schulen auf die Einstellung und Entlassung von Lehrkräften, die Festsetzung des Schulbudgets und die von ihnen angebotenen Fächer“ (S. 157).

Welche empirisch nachweisbaren Wirkungen auf Unterricht und Schülerlernen von Schulen ausgehen, die hinsichtlich Personalentwicklung, Budget- und pädagogischer Programmgestaltung über zusätzliche Freiheitsgrade verfügen, ist

zwar trotz langjähriger Debatte und Forschung zum Teil widersprüchlich und immer noch schmal und lückenhaft belegt (vgl. Pfeiffer 2004). Dies hängt u.a. auch damit zusammen, dass – wie Leithwood/Menzies (1998) in einer Metaanalyse feststellen – nur in wenigen (11 von 83) der vorhandenen amerikanischen Studien zu schulischer Autonomie Schülerwahrnehmungen oder Schülerleistungen überhaupt einbezogen wurden. In etwas mehr als 40% der Studien konnten jedoch positive Wirkungen auf die Lehrpersonen und auf die Gemeinschaft der Schule festgestellt werden. Auch Dempster (2000) nennt mögliche lernförderliche Wirkungen, welche durch koordinierte Planung und verbesserte Kommunikation entstünden, verneint aber im Anschluss an seine Metaanalyse englischer, australischer und neuseeländischer Schulen die Behauptung, dass teilautonome Strukturen *per se* zu verbesserten Schülerleistungen führen würden.

Im Rahmen von PISA 2000 wurde anhand von zwölf Indikatoren (unter anderem Einstellung von Lehrkräften, Festlegung des Schulbudgets, Festlegung von Kriterien für die Schülerbeurteilung und Bestimmung des Lehrstoffes) festgehalten, welchen durchschnittlichen Grad an Teilautonomie die teilnehmenden Länder erreichten. Dabei konnte festgestellt werden, dass diejenigen Länder, welche ein höheres Maß an Autonomie hinsichtlich des Fächerangebots aufweisen, höhere durchschnittliche Leseleistungen erbrachten. Selbstverständlich lassen sich daraus keine Ursache-Wirkungszusammenhänge ableiten, da die höheren Schülerleistungen auch auf andere Weise beeinflusst sein könnten (Pfeiffer 2004).

In eine ähnliche Richtung weisen die Ergebnisse einer schwedischen Untersuchung, mit welcher der Übergang von zentraler Steuerung zu schulischer Autonomie im Rahmen eines Längsschnitts mit drei Erhebungszeitpunkten über einen Zeitraum von 25 Jahren hinweg beobachtet werden konnte (Ekholm 1997). Dabei wurde der Fokus auf die Qualität der Arbeit der Schulen und auf die Qualität der Lernergebnisse gerichtet. Die Ergebnisse zeigen, dass die Unterschiede zwischen den Schulen über den untersuchten Zeitraum hinweg stabil blieben. Nach Ansicht des Autors ist es weniger das Steuerungsmodell, das Anlass für nachhaltige Veränderungen bietet, als vielmehr die Mentalität der beteiligten Lehrkräfte, welche die Qualität der Schule beeinflusst und damit die Grundlage für deren Arbeitskultur bildet. Eine wichtige weiterführende Frage wäre, wovon die beobachteten Mentalitätsunterschiede ihrerseits abhängig sind und welche Rolle lokale Theorien und Kulturen bei ihrer Stabilisierung spielen. In der Schweiz konzentrierte sich die Realisierung von schulischer Teilautonomie bisher auf folgende Bereiche: Gestaltungsspielraum, Schulprofil, Schulleitung, Schulaufsicht und Qualitätsmanagement. In einer Studie wurden die dies-

bezüglichen Reformprojekte verschiedener Kantone inhaltsanalytisch untersucht (Maag Merki/Büeler 2002). Dabei zeigte sich, dass die Lehrpersonen gesamthaft über projektbedingte Belastungen klagten, dass die Schulleitungen eine entscheidende Rolle bei der Gestaltung und Reorganisation der Schulen spielen, dass die Schülerinnen und Schüler eine recht hohe Zufriedenheit bekunden, dass der Informationsstand der Eltern in Bezug auf das Projekt relativ tief ist und dass die Schulbehörden den entstandenen Kompetenzverschiebungen insgesamt positiv gegenüberstehen. Hinsichtlich Wirkungen auf den Unterricht werden drei von vier laufenden Versuchen eher negativ bilanziert, und allfällige positive Wirkungen werden dem Einfluss anderer, die konkrete Unterrichtsarbeit betreffender Projekte zugeschrieben.¹⁴⁸

Insgesamt scheint eine größere Schulautonomie am ehesten dann positive Wirkungen zu haben, wenn die strukturelle Reform von einer (lern-)kulturellen Reform begleitet wird, die auf der Basis funktionierender Unterstützungssysteme ihrerseits in einen Kontext der langfristigen Personal- und Schulentwicklung eingebettet ist. Das heißt, dass nicht das Steuerungselement „Schulautonomie“ per se zu verbesserten Wirkungen führt, sondern vielmehr die Art und Weise, wie die betroffenen Lehrpersonen und Schulleitungen mit den neuen Handlungsspielräumen umgehen und sie zur Basis einer veränderten Arbeitskultur machen.

Für die Einführung von Bildungsstandards ist zu erwarten, dass vor allem starke, innerlich gefestigte, über einen tragenden Konsens und ein Ethos von Schul- und Unterrichtsqualität sowie über stabile Kooperations- und Selbstlernkulturen verfügende Schulen – als „lernende Organisationen“ – in der Lage sein werden, die Vorteile einer größeren Autonomie für das berufliche Handeln und für ihre Weiterentwicklung überhaupt zu erkennen und davon Gebrauch zu machen. Nur solche Schulen werden beispielsweise fähig sein, Leistungsrückmeldungen aus Tests für ihre produktive Weiterentwicklung zu nutzen.

Entscheidend bei alledem ist jedoch nicht eine neue Organisation, sondern der wahrgenommene Erfolg der Maßnahmen. Ohne sichtbaren Fortschritt werden die Schulen zu ihren bewährten Verfahren zurückkehren. In der Forschung ist immer wieder darauf hingewiesen worden, dass der Sinn von Wandel in einer sehr komplexen Situation mit der Erfahrung des Vorankommens ver-

¹⁴⁸ Auch in einer neueren Analyse der Schweizer PISA-Daten konnten keine Zusammenhänge zwischen schulischer Teilautonomie und Schülerleistungen in Mathematik nachgewiesen werden (Zahner Rossier/Berweger/Brühwiler et al. 2004). „Teilautonomie“ bedeutet allerdings in verschiedenen Schweizer Kantonen durchaus Unterschiedliches.

knüpft ist. Ohne Sense-Making gelingt im Bildungsbereich keine Implementation. Da Lehrpersonen die Frage, ob „mehr Autonomie“ wünschbar sei, deshalb nicht zuletzt aus der Perspektive betrachten, welche Vorteile (Nutzen für den Unterricht, Entlastung, weniger Unsicherheit usw.) ihnen diese in der alltäglichen Arbeit als Einzelkämpfer bietet (Tillmann 2001), ist es erforderlich, sowohl Schulen als kollektive Handlungsträger als auch individuelle Lehrkräfte in ihren autonomiebezogenen Entwicklungsbemühungen nicht allein zu lassen, sondern mittels Fortbildung mit neueren Konzepten und Architekturen von Schulorganisation in Kontakt zu bringen. Veränderungen werden sich dabei nicht sofort zeigen, Umwege und Rückschläge werden in Kauf zu nehmen sowie Beharrlichkeit und Ausdauer nötig sein, bis sich die Ergebnisse von Schulentwicklungsprozessen im unterrichtlichen Kern und in nachhaltigen Lernwirkungen bei den Schülerinnen und Schülern niederschlagen.

In Deutschland berührt eine neue, auf Selbstverantwortung und Vertrauen setzende Philosophie der Steuerung zudem die Wurzeln des Selbstverständnisses von Schulen und Lehrpersonen. Das Gewähren von Autonomie bedeutet sowohl angebotsseitig (Verwaltungshandeln, Politik) als auch auf der Seite ihrer Nutzer (Lehrerinnen und Lehrer) einen Kulturwandel – und damit viel Gewöhnung und Lernen.

Qualität der Fortbildung

Weder lassen sich Sinn für Verantwortung für den Bildungserfolg und eine damit verbundene, erneuerte Lern- und Arbeitskultur bei Lehrpersonen (und mittelbar bei Schülerinnen und Schülern) „top down“ auf dem Weg schulpolitischer Erlasse verordnen noch entsteht eine neue Schul- und Lernqualität allein oder vorwiegend aus selbstorganisierter und lokaler Schulentwicklung. Die in ihrem Anspruch keinesfalls zu unterschätzenden Reformziele der Einführung von Bildungsstandards bedürfen auf allen Ebenen des qualifizierten Ideen-Inputs von außen und der professionellen, handlungsorientierten Unterstützung durch neue Formen der Fort- und Weiterbildung.

Unter der Annahme, dass die erfolgreiche Einführung von Bildungsstandards eine hohe Schulqualität voraussetzt und diese große Anstrengungen bei der Personal- und Schulentwicklung bedingt, bilden innovative Methoden und Formen der Fortbildung einen weiteren Pfeiler des Projekts. Eine zentrale Forderung an eine wirksame Fortbildung von Lehrkräften ist, dass ihre Verfahren die professionell geformten Mentalitäten, Überzeugungen, Gewohnheiten und eingeschliffenen Handlungsrouninen auch tatsächlich treffen und erreichen – und diese bearbeitbar werden. Umlernen im Bereich implizit gewordenen Wis-

sens, habitierter subjektiver Theorien und Handlungsmuster verlangt, wie die Forschung klar zeigt, nach anderen als bloß akademischen Wissensvermittlungs- und Schulungsstrategien. Gesucht sind innovative Formen der Lehrerfort- und Weiterbildung, die einerseits prozessunterstützend und handlungsorientiert angelegt sind (indem sie sich z.B. an den Erkenntnissen zum situierten Lernen orientieren) und andererseits bewährte schulorganisatorische Konzepte und unterrichtliche Qualitätsvorstellungen an die Schulen und Lehrkräfte herantragen. Der Fokus muss sich dabei auf das Wechselspiel von individueller Lehrerebene und kooperativer Schulebene richten. Insbesondere jene Aspekte einer pädagogischen Schul- und Lernkultur, welche die Veränderung kollektiver und kollegialer Praxis im Auge haben, ergeben sich nicht einfach als Summe individueller Lernprozesse.

Alle vier von uns untersuchten Länder sehen in der Innovation der Formen und Werkzeuge der Fortbildung ein vorrangiges Handlungsfeld. Die ausführlich dargestellte amerikanische Kritik an den konventionellen, meist vorwiegend akademischen Formen der Lehrerfortbildung lässt sich mit Hawley/Valli (1999) pointiert wie folgt zusammenfassen: Diese Formen sind „shallow and fragmented; private, individual activity; brief, often one-shot sessions: offer unrelated topics, expect passive teacher-listeners; atheoretical; expect quick visible results“ (ebd., S. 134). Künftig dürften die Angebote zudem weniger individuell, als vielmehr auf ganze Schulen und ihre Bedürfnisse zugeschnitten werden. Auf der Basis einer umfangreichen Literatur zu Innovation und schulnahe Lernen legen die Autoren folgende Prinzipien für eine effektive professionelle Fortbildung von Lehrpersonen nahe (ebd.):

- (1) *Ausgerichtet an Zielen, Standards und Schülerleistungen:* Professionelle Schulentwicklung muss von der Analyse von Differenzen zwischen gesetzten Zielen und effektiven Schülerleistungen ausgehen. Lehrerinnen und Lehrer müssen – im Sinne eines schülerzentrierten Ansatzes – in der Lage sein, die Prozesse und Strategien der Lernenden in der Perspektive von zu erreichenden Zielen analysieren zu können.
- (2) *Orientiert an von den Lehrpersonen wahrgenommenen und anerkannten Bedürfnissen und Defiziten:* Lehrpersonen sollen erkennen, was sie in Bezug auf ihr professionelles Handeln dazulernen müssen, und sie sollen wissen, auf welche Weise sie dazu entsprechende Unterstützung erhalten. Dieses Prinzip verweist auf die Bedeutung von Schulleitungen. Wo professionelles Lernen von Lehrpersonen integraler Bestandteil einer schulischen Kultur ist, dürften entsprechende diagnostische Prozesse gefördert werden.

- (3) *Schulbasiert und in die tägliche Berufsarbeit eingebunden:* Fortbildung und professionelles Lernen sollen in die Arbeitsabläufe der jeweiligen Schule eingebunden oder auf diese bezogen sein. Effektives Lernen der Lehrpersonen wird vor allem dann produktiv, wenn authentische und dringliche Probleme als Ausgangspunkt genommen werden und die professionelle Weiterbildung im Rahmen der regulären Arbeitszeit erfolgen kann.
- (4) *Kooperative Problemlösung und Lehrerarbeit in Teams und Netzwerken:* Damit Schulen als Ganze sich erfolgreich weiterentwickeln, muss eine professionelle Weiterbildung von Lehrkräften neben deren individuellen Bedürfnissen auch gemeinschaftliche Problemlöseprozesse im Blick behalten. Im Hinblick auf die Einführung von Bildungsstandards sind deshalb Ansätze zu favorisieren, mit deren Hilfe eine gemeinsame Sprache und ein von möglichst allen geteiltes, konkretes Verständnis in Bezug auf Schul- und Unterrichtsqualität gefördert werden.
- (5) *Kontinuierlich und durch externe Ressourcen und Expertise kompetent unterstützt:* Um zu verhindern, dass Schulentwicklungsprozesse im Sande verlaufen, ist weiterführendes und kontinuierliches Lernen notwendig. Dabei sind auch Außenperspektiven fruchtbar, weil nur dadurch neue Sichtweisen und Ideen ins System der Einzelschule gelangen.
- (6) *Einbezug von aus multiplen Evaluationsquellen stammenden Rückmeldungen zu Schülerleistungen und Lehrerhandeln:* Datengestützte Evaluationen, die sich auf multiple Erfolgs- und Wirkungskriterien abstützen, sollten für die Beteiligten produktiv erlebte Rückmeldungen erzeugen. Günstig ist eine Kombination interner und externer Evaluationen. Wichtig ist dabei, dass Bewertungen des beruflichen Handelns für die Betroffenen keinen bedrohlichen Charakter aufweisen, sondern zur beruflichen Weiterentwicklung beizutragen vermögen.
- (7) *Ergebnisbezogenes und prozessnahes theoretisches Verständnis der im Unterricht zu erwerbenden Kompetenzen:* Lehrpersonen sollen ein theoretisches Verständnis der von den Schülerinnen und Schülern zu erwerbenden Wissensbestände und Fertigkeiten und der dazu hinführenden Prozesse entwickeln können. Dabei geht es neben allgemein pädagogischem und inhaltspezifischem Wissen um den Aufbau von fachdidaktischem Lernprozesswissen. Im Hinblick auf die Einführung von Bildungsstandards muss dieses Wissen an die tägliche Unterrichtstätigkeit gekoppelt sein, um eine tatsächliche Implementation zu ermöglichen.
- (8) *Teil eines langfristig angelegten, sowohl mit Lernfortschritten als auch mit Rückschlägen rechnenden Veränderungsprozesses unter fachlicher Unterstützung:* Lernprozesse können durch viele Faktoren behindert oder auch erleichtert werden. Es

muss berücksichtigt werden, dass Lehrpersonen das, was sie lernen, im Unterricht auf seine Praxistauglichkeit prüfen können. Um zu erkennen, ob bzw. unter welchen Bedingungen neue Unterrichtspraktiken funktionieren, benötigen sie sowohl Zeit als auch fachlich kompetente Unterstützung.

Neueren Ansätzen zum professionellen Lernen ist als Erkenntnis gemeinsam, dass Bemühungen um nachhaltiges Umlernen von Lehrpersonen in Bereichen, die erfahrungsgemäß gegenüber Korrekturen und Restrukturierung besonders resistent sind, nur erfolgreich sind, wenn sich diese vermehrt an Settings des situierten, handlungsbezogenen Lernens orientieren und insgesamt ein breites Repertoire von handlungs- und reflexionsorientierten Verfahren und Methoden eingesetzt wird; ein Repertoire, das über die akademische Vermittlung von (Forschungs-)Wissen hinausgeht (Wang/Odell 2002) und das sich auf das Kerngeschäft bzw. die Primärprozesse von Schule, d.h. den Unterricht und die daraus resultierende Lernqualität bei den Schülerinnen und Schülern richtet.

4.3 Ebenen und Akteure

Bildungssysteme sind vom Menschen geschaffene Regelwerke und konstituieren sich über verschiedene politische und funktionale Ebenen und Akteurgruppen, deren Existenz relativ konstant ist, deren Funktionen und Bedeutungen sowie Ausprägungen sich jedoch unterscheiden. Wer Standards im System verankern will, muss diese Ebenen berücksichtigen und den in ihnen spielenden sozialen Prozessen Rechnung tragen. Alle Ebenen und Gruppen sind gleichzeitig Akteure und werden betroffen sein von der Einführung neuer Modelle der Steuerung und Qualitätssicherung. Rollen, Funktionen und Bedeutungen von Behörden, Schulleitungen, Lehrkräften und nicht zuletzt von Eltern und Schülerinnen und Schülern werden sich mehr oder weniger stark verändern. Dies betrifft auch die Schulen als Handlungseinheiten, denen vermehrte Autonomie und somit erhöhte Verantwortung zugesprochen werden soll. Davon sind auch das Professionsbild und berufliche Ansprüche der Lehrerinnen und Lehrer betroffen. Auch den Schulleitungen wird künftig bei Innovationen eine wichtige Funktion zukommen.

In der neueren Theorie der Schule (vgl. Fend 2006) und in der Theorie der Schulentwicklung werden häufig drei Ebenen unterschieden, nämlich die *Mikroebene* des Unterrichts, die *Mesoebene* der einzelnen Schule und die *Makroebene* des Bildungssystems. Diese Unterscheidung wird von einer vertikalen Raumvor-

stellung geprägt. Es gibt in diesem Modell eine kleine, eine mittlere und eine große Einheit, die in einem normativen und pragmatischen Zusammenhang stehen und zwischen denen eine Art geregelter Austausch angenommen wird. Die Implementationsforschung hat, nicht nur im Bildungsbereich, gezeigt, dass auch horizontale sowie instrumentelle Faktoren den Verlauf von Innovationen bestimmen. Implementationen sind auf Prozesse der Verbreitung mittels tauglicher Handlungsressourcen und „Technologien“ angewiesen, die nie die gesamte Fläche auf dieselbe Weise erreichen. Dies gilt auch für Unterstützungssysteme, die nur mit Hilfe von Netzwerken und von leistungsfähigen Routinen und Werkzeugen ihre Wirksamkeit entfalten können.

Der Austausch zwischen den Ebenen kennt verschiedene Mechanismen und Kanäle der Kommunikation, die von Vorschriften in Papierform bis zu informellen Gesprächen zwischen Kollegen reichen. Eines ist ihnen gemeinsam: Sie sind sämtlich interpretationsabhängig. Aus diesem Grund sind in der Sozialpsychologie Modelle „interpretierender Systeme“ entwickelt worden (Daft/Weick 1984), die inzwischen auch in der Schulforschung angewendet werden (O’Day 2002). Hinter diesen Modellen stehen folgende Annahmen: Sozialer Austausch unter Bedingungen hoher Komplexität verlangt fortgesetzte Abstimmung und somit unablässige Deutungsarbeit der Akteure und ihrer Medien. Erforderlich sind Interpretationen sowohl in der Vertikalen als auch in der Horizontalen. Es gibt keine Meta-Interpretationen, die sich frei von anderen Interpretationen durchsetzen könnten, und somit auch keine privilegierte, oberste Interpretation oder Instanz der Reformsteuerung.

Unter diesen Voraussetzungen lassen sich folgende Ebenen der Implementation von Bildungsstandards unterscheiden:

- (1) das nationale Bildungssystem,
- (2) regionale Netzwerke,
- (3) einzelne Schulen,
- (4) Akteure,
- (5) Instrumente.

Bildungssysteme sind national organisiert; der heute verstärkt wahrgenommene internationale Vergleich und Wettbewerb setzt nationale Konfigurationen von Bildung voraus. Die Finanzierung der Systeme geschieht über das nationale Steueraufkommen, die politische Steuerung basiert auf zentralistischen und/oder föderativen Komponenten, bezieht sich aber immer auf ein nationales System; die Ausbildung der Lehrkräfte, die Lehrpläne und Lehrmittel, die Beschreibung der Ziele und Aufgaben folgen ebenfalls nationalen Vorgaben. Die

Internationalisierung ist demgegenüber mindestens im Bereich der obligatorischen Schule und bezogen auf den Alltag des Unterrichts schwach und marginal.

4.3.1 Nationale Ebene: Bund, Länder und Kantone

Die Diskussion und Entwicklung von Bildungsstandards folgt wohl einem internationalen Trend, weist aber im Quervergleich zwischen den nationalen Systemen deutliche Unterschiede auf. Bildungsstandards sind *nationale* Größen, die keinen internationalen Konsens voraussetzen. Was ähnlich ist, sind die Prinzipien, Instrumente und die Verfahren der Implementation, die aber je nur vor dem Hintergrund der nationalen Systeme eingesetzt werden können. Auch auf curricularer Ebene ist die Internationalisierung unterschiedlich ausgeprägt. Sie ist im Bereich des Elementarunterrichts oder der Mathematiklehrmittel erheblich größer als in den musischen, muttersprachlichen oder historischen Fächern. Und während der Englischunterricht inzwischen weitgehend normiert ist, gilt dies für andere Sprachfächer wie Französisch oder Italienisch wesentlich weniger.

Das Basisverständnis darüber, was als „Bildungsstandards“ bezeichnet werden soll, ist keineswegs dasselbe, sondern eher nur in den Absichten ähnlich, in den Ausführungen dagegen unterschiedlich. Unter den deutschsprachigen Ländern herrscht zum Beispiel Einigkeit in der Hinsicht, dass sich Bildungsstandards auf fachliche Kompetenzen beziehen und unterschieden werden müssen von inhaltsleeren „Schlüsselqualifikationen“ (Klieme/Avenarius/Blum et al. 2003, S. 134ff.). Andererseits sind etwa in der Schweiz deutliche Bestrebungen erkennbar, überfachlichen Kompetenzen ein deutliches Gewicht bei der Entwicklung von Standards zukommen zu lassen, nachdem versucht wurde, *Kompetenzen* in diesem Bereich begrifflich von „Schlüsselqualifikationen“ abzugrenzen (Grob/Maag Merki 2001). Dass überfachliche Kompetenzen in ihrer Bindung an den Unterricht und die Lernkultur ein zentrales Element der Schulentwicklung darstellen, ist unbestritten; wie weit sie sich wie fachliche Kompetenzen testen lassen, ist dagegen eine offene Frage.

Was nationale Differenzen anlangt, verlangen föderative Systeme wie Deutschland oder die Schweiz eine komplizierte Beschreibung. Während in Österreich oder Luxemburg nationale Ministerien die oberste Ebene der Steuerung darstellen, sind es in Deutschland und der Schweiz Konferenzen, in denen jeweils Mehrheiten gefunden werden müssen. In der Schweiz etwa bedurfte es der Annahme eines neuen Bildungsverfassungsartikels durch die Stimmberechtigten, um die Grundlage für eine interkantonale Vereinbarung zur Harmonisierung

der obligatorischen Schule zu schaffen (HarmoS). Hier bedarf es auch der Berücksichtigung regionaler und kantonaler Interessen in Bezug auf Kompetenz- und Rollenzuweisungen, um dem Gesamtprojekt die notwendige Unterstützung in einem Land mit mehreren Landessprachen zu sichern. In Österreich stellen ministeriale Entscheide einheitliche Vorgaben dar, in den föderativen Systemen ist dergleichen nur aufgrund von Abstimmungsprozessen möglich. Immerhin zeigen die beiden Projekte „nationale Bildungsstandards“ in Deutschland und „HarmoS“ in der Schweiz, dass nationale Vorgaben entweder angestrebt werden oder bereits beschlossen sind.

Auch in Österreich und Luxemburg heißt „Implementation von Bildungsstandards“ nicht, dass ministeriale Vorgaben mit Aussicht auf Erfolg einfach von der Makroebene auf die Mikroebene weitergegeben werden können. Auch hier stellen sich also Implementierungsprobleme, die den behördlichen Dienstweg überschreiten. Insofern stellt sich kein grundsätzlich anderes Problem, außer dass keine rivalisierenden Länderbehörden bestehen, die je eigene Wege der Implementation verfolgen können. Allerdings sind die Unterschiede zu beachten. In der Schweiz heißt Bildungsföderalismus etwas anderes als in Deutschland, so wie auch zentralistische Steuerungen in Österreich oder Luxemburg von solchen etwa in Frankreich zu unterscheiden sind. Vor allem sind die quantitativen Größenordnungen, das heißt die flächen- und Bevölkerungsmäßigen Einheiten, verschieden. Ein deutsches Bundesland ist kein Schweizer Kanton. Was in den verschiedenen Ländern als Region gilt, dürfte ebenfalls quantitativen und kulturellen Interpretationen unterliegen. Die Schweiz ist zudem vier-sprachig. Sodann entsprechen einige deutsche Bundesländer in ihrer Größe durchaus der ganzen Schweiz oder Österreich, Luxemburg demgegenüber entspricht einem städtischen Einzugsgebiet.

Die Größe der meisten Schweizer Kantone schließt aus, dass in sehr großem Umfang eigene Entwicklungsarbeit betrieben werden kann. Zumeist handelt es sich um die Anpassung der Instrumente und Verfahren an kantonale Gegebenheiten, wobei oft Beispiele übernommen werden. Das ist in den großen deutschen Bundesländern durchaus anders, und die laufende Föderalismusdiskussion verschärft den Trend zur eigenen Lösung oder mindestens zur Anpassung der nationalen Vorgaben an das eigene Verständnis. Ein zentrales Problem der Implementation von Bildungsstandards in Deutschland ist die Übersetzung der KMK-Standards in eigene Entwicklungsprogramme der Länder, die untereinander kaum abgestimmt sind. Eine „koordinierende Institution,“ wie sie die Expertise vorschlägt (Klieme/Avenarius/Blum et. al. 2003, S. 126), ist derzeit nicht in Sicht. Der Entscheid in der Föderalismusdiskussion

spricht zudem dagegen, mehr zu erwarten als die jetzigen Steuerungsmöglichkeiten der KMK.

Ein Beispiel für das Problem ist die Reaktion auf die KMK-Bildungsstandards im deutschen Bundesland Bayern. Hier wird betont, „dass die durch die KMK-Standards formulierten Lernziele in allen bayerischen Lehrplänen im Wesentlichen bereits abgedeckt werden.“¹⁴⁹ Die KMK-Standards werden von den bayerischen Lehrplänen unterschieden, in dem Sinne, als dass die Ersteren nur einen Teil des Curriculums abdecken, die Letzteren dagegen ein „pädagogisches Gesamtkonzept“ darstellen. Die Lehrpläne in Bayern:

- formulieren in ihren Lernzielen nicht nur überprüfbare Kompetenzen, sondern auch darüber hinausgehende Fähigkeiten und Wertorientierungen;
- beschreiben nicht nur Lernergebnisse, sondern auch wesentliche Lernprozesse vor dem Hintergrund didaktischer Prinzipien guten Unterrichts;
- machen Aussagen zur Sequenzierung der Ziele und Inhalte nach Jahrgangsstufen;
- machen schulartspezifische Unterschiede deutlich.

Die Bedeutung der Standards wird in einem „Perspektivenwechsel“ gesehen, der die Lehrkräfte anhält, den Blick nicht nur auf den Stoff des Unterrichts zu richten, sondern die Ergebnisse am Ende einer Jahrgangsstufe vor Augen zu haben. „Die Kompetenzentwicklung muss also gegenüber der Fachsystematik im Vordergrund stehen [...] Diesen Perspektivenwechsel, den die neueren bayerischen Lehrpläne schon vollzogen haben, im Bewusstsein aller Beteiligten zu verankern, ist die wesentliche Intention der Bildungsstandards“ (Bildungsstandards und bayerische Lehrpläne [o.J.]).

Von *nationalen* Bildungsstandards kann nur dann die Rede sein, wenn die Vorgaben bindenden Charakter haben. Dies gilt gleichermaßen für die Strategien der Implementation, sofern diese grundsätzlichen Charakter erhalten sollen. Über bestimmte Grundsatzfragen muss ein Konsens erreicht werden, wenn wirklich nationale Lösungen angestrebt werden sollen. Dies bedeutet nicht, dass die Entwicklung uniform oder linear vor sich gehen soll. Aber bestimmte Entschiede müssen einheitlich getroffen werden. Dies betrifft etwa die folgenden Fragen:

¹⁴⁹ Dieses Zitat geht zurück auf eine vergleichende Lehrplananalyse des Staatsinstituts für Schulqualität und Bildungsforschung in München (Staatsinstitut für Schulqualität und Bildungsforschung (ISB) 2005).

- Sind Regelstandards oder Mindeststandards verbindlich?
- Ersetzen oder ergänzen die Bildungsstandards die Lehrpläne?
- Werden nationale oder länderspezifische Tests eingesetzt?
- Sind die Tests zeitlich vernetzt und lassen sich damit längerfristige Trends messen?
- Dienen die Tests der Selektion und/oder der Förderung der Schülerinnen und Schüler?
- Werden Rückmeldesysteme aufgebaut, die die Schule, den Unterricht und die Akteure vor Ort erreichen?
- Sind Testergebnisse Teil der Schulevaluation?
- Werden die Ergebnisse veröffentlicht und wenn ja, in welcher Form?
- Werden mittelfristig Schulrankings angestrebt, die die Mittelzuweisung beeinflussen?
- Sind landesweite Unterstützungssysteme vorgesehen, die sowohl die Implementation von Bildungsstandards als auch deren weitere Ausgestaltung betreffen?

Ohne Beantwortung dieser Fragen ist es nicht möglich, zu (nationalen) und aufeinander abgestimmten Strategien zu gelangen, die tatsächlich das Gesamtsystem tangieren. Beispielsweise sind entscheidend, ob man sich an Regel- oder Mindeststandards orientiert, keineswegs triviale Auswirkungen auf den Stellenwert von Förderstrategien zu erwarten. Die deutsche Kultusministerkonferenz (KMK) hat sich für Regelstandards entschieden, die Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) dagegen strebt die Entwicklung von Mindeststandards an. Die deutschen Regelstandards sollen nach der bisher gültigen Lesart an den beiden Schnittstellen der Selektion überprüft werden, nämlich am Ende oder im Verlauf der 4. und am Ende der 9. bzw. 10. Klasse. Die schweizerischen Mindeststandards werden nach der Basis- bzw. Grundstufe (2. Klasse), an der Schnittstelle der ersten Selektion (6. Klasse) sowie beim Übergang in die Berufslehren oder die weiterführenden Schulen (9. Klasse) getestet. Die Daten sind mit klaren Förderabsichten verbunden, wobei die Frage der Verwendung auch für Selektionsentscheide noch ungeklärt ist.

Diese Unterschiede sind wichtig und verlangen nach klärenden Entscheiden. Wenn grundlegende Probleme zu lange in der Schwebe bleiben, ist das Projekt nationaler Bildungsstandards ebenso gefährdet wie bei Inkaufnahme zu grosser Unterschiede zwischen den Bundesländern oder den Kantonen. Auf der anderen Seite gibt es schwebende Probleme, die nicht gelöst werden, auch in zentralistischen Systemen. In Österreich werden Regelstandards und deren Überprüfung an den Schnittstellen der Selektion (4. und 8. Klasse) angestrebt, was un-

mittelbar die Frage aufwirft, welche Förderabsicht sich mit den Standards verbindet und wie diese realisiert werden soll.

Der politische und pädagogische Zweck der Standards muss klar bestimmt sein, wenn die Anstrengungen der Implementation Erfolg haben sollen. Zudem muss auf der Ebene des Gesamtsystems festgelegt werden, wie die Ziele überprüft werden und wer sie mit welchen Verfahren überprüft. Wenn Tests zum Einsatz kommen, sind valide Aufgaben, die fortlaufend entwickelt werden, bzw. ist eine strukturierte, vielfältig einsetzbare Itembank eine entscheidende Voraussetzung dafür. Es ist eine grundlegende Entscheidung, die möglichst rasch gefällt werden muss, ob wie in Schweden oder in England ein nationales Testbüro mit der Durchführung und Auswertung der Tests sowie mit der Rückmeldung der Daten beauftragt wird oder ob dies auf Länderebene geschehen soll. In der Schweiz wird es aller Voraussicht nach regionale Zusammenfassungen geben, die bereits heute in anderen Aufgabenfeldern der Bildung existieren. Es ist eine noch durchaus offene Frage, über welche institutionellen Vernetzungen sich die anstehenden Monitoring-, Evaluations- und Steuerungsaufgaben in den Auftragsländern dieser Expertise qualifiziert und auf für die Beteiligten transparente Weise werden lösen lassen.

Eine Schlüsselfrage ist auch, ob Daten aus Leistungstests Einfluss nehmen sollen auf die Prozesse der Qualitätssicherung in den Schulen. Konkret hieße dies etwa, externe Evaluationen einzelner Schulen mit Vergleichsdaten aus Leistungstests zu verknüpfen und so eine größere Objektivität der Qualitätseinschätzung zu erreichen. Rückmeldesysteme müssen die Akteure erreichen, aber wenn die Akteure mit Daten aus Leistungstests produktiv umgehen sollen, muss klar sein, dass damit keine persönlichen Risiken wie etwa Gehaltseinbußen verbunden sind. Vergleichsarbeiten, die in den meisten deutschen Bundesländern inzwischen üblich sind oder vor der Einführung stehen, sind nicht dasselbe wie Leistungstests. Nur bei der Einführung von Tests stellen sich Fragen der Rückmeldung oder der Datentransparenz über das Ergebnis hinaus. Auch hier sind politische Entscheide gefragt, wie im Übrigen auch für Schulrankings, die nicht einfach den regionalen Entwicklungen überlassen sein können.

Im Bundesland Bayern wird die operative Seite der KMK-Bildungsstandards für die Lehrkräfte so beschrieben: Standards

- lenken die Aufmerksamkeit noch stärker auf die Lernzielformulierungen in den Lehrplänen,
- konkretisieren die Angaben der Lehrpläne zum Grundwissen,

- betonen die Notwendigkeit der Förderung nachhaltigen, kumulativen Lernens, die sich auch in der Konzeption von Leistungserhebungen noch stärker widerspiegeln muss,
- schärfen den Blick für eine nach Kompetenzbereichen ausgewogene Zusammensetzung von Übungsaufgaben und Leistungserhebungen,
- geben konkrete Anregungen für neue Aufgabenformate, bei denen Anwendungs- und Handlungsorientierung im Vordergrund stehen,
- unterstreichen die Bedeutung der zentralen Vergleichsarbeiten (Orientierungsarbeiten, Jahrgangsstufentests) als objektive Diagnoseinstrumente, deren Ergebnisse von den Schulen konsequent genutzt werden müssen (Bildungsstandards und bayerische Lehrpläne [o.J.]).

4.3.2 Regionale Vernetzung der Schulentwicklung

Die zweite Ebene im Modell der Implementation von Bildungsstandards ist die regionale Vernetzung der Schulentwicklung. In den föderativen Systemen wird es komplizierte und vermutlich auch widersprüchliche Umsetzungen der nationalen Vorgaben geben, die aber wenig über reale Prozesse vor Ort aussagen. Seit jeher wurden landeseinheitliche Strukturen und Zielvorgaben immer nur unter Berücksichtigung regionaler Formen von Schulgestaltung durchgesetzt (Herrmann 1991). Bildungsstandards und damit verbundene Lehrpläne und Curricula, testfähige Beispielsammlungen, aber auch die Instrumente („Tools“) der Schul- und Unterrichtsentwicklung oder der Selbstevaluation von Schulen müssen eine lokale Ebene erreichen, wenn sie wirksam werden sollen. „Lokal“ meint nicht einfach die einzelne Schule oder den individuellen Unterricht. Wenn die Schulentwicklung regional vernetzt erfolgt, dann erfordert dies eine neue Steuerungsebene, die einmal „regionale Bildungsbüros“, einmal „regionale Steuerungsgruppen“ genannt wird, auf jeden Fall aber über Strategien der „Regional Governance“ (Fürst 2004) Einfluss auf die Qualitätsentwicklung vor Ort nimmt, ähnlich wie das in kanadischen oder amerikanischen Schuldistrikten der Fall ist. Auch hier ist nicht einfach die einzelne Schule die maßgebende Handlungseinheit. In Deutschland wurde „erst Mitte der 90er Jahre (...) in einigen Ländern ein schulpolitischer Perspektivenwechsel eingeleitet, in dessen Zusammenhang der lokalen und regionalen Schulentwicklung ein neuer Stellenwert zugemessen wurde“ (Zymek 2007, S. 279).¹⁵⁰ Vor dem Hintergrund dieser „neuen Aktualität der regionalen Schulentwicklung“ (ebd., S. 282) wären For-

¹⁵⁰ Zymek in seiner Einführung in den Themenschwerpunkt „Regionale Schulentwicklung“ in der Zeitschrift für Pädagogik (3/2007).

schungsdaten dringend notwendig; solche sind jedoch noch nicht in genügender Dichte und Qualität verfügbar (vgl. Weishaupt 2002).

Erfahrungen aus dem deutschen Bund-Länder-Projekt SINUS zeigen, dass eine fachbezogene, modularisierte Kooperation zwischen Lehrkräften in Netzwerken mindestens bei bestimmten Schularten¹⁵¹ leistungsfördernde Effekte mit sich bringt. Eine Bedingung dafür ist die klare Instrumentierung der Netzwerkarbeit mit Anreizen, Materialien, Werkzeugen und Ressourcen. Die Lehrkräfte müssen konkrete Anweisungen erhalten, wie sie vorgehen können, wenn sie eine Repertoireerweiterung ihres Unterrichts vornehmen sollen. Dabei sind schulische wie außerschulische, lokal organisierbare Netzwerke von großer Bedeutung; sie sorgen für wirksamen kollegialen Austausch im Fach und darauf bezogen für ein direktes Coaching.

Netzwerkarbeit im überfachlichen Sinn hat sich auch im Projekt „Selbstständige Schule“¹⁵² bewährt. Hier ist vor allem die Ebenenverschränkung in der Qualitätssicherung hervorzuheben. Aus Mitteln des Landes wie der Kommunen sind regionale Entwicklungsfonds aufgebaut worden, die die regionalen Steuerungsgruppen wirksam unterstützen. In verschiedenen Modellregionen entstanden Bildungsbüros, die mit der Koordination und Steuerung der Qualitätssicherung beauftragt wurden. Das sichert einen effizienten Einsatz der Mittel und verhindert Mehrspurigkeiten. Wenn nur die einzelne Schule im Fokus steht, dann muss sie auch die ganze Entwicklungsarbeit leisten, während Schulen dann am besten lernen, wenn sie ortsnahe und leicht zugängliche Beispiele zur Verfügung haben.

Die Implementation von Bildungsstandards verlangt von den Schulen den Aufbau neuer Verfahren, die Anpassung des Arbeitswissens der Organisation und nicht zuletzt die Überwindung hochgradiger Unsicherheiten. Die Erfolgskriterien ändern sich, die Standards steigern die Ansprüche an den Unterricht und die Ergebnisse werden transparent. Wenn diese durchaus riskanten Prozesse gelingen sollen, ist eine engmaschige Kooperation vor Ort und in der Region unausweichlich. Nur so können im Übrigen auch die beträchtlichen Investitionen in Implementationsstrategien wirksam eingesetzt werden. Andern-

¹⁵¹ Nach den vorliegenden Daten sind dies Hauptschulen, Schulen mit mehreren Bildungsgängen sowie Integrierte Gesamtschulen (Prenzel/Carstensen/Senkbeil/Ostermeier/Seidel 2005).

¹⁵² <http://www.selbststaendige-schule.nrw.de> sowie
<http://www.bertelsmann-stiftung.de/cps/rde/xchg/SID-0A000FOA-1>
(Stand 8.09.06).

falls wäre die einzelne Schule gehalten, jeweils den gesamten Entwicklungszyklus bei sich nachzuvollziehen.

Im Hinblick auf die neuen Instrumente der Implementation, also Kerncurricula, Tests, Aufgabenkulturen und Rückmeldesysteme, sollten frühzeitig regionale Spurguppen eingerichtet werden, die für eine wirkungsvolle Vernetzung der Entwicklung sorgen. Ähnlich wie bei SINUS müssen Instrumente und Handreichungen zur Verfügung stehen, in deren Gebrauch die Lehrkräfte eingeführt werden. Die einzelnen Schulen können damit bestimmte Lehrkräfte beauftragen, die dann als innerschulische Multiplikatoren eingesetzt werden. So hat die Implementation von Bildungsstandards einen konkreten Adressaten und zugleich eine regionale Verbreitung. Andernfalls ist die Gefahr groß, dass eine diffuse oder chaotische Praxis entsteht, die schnell hohe Motivationsverluste mit sich bringt und dann, wenn überhaupt, nur eine ganz schwache Akzeptanz erzeugt.

4.3.3 Einzelne Schulen

Die dritte Ebene ist die einzelne Schule, die sich wohl vernetzen kann, aber dadurch nicht aus der individuellen Verantwortung entlassen ist. Im Gegenteil bringt die höhere Selbstständigkeit der Schule eine erweiterte Verantwortung mit sich. Der Aufbau einer regionalen Steuerung impliziert:

- eine Verwaltungsreform an jenen Orten, an denen engmaschige bürokratische Vorschriften das selbstverantwortliche Handeln von Schulen gar nicht erst ermöglichen,
- neue von Einzelschulen abrufbare Unterstützungs- und Beratungsangebote sowie Instrumente der Qualitätssicherung und -entwicklung auf der Basis bejahter Ziele und selbst wahrgenommener Bedürfnisse,
- den Einbezug der Kompetenzen bisheriger Akteure und den Ausbau bestehender Qualitätsstrukturen und von „Best-practice“-Ansätzen,
- eine kompetente und in bestimmten Bereichen auch weisungsbefugte Schulleitung,
- die Akzeptanz der Innovationen bei den Akteuren
- und nicht zuletzt genügend Zeit und ausreichende Ressourcen bei der Umsetzung.

Auch die neuen Formen der Finanzierung in Nordrhein-Westfalen, etwa gemeinsame regionale Budgets von Land und Kommunen, setzen die Handlungsfähigkeit der einzelnen Schule voraus.

Bildungsstandards müssen den Unterricht in den Schulen erreichen. Das gelingt nur dann auf nachhaltige Weise, wenn dafür ein Klima der Akzeptanz oder gar des Aufbruchs erzeugt werden kann, wie dies etwa in bestimmten Projekten der Fall war (z.B. SINUS). Die Schulen müssen von den Vorteilen der Bildungsstandards überzeugt sein und hinter dem Projekt stehen, was neben der Ausbildung der Lehrkräfte auch interne Diskussionen voraussetzt. Die Akzeptanz ist umso größer,

- je weniger Rätsel oder Widersprüche mit dem Projekt verbunden sind,
- je offener die Vorteile zutage liegen,
- je besser die Wechselwirkung zwischen Anforderungen und Belastungssituationen bedacht werden,
- je besser die Kontinuität und Kohärenz im Prozess gesichert sind und
- je mehr der Verdacht entkräftet wird, dass all das schon einmal gescheitert sei.

Standards dürfen nicht als Neuauflage der (behavioristischen) Lernziel Diskussion der siebziger Jahre verstanden werden, aber auch nicht als Fortsetzung einer in der Regel lediglich rhetorisch zugestandenen Schulautonomie aus der Mitte der neunziger Jahre.

Die Meinungen entstehen im Kollegium und so im konkreten Austausch vor Ort. Bisläng nehmen Lehrkräfte „Bildungsstandards“ als Thema der Behörden- und Expertendiskussion wahr, welches ihren konkreten Erfahrungsraum nicht berührt. Dabei ist weitgehend unklar, was das Neue und Besondere der Bildungsstandards ausmacht. Der Ausdruck „Standards“ verleitet dazu, an überflüssige staatliche Normierungsversuche oder an Gleichmacherei zu denken, die die Freiheit der Lehrkräfte beschränken, ohne zu besseren Ergebnissen zu führen. Zudem verleitet eine bestimmte politische Rhetorik dazu, den Prozess als Rückkehr zu konservativen Formen des Unterrichts und des Umgangs zu deuten. Solche Barrieren lassen sich nur durch positive Beispiele und die Beteiligung an gelingenden Prozessen auflösen, die an der eigenen Schule oder in deren Umkreis erlebt werden. Deshalb sind auch die Unterstützungssysteme im Projekt „Bildungsstandards“ so wichtig, weil diesen die Aufgabe zukommt, die Einzelschulen mit den Kernideen der Reform in einen produktiven Kontakt zu bringen.

In diesem Sinne ist die Implementation von Bildungsstandards sowie von Verfahren der Zielsteuerung integraler Bestandteil der schulinternen Qualitätsentwicklung. Wir wiederholen diesen entscheidenden Punkt. Die Implementation kann also nicht isoliert betrachtet, sondern muss in ihrem Kontext ver-

ständig werden, und dies möglichst so, dass der Aufbau der Praxis in den Schulen die Befürchtungen überwindet, die die politische Rhetorik ausgelöst haben mag. Die verbreitete Skepsis gegenüber Bildungsstandards ist insofern verständlich, als starke Ankündigungen im Raume stehen, die ihre Probe in der Praxis noch vor sich haben. Andererseits hängt es von der Kommunikation vor Ort ab und ebenso von den Gelegenheiten, die lokal und regional bestehen, ob die Bildungsstandards auch als Chance und Herausforderung gesehen werden. Nur so stehen die Lehrkräfte hinter dem Projekt.

Damit zur internen und externen Evaluation verpflichtete Schulen ihre Aufgaben erfüllen können, benötigen sie entsprechende Rahmenbedingungen. Das heißt, die Philosophie einer „neuen Rechenschaftslegung“ verlangt die Koppelung einer verpflichtenden Qualitätssicherung mit Spielräumen von Schulautonomie und mit Ressourcen und Werkzeugen der lokalen und regionalen Schulentwicklung. Eine Steuerung teilautonomer Schulen ist nicht ohne eine *Reform des Verwaltungshandelns* möglich. Die administrative Regelungsdichte an der Basis muss abgebaut und es muss Raum für eigene Entscheidungen in wesentlichen Bereichen geschaffen werden, z.B. die Gestaltung der Schulverfassung, die Kompetenzen der Schulleitung, die Schulorganisation, individuelle Förderangebote oder die Nutzung schuleigener Budgets betreffend.

Dies bedeutet, dass auch die Kommunikationskultur gegen innen und außen verbessert werden muss und die Schulen Kompetenzen erwerben, mit Ergebnissen und Testdaten, die ihnen von außen oder von innen rückgemeldet werden, umzugehen. All dies bedingt die Etablierung einer neuen Kooperations- und Führungskultur und die Vernetzung der Einzelschulen dort, wo diese bestimmte Aufgaben nicht alleine erfüllen können. Dies betrifft vor allem die Aufgaben der Personalentwicklung oder Fortbildung.

4.3.4 Akteure: Schulleitungen, Lehrkräfte, Schüler und Eltern

Damit ist die vierte Ebene, die der individuellen Akteure, erreicht. Von einer gemeinsamen „Ebene“ kann gesprochen werden, weil die sehr verschiedenen Akteure an einer Schnittstelle, nämlich der des *Unterrichts*, zusammentreffen. Die Lehrkräfte, die Schülerinnen und Schüler, die Eltern oder die Erziehungsberechtigten haben ihre Gemeinsamkeit vor allem im Blick auf Unterricht und die damit verbundenen Entscheidungs-, Kooperations- und Führungsfunktionen. Die Lehrkräfte sehen hier ihr Kerngeschäft, nicht in der übergeordneten Schulentwicklung; die Schülerinnen und Schüler erleben Schule vor allem als Unterricht; und die Eltern oder die Erziehungsberechtigten sehen sich mit

Rückmeldungen aus dem täglichen Unterricht konfrontiert, auf den sie in irgendeiner Form reagieren.

Es gibt weitere Akteure, die nach ihrer Distanz zur Ebene des Unterrichts gruppiert werden können: die Schulleitungen, das nicht-pädagogische Personal einer Schule, die Schulaufsicht oder auch Vertreterinnen und Vertreter der weiteren Öffentlichkeit. Dort, wo die Schulaufsicht wie in den Schweizer Schulgemeinden vor allem lokal organisiert ist, ist die Nähe zum Unterricht vergleichsweise groß, während sie mit der Eingliederung in hierarchische Behörden abnimmt. Das gilt ähnlich für Schulleitungen, die dem Unterricht umso ferner stehen, je mehr sie mit professionellen Aufgaben betraut sind und eigene Kompetenzen wahrnehmen können. Aber selbstverständlich handelt es sich in jedem Fall um wichtige Akteure, die auch im Hinblick auf die Implementation von Bildungsstandards eine Rolle spielen.

Während Schulleitungen und Behörden der Schulaufsicht in diesem Prozess eine „offizielle“ Rolle spielen können, müssen die Lehrkräfte, die Schülerinnen und Schüler sowie die Eltern und Erziehungsberechtigten von den Vorteilen des neuen Verfahrens überzeugt werden. Die größte Rolle in diesem Feld der Akteure haben die Lehrkräfte. Bei der Einführung von Standards müssen sie einen doppelten Paradigmenwechsel vollziehen; einerseits indem Lernen als Entwicklung von wissensbasierten Kompetenzen verstanden werden soll, andererseits muss der Beherrschung der Grundlagen dieser Lernprozesse im Hinblick auf standardbasierte Tests Aufmerksamkeit geschenkt werden. Lehrkräfte können jede Reform unterlaufen und haben oft Erwartungen, die kaum eine Reform je erfüllen könnte. Sollen sie gewonnen werden, müssen sie das Innovationspotenzial im Klassenzimmer erkennen und Gelegenheit erhalten, ihr Repertoire mit Blick auf die neuen Anforderungen weiterzuentwickeln, ohne dabei zu erleben, dass sich negative Erwartungen bestätigen. Dies verlangt realistische Zielsetzungen und einen angemessenen Zeitrahmen für die Erprobung vor Ort.

Auch die *Position der Schülerinnen und Schüler* in der ihnen von der neueren Lerntheorie zugeschriebenen Rolle als zunehmend selbstverantwortlich Lernende und Handelnde muss auf die neuen Anforderungen eingestellt werden. Das heißt, dass sich den Kindern die in Bildungsstandards ausgedrückten Kompetenzen nicht einfach einpflanzen lassen, schon gar nicht in gleichförmigen Portionen und in für alle gleich abgemessenen Zeiträumen (Fend 2005). Mit anderen Worten: „Die schulischen Lernangebote treffen auf kindliche Nutzungsprozesse, die freudige Zuwendung aber auch defensive Abwehr erzeugen können“ (ebd., S. 26). Ob motivierte Zuwendung oder demotivierte Abwehr wird einen Unterschied machen. Es ist nämlich nicht so, dass sich im Unterricht – weder für Lehrpersonen noch für die Schülerinnen und Schüler – ei-

gentlich gar nicht viel ändert und nur die Leistungsbeurteilung umgestellt wird. Die neue, vermehrt auf Kompetenzen und weniger auf stoffliche Systematik bezogene curriculare Kultur steigert die Anforderungen an die Unterrichts- und Lernqualität, d.h. *die schulische Angebotsqualität und die Fähigkeiten zur Nutzung der angebotenen Lerngelegenheiten*. Dies, weil sie das Lernen verbindlicher auf Ziele ausrichtet, deren Erreichen auch überprüft wird, und weil sie den Unterricht mit Anwendungs- und Handlungsorientierung verbindet und vertieftes Verstehen oder Problemlösen erfordert. Überdies wird das Lernen gleichzeitig aktiver und selbstgesteuerter, was nicht produktiv möglich ist, ohne die Schülerinnen und Schüler dafür zu gewinnen. Dazu sind Anstrengungen zur Förderung ihrer Lernfähigkeiten und ihrer Motivation, die Lerngelegenheiten des Bildungssystems auch zu nutzen, in Rechnung zu stellen. Gegenüber dem Verdacht einer „mechanischen“ oder „technokratischen“ Bildungsreform kann deshalb darauf verwiesen werden, dass die Ansprüche an selbständiges Lernen, aber auch an Haftbarkeit und das Tragen von Verantwortung für (nicht) erreichte Ergebnisse sicher zunehmen werden, sofern tatsächlich Leistungstests, deren Qualität von den Nutzern anerkannt wird, als Steuerelemente eingeführt werden.

Tests, die das Erreichen von Kompetenzstufen messen, präzisieren nicht nur die Leistungserwartungen, sondern sie bedeuten für die Akteure auch, dass mehr Transparenz erzeugt wird. Der Aufbau von Rückmeldesystemen auf der Basis von Leistungstests schafft zudem neue Möglichkeiten der Partizipation. Eltern oder Erziehungspersonen können an der Offenlegung und Diskussion der Daten ebenso beteiligt werden wie die Schülerinnen und Schüler. Auf Klasebene ist dies in der Schweiz inzwischen erfolgreich ausprobiert worden (Tresch/Moser 2005). Eltern, Erziehungspersonen, Schülerinnen und Schüler sind für Bildungsstandards nur dann zu gewinnen, wenn sie neue Möglichkeiten und Vorteile gegenüber dem alten System erfahren können. Für die Akzeptanz des Projekts ist dies eine entscheidende Größe, die bei der Implementation daher gebührende Aufmerksamkeit erhalten muss.

Bislang stehen von den Akteuren allein die Lehrkräfte im Mittelpunkt der Aufmerksamkeit von Implementationsprogrammen. Wenn, dann sprechen sie die Lehrerinnen und Lehrer an. Als wichtige Bezugsgruppe nicht zu vergessen sind neben den bisher genannten zudem die Eltern. Für sie als Erziehungsberechtigte mögen oftmals Ängste vor dem Neuen oder überzogene Erwartungen dominieren. Viele besitzen in Bezug auf die eigenen Kinder jedoch auch klare Aspirationen. Gleichzeitig haben sie als ehemalige Angehörige aller Stufen des Bildungssystems Vorstellungen dessen, was Schule leisten kann und soll. Dies betrifft sowohl die Qualifikations-, die Sozialisations- als auch die Selektions- und Allokationsfunktion des Bildungswesens. Hinzu kommen Ansprüche be-

züglich der Vermittlungsinhalte von Schule. So zeigen die Ergebnisse des sogenannten Bildungsbarometers, einer Panelbefragung von 7000 Personen verschiedener Altersgruppen der Bundesrepublik, dass Eigeninitiative, selbständiges Lernen und eine breite Allgemeinbildung höher gewichtet werden als Fachwissen (Jäger/Mengelkamp 2006). Daraus wäre zu folgern, dass der Kommunikation und Information über die Ziele einer Schule, die ihre Leistungen zunehmend an der Erreichung von (fachbezogenen) Bildungsstandards bemessen will, die notwendige Aufmerksamkeit geschenkt werden sollte.

Programme für Eltern und Schülerinnen und Schüler liegen bislang nicht vor, weshalb sie entsprechend auf die kommende Situation neuer Lernanforderungen, anderer Formen der Leistungsüberprüfung und der Nutzung von Rückmeldesystemen nicht vorbereitet sind. Die Akzeptanz von Bildungsstandards und damit verbundenen neuen Instrumenten hängt aber stark davon ab, dass auch Eltern und Schülerinnen und Schüler damit umgehen können. Sie sollten, wie z.B. in Kanada, offensiv als Ressourcen – und nicht als „letzte Barriere“ – der Reform verstanden und behandelt werden.

4.3.5 Werkzeuge und Verfahren

Der Auftrag der Standardsicherung ist ein Rahmen, der auf der Ebene der Lehrpersonen durch Fortbildung und Schulentwicklung und auf der Seite der Schülerinnen und Schüler durch pädagogische Förderung ausgefüllt werden muss. Die letzte Ebene dieser Entwicklungs- und Förderanstrengungen ist diejenige der Verfahren und Instrumente. Nicht nur das Bildungssystem als solches, sondern auch seine Reformen brauchen Instrumente. Ohne geeignete „Tools“, die zu den Aufgaben und zu den Schulen bzw. den Systemebenen, in und auf denen sie eingesetzt werden, passen, wird wenig bewegt. Dass die instrumentelle Ausrüstung eines Unterrichtssystems und von Reformprojekten im Bildungsbereich zu den entscheidenden Erfolgsfaktoren zählt, ist erst in den letzten Jahren auch in der Forschung anerkannt worden (vgl. West/Staub 2002). Ein Beispiel lässt sich wiederum dem Programm der Bertelsmann-Stiftung entnehmen. Das Instrument bezieht sich auf die Gestaltung von „Selbstevaluationen in Schulen“ (SEIS) und ist als Unterstützung der internen Schulsteuerung zu verstehen. Zudem können damit externe Evaluationen vorbereitet und verknüpft werden. Konkret handelt es sich um einen Fragebogen, der mit Blick auf fünf Dimensionen der Schulqualität eingesetzt werden kann.¹⁵³ Die

¹⁵³ Bildungs- und Erziehungsauftrag, Lernen und Lehren, Führung und Management, Schulklima und Schulkultur, Zufriedenheit (Lehrer, Eltern und Schüler) (vgl.: *SEIS*)

Fragebögen können auch online ausgefüllt werden. Die Datenauswertung erfolgt elektronisch und lässt überdies auch Vergleiche mit anderen Schulen zu. Eine regionale Erprobung mit politischer Unterstützung durch die Behörden läuft seit Mitte 2005 in den beiden deutschen Bundesländern Niedersachsen und Baden-Württemberg. Unabhängig von dieser Erprobung waren zu Beginn des Jahres 2006 1.073 Schulen aus 15 deutschen Bundesländern registriert, die nach dem Verfahren arbeiten. Die Zahl steigt vergleichsweise schnell an, wobei berücksichtigt werden muss, dass die Internetplattform mit Auswertungssoftware erst seit dem 1. Mai 2005 zur Verfügung steht. Ähnliche Erfolge (und Zuwachsraten) zeigt die Schweizer Plattform *Klassencockpit* (Klassencockpit [o.J.]), die auf Leistungsvergleiche im Klassenzimmer zugeschnitten ist. Die Akzeptanz bei den Lehrkräften erklärt sich durch den direkten Nutzen und den in der Durchführung geringen Aufwand (Moser 2003).

Bei der Implementation von Bildungsstandards werden elektronische Instrumente auch in anderen Sektoren eine zentrale Rolle spielen. Das gilt für die Kommunikation des Projekts über die nationalen, regionalen und lokalen Plattformen ebenso wie für die Rückmeldesysteme oder die Eingabe und Auswertung von projektspezifischen Fragebögen. Schnelle und in der Handhabung unproblematische Instrumente steigern die Projektakzeptanz und sichern den Rückhalt bei den Akteuren. Sie lernen nicht zuletzt auf diese Weise, was der Paradigmenwechsel in Richtung „Outputsteuerung“ konkret bedeutet und welche neuen Möglichkeiten er eröffnet.

4.4 Kontext- und Unterstützungssysteme

Die zentralen Konzepte dieser Expertise in Bezug auf die vorgeschlagene Strategie zur Innovationssteuerung im Zusammenhang mit der Einführung und Umsetzung von Bildungsstandards sind *Lernen* und *Schulentwicklung* unter der Spezifik gegebener Schulstrukturen und -kulturen. Der Fokus richtet sich dabei auf die Möglichkeiten und Voraussetzungen für anhaltendes (individuelles und soziales, kooperatives) Lernen von Lehrpersonen, Schulleiterinnen und -leitern sowie der anderen Akteure. Da Schulkulturen sich jedoch nicht isoliert, sondern im Kontext von Schulaufsichts- und Unterstützungssystemen weiterentwickeln,

müssen auch diese in den Implementations- und Transferprozess einbezogen werden.

Zu den Kontextsystemen gehören die Einrichtungen der Lehreraus-, -fort- und -weiterbildung ebenso wie die Schulaufsicht und die anderen Behörden sowie die Landesinstitute und pädagogischen Arbeitsstellen. Ohne die längerfristige Schaffung von Zielverpflichtung und Akzeptanz bei allen Beteiligten kann ein Projekt wie die Etablierung eines neuen Referenzsystems für Bildungsqualität nicht gelingen. Wenn die Implementation von Bildungsstandards gelingen soll, müssen alle Einrichtungen, die bisher an der Aus- und Weiterbildung der Lehrpersonen sowie an der Schulverwaltung beteiligt waren, eine neue Akzentuierung ihrer Aufgaben vornehmen und an deren Ausführung mitwirken. Da die bildungspolitischen und strukturellen Voraussetzungen in den Auftragsländern dieser Expertise und teilweise auch innerhalb eines einzelnen Auftragslandes und seiner Regionen verschieden sind, kann dies nicht auf universell gleiche Weise geschehen, sondern muss den Bedingungen von kultureller Differenz Rechnung tragen.

Das Ziel der Politik muss es sein, für eine Abstimmung („alignment“) der Arbeit der Unterstützungssysteme mit den Zielen der Reforminitiative zu sorgen, ohne damit die Ausrichtung der Unterstützungssysteme allein auf diese Aufgabe konzentrieren zu wollen. Es geht nicht darum, künftig *alle* Ressourcen in Richtung Bildungsstandards zu lenken, wohl aber muss es möglich sein, Prioritäten festzulegen und in den Budgets Umschichtungen vorzunehmen.

Was sich vielerorts auch ändern dürfte, sind die Funktionen der *Schulaufsicht* in ihrer Beziehung und Komplementarität zu Funktionen der Schulinspektion und zu neu zu etablierenden Verfahren der externen Schulevaluation einerseits sowie der Schulentwicklungs- und Evaluationsberatung andererseits. Die Auftragsländer zeigen hier unterschiedliche Konfigurationen und Ausgangspunkte, so dass nicht mit einheitlichen Lösungen gerechnet werden kann. Aufgrund von Erfahrungen und Forschungsergebnissen im internationalen Raum (vgl. Kapitel 3) scheint es uns wichtig anzustreben, Kontrollfunktionen von Beratungsfunktionen zu trennen, ohne eine der beiden Funktionen aufzugeben (vgl. nochmals das Projekt „Selbständige Schule“ in Nordrhein-Westfalen)

Als eine *Pièce de Résistance* des gesamten Unternehmens sind die Einrichtungen der Lehrergrundausbildung sowie der Lehrerfort- und -weiterbildung zu betrachten. Was die *Lehrergrundausbildung* anbelangt, dürfte es wegen der angestammten Autonomie, die sie vor allem in Deutschland an den Universitäten (erste Phase) genießt, aber auch infolge ihrer langjährigen Tradition der Disparität und der Beliebigkeit in der curricularen Ausrichtung ihrer Angebote nicht einfach sein, sie in eine kohärente Qualitätspolitik in Bezug auf Standards ein-

zubinden. Auch im Blick auf die *Lehrerfort- und -weiterbildung* sind neue Formen zu finden (vgl. die weiter oben dargelegten Prinzipien in Hawley/Valli 1999). Durch traditionelle, sich an einzelne Lehrpersonen richtende Formen wird die Praxis an Schulen kaum verändert. Wir haben das ausführlich in den Länderberichten in Kapitel 3 dargestellt.

Der Stand der Entwicklung ist in den Auftragsländern verschieden. In der Schweiz ist mit der Etablierung von Pädagogischen Hochschulen der Status der Ausbildung der Lehrkräfte verändert worden. Ein ähnlicher Schritt besteht in der Überführung der jetzigen Pädagogischen Akademien in eigenständige Pädagogische Hochschulen in Österreich, die für 2007 vorgesehen ist.¹⁵⁴ In Deutschland ist die zweiphasige Lehrerbildung in der Diskussion, ein grundlegender Wandel ist allerdings nicht zu erwarten. Einige Bundesländer haben sogenannte „Praxissemester“ eingeführt, andere haben die Studienseminare stärker an der ersten Phase beteiligt; eine Rückkehr zu einer einphasigen Ausbildung ist allerdings nirgendwo vorgesehen.

4.4.1 Aus- und Weiterbildung der Lehrkräfte

Die verschiedenen Phasen der Ausbildung der Lehrkräfte in Deutschland schaffen eine spezifische Problemlage, die in den anderen Auftragsländern so nicht gegeben ist. Gewisse Transferprobleme gibt es in jeder Lehrerbildung, im deutschen System müssen die Lernenden aber nicht nur einen Wissenstransfer zwischen den Orten der akademischen Ausbildung und denen der praktischen Anwendung vornehmen, sondern das Ausbildungswissen darüber hinaus konsekutiv von einer Phase in die nächste übertragen, ohne dass die Ausbildung dafür auch nur für eine curriculare Abstimmung sorgen würde. Das hat Folgen, die in der Schweizer Lehrerbildung aufgrund ihrer auch nach der Akademisierung bestehenden Einphasigkeit sowie der traditionell engen Theorie-Praxis-Bezüge nicht zu beobachten sind.

Im Mittelpunkt der Lehrerbildung in Deutschland steht nur sehr bedingt die Profession der Lehrperson oder der Unterricht als das „Kerngeschäft“ des Berufs. Ausgebildet wird für verschiedene Lehrämter. Die Erfahrung des Studiums ist oft gekennzeichnet durch Zusammenhangslosigkeit, Beliebigkeit und

¹⁵⁴ Die bisher zwischen verschiedenen Anbietern aufgeteilte Ausbildung, Fort- und Weiterbildung der Lehrkräfte wird in Zukunft einer einzigen Institution übertragen. Damit soll ein Professionalisierungskontinuum geschaffen werden. Nötig ist dafür ein eigenes Gesetz für die Pädagogischen Hochschulen (Vorblatt zum Hochschulgesetz 2005).

Praxisferne. Die Lehrerbildung kennt keine klar geregelte Verantwortung für die gesamte Ausbildung, die Entwicklung berufsfeldbezogener Standards hat in der Breite gerade erst begonnen und ist in den deutschen Bundesländern bislang sehr unterschiedlich. Zwischen den *Lehrerbildungsstandards* und den *Bildungsstandards* für die Schulen gibt es bislang kein Verbindungsglied, das auf eine zusammenhängende Entwicklung verweisen würde.

Ob Ziele der Ausbildung erreicht werden oder nicht, ist schwer zu sagen, weil klare und verbindliche Zielsetzungen in aller Regel fehlen. Wenn man die deutsche Lehrerbildung mit Ausbildungen im Ausland vergleicht, dann fällt auf, dass die Fachausbildung nirgendwo so dominant ist wie in Deutschland. Wir verweisen etwa auf die oben erwähnte Entwicklung in Schweden. Nirgendwo sonst gibt es auch die Trennung von zwei Ausbildungsphasen, die wenig bis nichts miteinander zu tun haben, und nirgendwo zerfällt die Ausbildung in Segmente, die weitgehend unkoordiniert arbeiten können oder müssen.

Amerikanische Erfahrungen verweisen darauf, dass neue Steuerungselemente für die Entwicklung der Lehrerbildung unerlässlich sind (Darling-Hammond 2000a, 2000b). Die neuen Instrumente zur Veränderung des Systems sind international sehr ähnlich und haben sich inzwischen auch bewährt. Grundlegend sind überall Zielsteuerung, Entwicklung von eigenen Standards der Lehrerbildung sowie regelmässige interne und externe Evaluationen. Hinzu kommen ein gewisser Wettbewerb zwischen den Ausbildungsinstitutionen sowie die Beteiligung der Abnehmer und Absolventen, in dem Sinne, dass die Ausbildungsinstitutionen von den Feedbacks ihrer Kunden lernen. Letztlich geht es auch hier um die Umstellung der Ausbildungskultur im Sinne der Orientierung an anzustrebenden Ergebnissen: Konkrete Standards für die Lehrerbildung sollen möglichst präzise Ziele und Ergebnisse des Unternehmens Lehrerbildung festlegen und die zu erwerbenden Fähigkeiten beschreiben können. Weiter sollen sie als Absolventenstandards einen „idealen“ Erwartungsraum definieren (vgl. Terhart 2007). Dabei muss jedoch unterstrichen werden, dass es nicht darum gehen kann, die Lehrerbildung auf eine eindimensionale Funktionslogik festzulegen, im Sinne von: Je besser ausgebildet die Lehrpersonen sind, desto höhere Fachleistungen erzeugen sie bei den Schülerinnen und Schülern. Eine anzustrebende Standardorientierung der Lehrerbildung soll dazu beitragen, dass Absolventen den herausforderungsreichen und vielseitigen Handlungsfeldern ihres Berufes gewachsen sind und dies anhand berufsrelevanter Kompetenzen und mit Blick auf multikriteriale Bildungsziele, welche mehr als Fachleistungen umfassen, demonstrieren können.

In Deutschland hat die Lehrerbildung von programmatischen Aussagen gelebt, ohne durch nachhaltige Evaluationen gesteuert zu werden, die nur möglich

und im Übrigen auch nur fair sind, wenn sie von Standards ausgehen können. In den Institutionen der Lehrerbildung gibt es bislang noch kaum entwickelte Systeme der Qualitätssicherung, und wenn, dann finden sie sich an den Studienseminaren und nicht an den Universitäten. Einige Studienseminare haben begonnen, eigene Standards aufzubauen, was allerdings uneinheitlich und wenig gesteuert geschieht. Eine Schwäche ist die zu geringe Kooperation zwischen den Praxislehrkräften und den Studienseminaren. Was hier möglich ist, lässt sich am Beispiel der Ausbildung und Qualifizierung von Ausbildungslehrkräften am Institut für Qualitätsentwicklung in Schleswig-Holstein zeigen.¹⁵⁵

Zum Vergleich: Die vor einigen Jahren neu gegründeten Pädagogischen Hochschulen in der Schweiz haben fast alle fachliche sowie überfachliche Standards entwickelt, führen interne Evaluationen der Lehre durch und werden in Zukunft einem nationalen Assessment unterliegen. Einige Institutionen wie etwa die Pädagogische Hochschule in Zug haben eine sehr weitgehende Steuerung durch berufsfeldbezogene Standards entwickelt, die ausbildungsübergreifend angelegt sind. Es sind in diesem Sinne echte Standards für die Ausübung des Berufs, die nicht mit dem Ende der Ausbildung erreicht sind (PHZ Schwyz/PHZ Zug 2004).

Die Pädagogische Hochschule Zürich hat im Januar 2004 ein Strategiepapier verabschiedet, das auf der Basis einer detaillierten Stärke-Schwäche-Analyse klare Entwicklungsziele in allen Bereichen, also Lehre, Forschung, Weiterbildung, Dienstleistung und Verwaltung, auflistet. Diese Ziele sollen bis 2008 erreicht werden; es handelt sich um einen Businessplan, der verbindliche Vorgaben macht und an dem die Hochschule gemessen werden will (Pädagogische Hochschule Zürich 2004). Zu diesen Aufgaben gehört auch die Entwicklung eines Verfahrens zur Eignungsabklärung. Studierende können von der Ausbildung ausgeschlossen werden, wenn sie pädagogisch nicht geeignet sind.

Qualitätssicherung in diesem Sinne ist eine Kernaufgabe für die Entwicklung der Lehrerbildung. Sie muss eine echte Zielsteuerung ausbilden und sich an ihren Ergebnissen orientieren, was bislang nicht annähernd der Fall ist. Wesentlich für die Qualitätsentwicklung der Lehrerbildung in Deutschland sind neue Formen der Organisation wie verantwortliche Leitungen, gemeinsame Zielsetzungen zwischen den Phasen oder die Steuerung der Entwicklung durch externe Evaluationen. Auf curricularer Ebene geht es um fachliche sowie über-

¹⁵⁵ Das Ministerium für Bildung und Frauen des Landes Schleswig-Holstein hat am 16. Juni 2006 einen Erlass herausgegeben, der das Zertifikat für die Tätigkeit als Ausbildungslehrkraft regelt.

fachliche Standards, die verbindlichen Charakter annehmen und den Kern des Angebots bestimmen.

In Deutschland haben verschiedene Länder und auch die Kultusministerkonferenz (KMK) mit einer Reform der Lehrerbildung begonnen, die inhaltlich bestimmt ist. Bislang gab es keine verbindlichen oder gar unter den Phasen abgestimmte Curricula, und dies weder innerhalb der Länder noch bundesweit. Die neuen Curricula sind nicht wie die bisherigen Studienordnungen formuliert und stellen auch keine Lehrpläne alter Art dar, in denen nichts festgelegt wurde und die jeder Nutzer mit eigenen Freiheiten interpretieren konnte. Die jetzige Lehrerbildung ist dozentenorientiert, jeder kann anbieten, was er oder sie will, ohne auf verbindliche Vorgaben achten zu müssen. Wenn aber die Ausbildung mit berufsfähigen Kompetenzen abschließen soll, muss festgelegt sein, was diese ausmacht. Damit soll nicht gesagt sein, dass nichts anderes gelernt wird, die Ausbildungserfahrungen übersteigen immer das Curriculum; aber keine Ausbildung erreicht Ziele, wenn sie nicht durch verbindliche Erwartungen gesteuert wird.

Genau diese Einsicht bestimmt den Systemwandel in der Bildungspolitik. Die deutsche Kultusministerkonferenz (KMK) hat am 16. Dezember 2004 erste Vorschläge für Standards in der Lehrerbildung verabschiedet, die die sogenannten „Bildungswissenschaften“ betreffen (Kultusministerkonferenz (KMK) 2004b). Standards für die Fachwissenschaften werden folgen. Der Beschluss der KMK geht zum Teil zurück auf Vorschläge, die die „Arbeitsgruppe Curriculare Standards Bildungswissenschaften beim Ministerium für Wissenschaft, Weiterbildung, Forschung und Kultur“ in Rheinland-Pfalz vorgelegt hat. Das Gutachten (Arbeitsgruppe Curriculare Standards Bildungswissenschaften beim Ministerium für Wissenschaft, Weiterbildung, Forschung und Kultur in Rheinland-Pfalz 2004) lag im März 2004 vor, danach wurde die Kommission der KMK tätig.

Im Beschluss werden zwei zentrale begriffliche Konzepte verwendet, nämlich Kompetenzen und Standards. „Kompetenzen“ sind die erreichten Fähigkeiten der Absolventen am Ende der Ausbildung und weiter im Berufsfeld. Gemeint ist die persönliche Fähigkeit und Fertigkeit, Unterricht zu erteilen und Schulstunden abzuhalten. Der Anspruch ist nicht neu, wurde aber nie wirklich kodifiziert. Mit Standards für die Lehrerbildung wird es möglich sein, in präziser Weise verbindliche Ziele zu formulieren, die in allen Ausbildungsteilen gelten und auch erreicht werden können. Heutige Ziele, wenn es sie denn gibt, sind in aller Regel abstrakt, vage formuliert und letztlich unerreichbar.

Standards implizieren unterschiedliche Niveaus. Wer standardbezogen lernt, kann *fachliche* und *überfachliche* Kompetenzen auf verschiedenen Niveaus errei-

chen. Diese Unterscheidung wird die weitere Diskussion bestimmen, weil anders kaum ein curriculares Verhältnis zwischen Fachwissenschaften und Bildungswissenschaften entwickelt werden kann. *Fachliche Kompetenzen* in der Lehrerbildung beziehen sich auf Fähigkeiten, ein schulisches Unterrichtsfach oder eine Lerndomäne zu unterrichten, *überfachliche Kompetenzen* sind allgemeine professionelle Kompetenzen des Lernens, der Reflexion und der Entwicklung.

Dabei hat jedes akademische Studium, das sich auf einen Beruf bezieht, eine grundlegende Schwierigkeit, die besser oder schlechter überwunden werden kann. Die Fachlichkeiten von Studium und Beruf sind nicht identisch. Themen oder Inhalte der Hochschulausbildung beziehen sich auf wissenschaftliches Wissen und nicht zugleich auch auf professionelles Können. Curriculare Standards eines akademischen Studiums verweisen auf Forschungsfelder, Theorien, Modelle und Datensätze, die vor allem reflexiv erfahren werden. Es sind Seminar- und Vorlesungsthemen, denen ein direkter Anwendungsbezug fehlt. Professionelles Können aber entsteht und festigt sich nur unter der Voraussetzung eigener Erfahrungen im Zielbereich, also situiert im schulischen Unterricht und seinem Umfeld.

Lösbar ist dieses Problem dann, wenn Wissen und Können parallel und zusammenhängend, nach Möglichkeit situiert und problemorientiert (Gräsel 1997; Reusser 2005) aufgebaut werden. Wenn wirklich *Kompetenzen* die Ausbildung beschließen und in den Beruf hineinführen sollen, dann müssen die wissenschaftlichen und die berufspraktischen Teile nicht nur aufeinander bezogen, sondern so weit wie möglich verzahnt werden. Seminarerfahrungen und schulpraktische Studien haben dann thematische Bezüge und sind aufeinander abgestimmt. Die Universitätsangebote sind curricular auf das Berufsfeld eingestellt, die Studierenden erleben eigenen Unterricht, den sie reflektieren und mit wissenschaftlichem Know-how in Verbindung bringen können. Sie machen in diesem Sinne klinische Erfahrungen, die die beiden Welten von Theorie und Praxis so weit und so gut wie möglich integrieren.

Die neue Form der Modularisierung erleichtert im Prinzip diese Entwicklung, weil sich verschiedene Formen der Lehre verknüpfen lassen und Lehrpersonal aus unterschiedlichen Bereichen kooperieren kann. Die Universität Frankfurt¹⁵⁶ etwa hat unlängst ein Seminarangebot im Bereich Erziehungs- und Gesellschaftswissenschaften eingerichtet, in dessen Rahmen Lehramtsstudierende und Referendare für das Lehramt an Gymnasien gemeinsam Lehrveranstaltungen

¹⁵⁶ Pressemitteilung der Johann Wolfgang Goethe-Universität Frankfurt vom 20. Juni 2005.

besuchen können. Die Module werden einheitlich bewertet und für die Staatsexamen angerechnet. In diesem Sinne untertunnelt die Modularisierung die Phasentrennung, auch weil die Lernleistungen von Studierenden und Referendaren nach einem einheitlichen Punktemodell berechnet werden.

Es gibt keine universelle Form, und es wird eine solche aufgrund der Komplexität der Aufgabe auch kaum je geben, in der der „Praxisbezug“ von Lehrerausbildungen auf ideale Weise wirkungsvoll und effizient organisiert werden kann. Jedoch gibt es zweifellos effektivere und wenig effektive Realisierungen der Verknüpfung der akademischen und der berufspraktischen Ausbildungselemente an Hochschulen. Dass das zweiphasige deutsche System zu den wirkungsvollsten gehört, muss dabei bezweifelt werden. Heute beschränkt sich die Zuständigkeit in der ersten Phase auf die Praktika, wobei die übrigen Lehrveranstaltungen von der Verantwortung entlastet sind. Die Erfahrungen der Studierenden werden damit nicht wirklich oder viel zu wenig für das anschließende Themenaufkommen genutzt; die „schulpraktischen Studien“ sind daher oft isoliert, während sie für die Studierenden ein zentrales Studienelement darstellen, das nicht zuletzt als Probe für das sich entwickelnde eigene Können verstanden wird. Die Erfahrungen in der Praxis generieren Fragen, die im weiteren Studium genutzt werden müssen, einschließlich einer Beratung über die Eignung für den Beruf, die in der deutschen Ausbildung immer noch weitgehend fehlt.

Das kann und muss sich mit der Modularisierung des Studiums ändern. Die Mainzer Arbeitsgruppe zu den Bildungswissenschaften hat dafür ein bestimmtes Format festgelegt. Gemäß dieser Beschreibung erhält ein Studienmodul folgende Merkmale:

- das Thema,
- verbindliche Inhalte,
- Qualifikationsziele/erwartete Kompetenzen,
- Stellung im Studium,
- Lehr- und Lernformen (didaktische und methodische Standards),
- Teilnahmevoraussetzungen (Kenntnisse, Fähigkeiten, vorausgehende Module, Lektüreforderungen),
- Dauer in SWS,
- Anrechnungspunkte (credits) und Arbeitsaufwand (workload) sowie
- Form und Umfang studienbegleitender Prüfungen.

Wenn alle Module so formatiert sind, wird es möglich sein, fortlaufend und fließend die tatsächlich erbrachten Leistungen nachzuweisen. Das entsprechende Studien-Portfolio würde die Leistungsbewertung dokumentieren und die Noten hätten eine transparente Basis. Am Ende wird sichtbar, was gelernt

wurde und was genau die Kompetenz ausmacht. Werden die schulpraktischen Studien als Teil eines Moduls verstanden, sind ganz andere Kooperationen und inhaltliche Verzahnungen möglich, als dies heute der Fall ist. Die zweite Phase könnte eingebunden werden, Vertreter verschiedener Disziplinen können zusammenarbeiten und die Erfahrungen der Studierenden hätten thematisches Gewicht. Ein Modell wie in Frankfurt wäre überall möglich.

Anzumerken ist, dass eine konsequente Modularisierung viele Vorteile aufweist. Neben der transparenten Ausrichtung auf Ziele, Kompetenzen und Standards sind dies etwa die Möglichkeiten zu einem problemorientierten, Sinnzusammenhänge schaffenden Unterricht. Jedoch bleiben auch einige Probleme ungelöst. Das wichtigste betrifft die *Kohärenz bzw. Fragmentierung* des Curriculums. Durch Kooperations- und andere Abstimmungsmaßnahmen zur Integration der modularen Ausbildungselemente muss sichergestellt werden, dass auch in einem modularisierten Lehrbildungscurriculum Teilkompetenzen und Partikulareinsichten in ihren gegenseitigen Bezügen und wechselseitigen Verknüpfungen gesehen werden. Das heißt, dass modulübergreifende Handlungs- und Sinnzusammenhänge nicht verloren gehen dürfen, sondern auf der Rezeptionsebene von den Studierenden gesichert werden müssen.

Modularisierung ist eine notwendige Form für den Anfang des Systemwechsels. In einem weiteren Schritt muss die Ausbildung auf *Zielsteuerung und Evaluation* hin ausgerichtet werden, andernfalls ist sie nicht wirklich lernfähig. Darauf sind die deutschen Universitäten und die lehrerbildenden Hochschulen bislang nicht oder nur sehr unzureichend eingestellt. Weil das so ist, muss die Reform umfassend angelegt sein und kann nicht nur auf die konsekutive Struktur bezogen werden. Verlangt wird von der Lehrerbildung auch die Neugestaltung ihrer Curricula und damit einhergehend die Orientierung am Resultat, was eine Umstellung der gesamten Organisation bedeutet.

Die Lehrerbildung wird lernen müssen, ihre Absolventen und Abnehmer zu befragen, das Lehrangebot wird zunehmend vom Erfolg bei den Studierenden abhängig, und zwischen den Phasen werden Feedback-Systeme etabliert, mit denen der konkrete Verlauf des Kompetenzaufbaus begleitet wird. „Praxisbezug“ wird auf diesem Weg weit mehr sein als die Antwort auf die Frage der Studierenden „What works?“. Auch diese Antworten müssen überzeugend sein, wenn das Vertrauen darauf erhalten bleiben soll, dass die Ausbildung kontinuierlich berufliche Kompetenzen aufbaut, die sich auch tatsächlich anwenden lassen. Aber „Verwendbarkeit“ ist mehr als das, die Ausbildung muss fortlaufend überprüfen, ob sie im Hinblick auf ihren Zweck die richtigen Angebote macht und wie diese optimiert werden können. Dazu braucht sie die Rückmeldung nicht nur ihrer Absolventen, sondern auch ihrer Abnehmer.

Geklärt werden muss damit insbesondere, welcher Zusammenhang zwischen der Art der von den künftigen Lehrerinnen und Lehrern erfahrenen Lehrerbildung und ihrem späteren beruflichen Handeln bzw. dem dadurch bewirkten Lernen und seinen Ergebnissen bei den Schülerinnen und Schülern besteht (vgl. Terhart 2006). Erzielen besser ausgebildete Lehrpersonen unter relevanten, multidimensionalen Zielkriterien, darunter Bildungsstandards bezogen auf die Unterrichtsfächer der Zielstufe, bessere Ergebnisse bei den Schülerinnen und Schülern? Lehrerbildung ist ein teures Investitionsgut der Gesellschaft. Da erscheint es legitim, wenn diese genauer und jenseits rhetorischer Versprechungen wissen möchte, welche gesamt-systemischen Anstrengungen bzw. welche Varianten von Lehrerbildung welche Wirkungen erzeugen. Das heißt, auch in der Lehrerbildung muss angestrebt werden, dass ihre Absichten nicht unverbindlich und unpräzise, und ihre tatsächlichen Effekte nicht zufällig bleiben, sondern sich an Standards – einerseits beschreibbar als Anforderungen an die Institutionen selbst (Institutions- und Prozessstandards), andererseits als erwartbare Fähigkeiten ihrer Absolventen (Wissens- und Kompetenzstandards) – messen lassen.

Inzwischen liegen in Rheinland-Pfalz auch Ergebnisse von Arbeitsgruppen für die Entwicklung curricularer Standards in den Studienfächern vor. Weil die Fachausbildung den eigentlichen Kern der Ausbildung darstellt, bislang aber nur unverbindliche Studienordnungen vorlagen, die gegenüber den Anforderungen der Lehrerbildung unspezifisch waren, liegt hier der eigentliche Meilenstein der Reform. Die Fachwissenschaften standen immer im Verdacht, sich um die Belange der Lehrerbildung nicht gebührend gekümmert zu haben, vor allem, weil sie ihre Aufgabe nicht deutlich genug vor Augen hatten. Die heftigen Diskussionen an den Universitäten von Rheinland-Pfalz zeigen, um welchen Umbruch es sich handelt.

Die neue Struktur, die Standards, Kompetenzbeschreibungen und Studienmodule umfasst, ist für alle Fächer entwickelt worden. Die Entwürfe zeigen, dass möglich ist, was bislang immer ideologisch ausgeschlossen wurde, nämlich Leitbilder für künftige Lehrpersonen aus der Sicht der Fachwissenschaften, Kompetenzprofile sowie klar strukturierte Studienmodule, die verbindlich vorschreiben, was und wie studiert werden muss. Das ist für Studiengänge in anderen Bereichen und im Ausland seit langem selbstverständlich. Niemand kann sich seine Ausbildung selbst zusammenstellen, wie dies in der Lehrerbildung in Deutschland üblich war und immer noch ist.

Wenn es im Entwurf für die Kompetenzen der angehenden Lehrkräfte im Fach Deutsch heißt, am Ende der Ausbildung müssten die Lehrkräfte imstande sein, „fachspezifische Lernprozesse zu initiieren, zu begleiten und zu beurteilen

– vornehmlich mit dem Ziel, bei Kindern und Jugendlichen sichere Kenntnisse über Sprache und Literatur zu fördern“ (Curriculare Standards des Fachs Deutsch [o.J.], S. 4), dann ist das mehr als eine Selbstverpflichtung. Die Erreichung dieses Ziels kann überprüft werden, sofern es tatsächlich Rückmeldesysteme gibt. Die Ausbildung kann damit auch lernen, was zu tun ist, wenn derartige Ziele nicht erreicht werden. Heute hat man höchstens die Möglichkeit zu lamentieren.

Auch der Bericht der Mainzer Arbeitsgruppe zu den Bildungswissenschaften unterscheidet Kompetenzen und Standards. Kompetenzen sind Erwartungen des Könnens am Ende der Ausbildung (Arbeitsgruppe Curriculare Standards Bildungswissenschaften beim Ministerium für Wissenschaft, Weiterbildung, Forschung und Kultur in Rheinland-Pfalz 2004, S. 9f.), Standards jedoch beziehen sich auf mehr als nur die Inhalte der Bildungswissenschaften, nämlich ebenso auf die Formen der Vermittlung im Hochschulunterricht sowie auf die Ressourcen der Lehrerbildungssysteme. Dies festzuhalten ist wichtig, weil unter „Standards“ oft nur die curricularen Inhalte verstanden werden, was dazu verleiten würde, einfach nur die vorhandenen Stichwortlisten umzuschreiben. In der Vergangenheit sind abstrakt gut klingende Begriffssequenzen formuliert worden, die nichts festlegen und alles offen lassen, damit möglichst viel vom Status quo bewahrt wird, während genau das Umgekehrte angebracht ist, wenn die Reform gelingen soll.

Standards in der Lehrerbildung legen fest und lassen nicht alles offen, was in vielen Studiengängen außerhalb der Lehrerbildung längst selbstverständlich ist. Gemäß der Definition der Arbeitsgruppe Bildungswissenschaften

- müssen die Curricula auf die Ausbildungsziele ausgerichtet werden,
- darf die Qualität der Lehre bestimmte Niveaus nicht unterschreiten,
- müssen ausreichend Ressourcen vorhanden sein,
- sind Formen der Qualitätssicherung angesprochen
- und wird die Systementwicklung qualifiziert,
- einschließlich der Anforderungen an das künftige Personal der Lehrerbildung in allen Phasen.

Wenn also von „Standards“ in der Lehrerbildung gesprochen wird, dann sollte das Gesamt der Aufgabe vor Augen stehen. Es geht um die Themen und Inhalte der Ausbildung, aber auch um die Forschungsprofile der künftigen Professoren in den Lehrerbildungsbereichen sowie um die Ausbildung der Lehrkräfte an den Studienseminaren oder die scheinbare Randfrage des Curricular-

Norm-wertes der Pädagogik, an dem sich ein Grossteil der Finanzierung entscheiden wird (ebd., S. 22).

Die Lage der deutschen Lehrerbildung war seit der Integration der Pädagogischen Hochschulen¹⁵⁷ in den siebziger Jahren festgefahren. Das gilt für die Aufteilung der Ressourcen ebenso wie für die Interpretation der Aufgaben und die wechselseitige Beobachtung der Phasen. Entwicklung war lange Jahre nicht zu beobachten, obwohl die Mängel der Lehrerbildung durchgehend und immer wieder thematisiert wurden. Aber erst nach dem Gutachten der KMK-Kommission (Terhart 2000) und nicht zuletzt unter dem Eindruck der PISA-Diskussion ist in Deutschland Interesse für die Weiterentwicklung der Lehrerbildung entstanden und das Thema dringlich geworden.

Inzwischen ist die politische Bewegung in diesem Bereich unverkennbar. Die 310. Plenarsitzung der Kultusministerkonferenz vom 2. und 3. Juni 2005 hat Eckwerte für die Lehrerbildung beschlossen und Regeln für die gegenseitige Anerkennung von Bachelor- und Masterabschlüssen bestimmt. Dabei geht es zunächst um „ländergemeinsame“ Standards. Standards in den Bildungswissenschaften sollen weiterentwickelt, Standards für die Fachwissenschaften und deren Didaktik erarbeitet werden. Schulpraktische Studien werden bereits in den Bachelor-Studiengängen angeboten, ihr Anteil wird insgesamt deutlich erhöht. Die einzelnen Ausbildungsteile sollen stärker vernetzt werden, die Studiengänge erhalten eine modulare Struktur und auch alle Studiengänge der Lehrerbildung durchlaufen einen Akkreditierungsprozess.

Die Implementation von Standards der Lehrerbildung setzt voraus, dass zwischen Universitäten/Hochschulen und Studienseminaren sowie der Schulpraxis curriculare Abstimmungen erreicht werden, die durch die neuen Formen der Modularisierung und der Kreditierung der Studienleistungen erleichtert werden, aber immer noch auf erheblichen Widerstand stoßen. Letztlich muss auch für die Vision der Reform geworben werden, und dies in einem Bereich, der gegenüber neuen Instrumenten eher skeptisch eingestellt ist.

Die Aufgabe wird sein, die Dozierenden zu veranlassen, mit einem Curriculum zu arbeiten, das auf verbindlichen Vorgaben beruht. Heute, wie gesagt, kann jeder Dozent anbieten, was ihm im Rahmen seiner wissenschaftlichen Spezialisierung als sinnvoll erscheint. Das neue Curriculum müsste unter den einzelnen Dozierenden abgestimmt werden und an die zweite Phase sowie die Berufspraxis anschlussfähig sein. Zudem müsste die Ausbildung in der Lage

¹⁵⁷ Pädagogische Hochschulen gibt es heute nur noch in Baden-Württemberg. Sie haben inzwischen die akademischen Rechte erhalten, sind also keine Fachhochschulen.

sein, den Aufbau berufsbezogener Kompetenzen zu überprüfen, was nur möglich ist, wenn das jetzige Prüfungswesen einer Revision unterzogen wird. Das Hauptproblem aber dürfte sein, die Professorinnen und Professoren und die Mitarbeitenden der ersten Phase für die Entwicklung der Lehrerbildung dermaßen zu interessieren, dass sie sich auf die neuen curricularen Standards und die Kompetenzorientierung des Studiums ernsthaft einlassen.

In der bisherigen Organisation der Lehrerbildung sind Instrumente wie Standards und Zielsteuerung noch unterentwickelt. Ein Wechsel der Perspektive wird sich nur dann vollziehen lassen, wenn der Prozess den Beteiligten aufwändig kommuniziert wird. Zwar sind sich alle Akteure weitgehend einig, dass die Lehrerbildung gründlich reformiert werden müsse; wie das aber geschehen soll, darüber gehen die Meinungen weit auseinander. Der erwartbare Widerstand hat nicht nur mit Besitzstandswahrung und strukturkonservativem Denken zu tun, er ist daneben ebenfalls statusbedingt und könnte auch mit Vorteilen zu tun haben, die aus dem diffusen Zustand der jetzigen Lehrerbildung gewonnen werden.

Neben organisatorischen Veränderungen und Zielsteuerung ist das zentrale Element der Reform das *Durchsetzen von Verbindlichkeit*, was ohne geregelte Kooperation und Zuständigkeit nicht möglich ist. Der Aufbau von Zentren für Lehrerbildung ist dafür ein geeigneter Schritt. An diesen Zentren sollten alle Teile der Lehrerbildung beteiligt werden, was effiziente Verantwortungs- und Leitungsstrukturen erforderlich macht. Die Verteilung von Zuständigkeit setzt einen inneren Konsens voraus, der sich auf die Zielsetzungen, das Leitbild und die Organisation der „neuen Lehrerbildung“ bezieht.

Bei den bislang vorliegenden fachwissenschaftlichen Standards für die Lehrerbildung fällt auf, dass sie nicht mit den Bildungsstandards der Schule verzahnt sind. Der Unterricht auf der Zielstufe einzelner Lehrämter wird aber zunehmend mehr von Kerncurricula, Standards und Tests bestimmt sein. Die Lehrkräfte können so nicht einfach durch Standards ihrer Fachwissenschaften auf Unterricht in den verschiedenen Schultypen vorbereitet werden, sie müssen auch wissen, was auf der Zielstufe von ihnen verlangt wird. Hier werden in Deutschland vor allem die zweite Phase der Ausbildung sowie die Berufseinführung gefragt sein, wobei konkrete Modelle, wie auf standardbasierten Unterricht in den verschiedenen Lehrämtern vorbereitet werden soll, bislang nicht vorliegen. Ein weiteres Problem liegt darin, dass der Konsens über eine Vielzahl aus der einschlägigen Literatur bekannten Schlüsselqualifikationen und Kompetenzen, über welche Lehrpersonen verfügen sollten, schwierig zu erreichen sein wird. So schlägt etwa die KMK allein schon für den Teilbereich „Bildungswissenschaften“ 84 Kompetenzen vor. Zudem stellt sich die Frage, wie

sich die zumeist komplex formulierten Lehrerkompetenzen zufriedenstellend und valide diagnostizieren lassen (Beck 2006).

Da sich die Ausbildungsprogramme nicht kurzfristig ändern und der Umbau der Lehrerbildung Jahre in Anspruch nehmen wird, müssen die unmittelbaren Anstöße von der *Fort- und Weiterbildung* ausgehen.¹⁵⁸ In der internationalen Forschungsliteratur, auf die wir oben bereits verwiesen haben, wird der kontinuierlichen Fort- und Weiterbildung von Lehrkräften inzwischen hohe Aufmerksamkeit zuerkannt. Hier wird ein zentraler Faktor für die gesamte Schulentwicklung gesehen. Allerdings gilt diese These nur dann, wenn sich die Struktur der Weiterbildung grundlegend verändert. Heute nehmen Lehrkräfte im Allgemeinen kurzfristig angelegte Kursangebote wahr, die sie in aller Regel gemäß ihrem persönlichen Interesse auswählen. Dieses Interesse ist oft wenig spezifisch; auf der anderen Seite sind die Kursangebote ebenso breit wie diffus. Der Bedarf der Schule an Qualifizierung ihrer Lehrkräfte spielt dabei bislang keine Rolle.

Die vorliegenden, nicht sehr zahlreichen Erkenntnisse aus der Forschung (vgl. Messner/Reusser 2000; Terhart 2003; Lipowsky 2004) besagen, dass Maßnahmen zur Fort- und Weiterbildung nur dann Erfolg haben, wenn sie auf die jeweilige Schule und die Bedürfnisse des Klassenzimmers zugeschnitten sind. Die Lehrkräfte müssen mit neuen und überzeugenden Ansätzen zur Verbesserung ihrer täglichen Praxis konfrontiert werden, die einzelne Weiterbildung muss sich im Blick auf die Entwicklung von Unterricht und Schule als nützlich erweisen und die Kurserfahrungen müssen zur Bereicherung des Handlungsrepertoires der Lehrkräfte beitragen. Angebote, die diese Bedingungen *nicht* erfüllen, werden nicht verwertet und bleiben unbeachtet. Nach Lipowsky (2004) ist es deshalb kein Zufall, dass als Folge einer zunehmend kritischen Einschätzung der Wirksamkeit von Lehrerfortbildungen vielerorts dazu übergegangen wurde, diese vermehrt zu dezentralisieren.

„Lehrerfortbildungen werden zunehmend schulintern durchgeführt. Dahinter steht die Erwartung, dass schulinterne Maßnahmen effektiver und kostengünstiger gestaltet werden können, weil sie gezielter auf die spezifischen Erfordernisse der jeweiligen Schule und auf die Bedürfnisse der jeweiligen Kolleginnen und Kollegen reagieren können. Für die Schulen entstehen mit dieser Dezent-

¹⁵⁸ Nicht überall werden die Termini „Fortbildung“ und „Weiterbildung“ gleich verwendet. Wir verwenden sie hier austauschbar und ohne Anspruch auf klare Trennschärfe.

ralisierung auf der einen Seite größere Entscheidungs- und Gestaltungsspielräume, andererseits gewinnt für sie die Frage nach der Wirksamkeit von Fortbildungen eine besondere Relevanz“ (ebd., S. 1).

Gemäß Lipowsky (2004), der den Forschungsstand zur Frage, welche Wirkungen insgesamt von Lehrerfortbildung zu erwarten sind und welche Merkmale diese positiv beeinflussen, zusammenfasst, sind Fortbildungen vor allem dann erfolgreich, wenn sie:

- unterrichtsnah erfolgen,
- sich über einen längeren Zeitraum erstrecken,
- aus einer Kombination von Input-, unterrichtsbezogenen Trainings-, Reflexions- und Feedbackphasen bestehen,
- zum gegenseitigen Austausch und zur Kooperation – sowohl innerschulisch als auch schulübergreifend – anregen,
- einen klar definierten (fach)didaktischen Fokus und damit ein klares Ziel aufweisen,
- an den kognitiven Wissensvoraussetzungen und den Überzeugungen der Lehrkräfte ansetzen.

In diesem Zusammenhang hat sich in den letzten Jahren insbesondere der Einsatz von videografierten – eigenen und fremden – Unterrichtsbeispielen als Ausgangspunkt für die Reflexion und die Weiterentwicklung des eigenen Unterrichts als vielversprechend erwiesen (Krammer/Reusser 2005).

Die Lehrkräfte bauen ihre professionelle Strategie des Lehrens und Lernens in den ersten Berufsjahren auf. Die ersten guten Lösungen der Basisprobleme des Unterrichts sind nachhaltig wirksam, das heißt, alle Ansätze zur Weiterbildung im Beruf haben ein *entwickeltes* Know-how zu ihrer Voraussetzung, das nicht beliebig verändert werden kann und das die Lehrkräfte als überzeugend erleben. Die Anreicherung des professionellen Repertoires geschieht heute mehr oder weniger ungesteuert, ganzheitliche Strategien und strukturierte Programme mit expliziten professionsbezogenen Entwicklungszielsetzungen sind immer noch selten vorhanden und auch eine Steuerung durch evaluatives Feedback oder die Schulleitung fehlt in der Regel.

Im Blick auf das Problem der Bildungsstandards lässt sich sagen, dass mit der vorherrschenden diffusen Struktur der Lehrerfort- und Weiterbildung eine effektive Vorbereitung auf standardbasierten Unterricht nicht möglich ist. Es müssen eigene Programme entwickelt werden, die das Angebot nachhaltig verändern und die eine deutliche Priorität erhalten. Ihr Besuch wäre obligatorisch

und müsste von den Schulen veranlasst werden. Bildungsstandards müssen Eingang finden in die Entwicklungsziele einer Schule oder einer Region, denn von ihnen leitet sich der Weiterbildungsbedarf ab. Die Entwicklungsziele geben an, was die Schule in einem bestimmten Zeitraum erreichen will; wenn standardbasierter Unterricht aufgebaut werden soll, müssen Weiterbildungsmaßnahmen in diesem Bereich eine Priorität erhalten. Die Schulleitung legt in Absprache mit dem Kollegium den spezifischen Bedarf fest und organisiert die dazupassenden Angebote.

Auf Seiten der Anbieter existieren zu bestimmten Problemfeldern Standardkurse, die von den Schulen *on demand* abgerufen werden können. Dieser Service ist inzwischen in einigen deutschen Bundesländern vorhanden, wenngleich noch kaum in Verbindung mit Entwicklungszielen. Das Angebot der Weiterbildung wäre so nicht länger breit und beliebig, sondern geordnet nach Prioritäten, die sich letztlich bildungspolitisch bestimmen lassen. Prioritäten bestimmen Programme, die den Besuch von Tageskursen – immer noch die hauptsächliche Form der Weiterbildung – ablösen.

Ein vorrangiges Beispiel sind die Bildungsstandards. Im Ausland, etwa in Schweden oder England, sind seit Beginn der neunziger Jahre Teile der Weiterbildung auf die Implementation und Sicherung von Bildungsstandards umgestellt worden. Ein weiteres Beispiel sind größer angelegte fachdidaktische Weiterbildungen in Netzwerken, wie sie etwa das SINUS-Projekt darstellt. Die Lehrkräfte haben hier ein strukturiertes Angebot in Form eines Programms zur Verfügung, das sie unmittelbar nutzen können und das eine klare Priorität darstellt.

Die Wirksamkeit der Weiterbildungsmaßnahmen muss fortlaufend erhoben werden. Auch das spricht dafür, die Weiterbildung in größer angelegten Programmen zu organisieren. Das bisherige „Kurswesen“ würde damit der Vergangenheit angehören. Wiederum kommen die überzeugendsten Forschungsdaten zu diesem Punkt aus dem Ausland (vgl. Kennedy 1998; Garet/Porter/Desimore/Birman/Yoon 2001; Adey/Hewitt/Hewitt/Landau 2004). Wenn die Weiterbildung lediglich punktuell angeboten wird und nicht mit Transferstrategien verbunden ist, sind Effekte auf der Unterrichtsebene nicht nachweisbar. Auch die persönliche Anreicherung ist wenig dauerhaft. Die Rückübertragung des Gelernten auf Situationen im Klassenzimmer findet dann entweder gar nicht oder nur zufällig statt. In der Konsequenz dieser Vorschläge müssten Weiterbildungsangebote dann abgerufen werden können, wenn Bedarf besteht. Nur so kann ein Streueffekt verhindert werden und nur so ist es möglich, die Ressourcen zielgerecht einzusetzen.

Der Aufbau professioneller Kompetenz erfolgt in ausländischen Lehrerbildungssystemen mindestens in den Bereichen Primar- und Sekundarschule zunehmend standardbasiert. Die Programme der Lehrerbildung werden einem Assessment unterworfen und müssen akkreditiert werden. Beispiele hierfür sind die neuen Pädagogischen Hochschulen in der Schweiz. Dasselbe – Programmstruktur, Assessment und Akkreditierung – ist auch von der Weiterbildung zu fordern, was dann möglich ist, wenn sie einer einheitlichen Organisation unterliegt. Es gibt Versuche in der Schweiz, Standards der Ausbildung mit solchen der Weiterbildung zu verzahnen. Grundlegend hierfür sind die erwähnten Anforderungsprofile, die für die Berufsausübung insgesamt notwendig sind und die nicht mit der Erstausbildung allein erreicht werden können. Die Ausbildung sorgt für eine Initiation in den Beruf, nicht für die abschließende Kompetenz.

In diesem Sinne ist die Neustrukturierung der Weiterbildung nicht einfach eine Variante der Erwachsenenbildung, die sich auf eine Nachfrage richtet, welche sie kaum selbst beeinflussen kann. Die Weiterbildung von Lehrkräften ist vielmehr – um diesen Punkt nochmals zu wiederholen – berufsbezogen und muss Wirkung auf der Unterrichtsebene nachweisen. Daraus folgt, dass sämtliche Angebote der Weiterbildung zielorientiert verfahren müssen. Die Ziele beziehen sich auf Prioritäten wie die Implementation von Bildungsstandards und verlangen eine Programmstruktur. Die Organisation ist auch hier modulförmig.

Der Test der Ausbildung erfolgt im Wesentlichen in der Berufseingangsphase, also in den ersten drei bis fünf Jahren selbst verantworteter Praxis (Larcher Klee 2005). Diese Phase muss intensiviert werden, nicht als zusätzliche Verlängerung der Ausbildung, sondern durch besondere Betreuung und Reflexion des Berufsanfangs. Bildungsstandards und Kompetenzen der Ausbildung müssen ihren berufsbiografischen Stellenwert erhalten, also zum professionellen Habitus werden, was besondere Formen des Coachings von und mit Berufsanfängern verlangt. Hier sind die Schulleitungen sowie die Fort- und Weiterbildungen in besonderer Hinsicht gefordert. Für den Austausch zwischen und die Beratung von Novizen müssen spezielle Programme entwickelt werden, worauf in Kapitel 5 einzugehen sein wird. Damit wird verdeutlicht, welche Relevanz brauchbaren Instrumenten in der Lehrer(fort)bildung zukommt.

Bei der Reorganisation von Lehrerbildung ist neben dem „Blick nach außen“ ebenso der „Blick nach innen“ notwendig. Rolle und Status von Lehrpersonen sind je nach Schulsystem und Kultur unterschiedlich. Anreiz- und Belohnungssysteme sehen je nach Ausmaß der Marktorientierung, der Breite des Bildungsverständnisses oder der Bildungsachtung in einem Schulsystem anders aus. Insofern ist die vorschnelle Übernahme von Merkmalen anderer Bildungs-

systeme problematisch und bedarf immer auch der Reflexion der kulturellen Rahmenbedingungen (vgl. Blömeke 2006).

Probleme wie die Gestaltung der Berufseingangsphase treten in allen Lehrerbildungssystemen auf, sie sind nicht spezifisch für das deutsche Modell. Das gilt auch für die Verknüpfung zwischen den Standards der Lehrerbildung und den Bildungsstandards für die Schulen. In keinem der vier Auftragsländer ist dieses Problem bislang in überzeugender Weise bearbeitet worden. Wir verweisen auf die besondere Dringlichkeit dieser Frage. Beide Entwicklungen laufen nebeneinander her und sind allenfalls nominell verbunden. Aber wenn Lehrkräfte in Zukunft mindestens teilweise standardbasierten Unterricht erteilen, mit Rückmeldesystemen umgehen und ihre Benotungspraxis verändern sollen, dann *muss* die Ausbildung darauf reagieren, und dies zielstufengerecht.

5 Werkzeuge und Verfahren der Implementation

Nach der in Kapitel 4 dargelegten grundsätzlichen Position dieser Expertise werden Steuerungsmaßnahmen wie die Formulierung von Bildungsstandards, ihre schulrechtliche „Akkreditierung“ und testmäßige Umsetzung sowie die Rückmeldung über ihre Einlösung an Schulen nur dann Erfolg haben, wenn es gelingt, sie gleichzeitig für den Kern der schulpädagogischen Qualitätsentwicklung, die Förderung der Professionalität der Lehrkräfte und deren fachdidaktische Schulung für guten Unterricht fruchtbar zu machen. Um dieses langfristige Ziel der eingeleiteten bildungspolitischen Weichenstellung zu erreichen, bedarf es mehr als der Vision ausgearbeiteter Standards und einer bildungspolitischen Strategie zu ihrer Implementation auf den verschiedenen Ebenen des Bildungssystems. Wie professionelle Systeme allgemein, benötigen Reformen – außer Konzepten und ihnen gewogenen Personen – wirkungsvolle *Instrumente* zu ihrer Initiierung, Steuerung und nachhaltigen Umsetzung im Sinne eines Prozesses des (Um-)Lernens und des Wissenstransfers. Solche Instrumente sind auf der politischen Seite die Bildungsstandards selber sowie deren evaluative Anwendung in Tests, im Bildungsmonitoring und in weiteren Verfahren der Schulevaluation. Auf der Seite der Schularbeit und des professionellen Lernens sind es die Werkzeuge und Verfahren der Lehreraus- und -fortbildung, der Selbstevaluation, der kooperativen Entwicklung von Schulen und der Unterrichtsentwicklung.

Lange Zeit wurde die Bedeutung von Werkzeugen und Instrumenten („Tools“) in Bildung und Unterricht unterschätzt. Nicht zuletzt von einer (deutschen) Pädagogik, die ihre Praxis weniger als Profession, denn als individuelle Lehrkunst verstand oder immer noch versteht, wurden und werden „Methoden“ und wird „technologisches Wissen“ bis heute beargwöhnt und nicht selten abgewertet. Übersehen wird, dass jede Profession ihre (human-)technologischen, kognitiven, organisationalen und sozialen Werkzeuge benötigt. Desgleichen brauchen auch Reformen und Innovationen im System öffentlicher Bildung Instrumente, wenn ihre Ziele verwirklicht werden sollen. Gute Ideen, eine überzeugende Rhetorik und umsetzungswillige Personen allein reichen dazu nicht aus. Ohne geeignete Verfahren und Werkzeuge, die zu den Aufgaben und zu den Schulen bzw. den Systemebenen, auf denen sie eingesetzt werden, passen, ist wenig zu bewegen. Dass die instrumentelle

Ausrüstung von Reformprojekten im Sozial- und Bildungsbereich, die Verfügbarkeit einer professionsbezogenen Infrastruktur überhaupt zu den entscheidenden Erfolgsfaktoren zählt, ist auch in der Innovations- und Design-Forschung längst anerkannt. Gemäß dem Nobelpreisträger Herbert Simon (1981), einem der geistigen Väter der „künstlichen Intelligenz“, basiert professionelles Handeln nicht einfach auf Theorien, sondern ist auf das „Design“ von Werkzeugen – „Tools“ – angewiesen. Nicht nur Ingenieure und Ärzte brauchen solche, sondern auch die Lehrkräfte.

Die Implementation von Bildungsstandards wird im vorliegenden Kapitel deshalb auch als *werkzeugbasierter Prozess* verstanden. Die Vision einer neuen Steuerung bedarf der Einbettung in einen Kranz von Instrumenten und Verfahren ihrer Durchsetzung und der pädagogischen Qualitätsentwicklung. Diese „Tools“ müssen sich auf die verschiedenen Handlungsebenen und ihre Akteure richten.

Einige Kriterien für die Auswahl bzw. die Entwicklung (Design und Re-Design) von Verfahren und Instrumenten im Zusammenhang mit der Aufgabe der Implementation von Bildungsstandards sind:

- Zielbezug, Zieltransparenz und Fokussierung auf professionelle Kernaufgaben, Abkehr von Beliebigkeit;
- Ansetzen am Innovationspotenzial von Einzelschulen;
- Settings und Tools für individuelles, kooperatives und organisationales Lernen;
- Anschlussfähigkeit an ein konstruktivistisches Selbstverständnis vom Lernen der Lehrpersonen, aber auch der Schülerinnen und Schüler;
- Förderung der Routinisierung von Verfahren der Qualitätssicherung und -entwicklung;
- Robustheit, Verständlichkeit, Handhabbarkeit und Ökonomie;
- Potenzial für eine nachhaltige Schul- und Unterrichtsentwicklung mit Breitenwirkung;
- Möglichkeiten und Ressourcen für kontinuierliches und situiertes, handlungsbezogenes Lernen „on the job“, Möglichkeit zur Selbstprüfung von Lernfortschritten;
- Anbieten von neuen Strategien und Modellen des Unterrichts (*Best-practice*-Beispiele);
- Möglichkeiten, neue Praktiken mit bereits existierenden Routinen und Kompetenzen zu verknüpfen;

- Verfahren und Instrumente eines integrierten, von den Komponenten her aufeinander abgestimmten Systems von Qualitätssicherung und Qualitätsentwicklung;
- nicht nur Reformwerkzeuge, sondern solche, die als Professionswerkzeuge in die Alltagspraxis übernommen werden können.

Im Folgenden werden Verfahren und Instrumente eines integrierten Systems der Qualitätssicherung und Qualitätsentwicklung für die verschiedenen Ebenen und Akteure im Kontext der Implementation von Bildungsstandards dargestellt und es wird ein Standard für ihre Nutzung formuliert. Inwiefern die Instrumente greifen und von den Praktikern genutzt werden, hängt von der Gesamtstrategie der Innovationssteuerung ab, d.h. vom Mitteleinsatz und der Infrastruktur, die das System im Sinne von Anreizen, wählbaren Angeboten, Kontrollen, verbindlichen Vorgaben und Empfehlungen den Akteuren auf den unterschiedlichen Ebenen, insbesondere den einzelnen Schulen und ihrem Personal, zur Verfügung stellt.

5.1 Systemebene

„System“ meint in diesem Zusammenhang die politisch-gesellschaftliche Ebene bzw. die Ebenen, auf der oder auf denen „Bildungsmanagement“ stattfindet, das heißt, wo geplant, kontrolliert, unterstützt und im politischen Sinne Rechenschaft abgelegt wird. Im Sinne der Überlegungen von Kapitel 4 kann es sich um die *nationale* Ebene, aufgrund einer föderativen Bildungsorganisation der Staaten aber auch um die *Länder-* (Deutschland) oder *Kantonsebene* (Schweiz) handeln. Oftmals sind mehrere Ebenen mit geteilten Aufgaben beteiligt. Auch Regionen oder Bezirke können Ebenen des Bildungswesens darstellen, sofern sie über eigene Kulturen, z.B. des Qualitätsmanagements oder in Bezug auf schulorganisatorische oder curriculare Merkmale verfügen. In der Schweiz haben sich in den verschiedenen Sprachregionen zum Teil unterschiedliche Subkulturen des Bildungswesens herauskristallisiert. Neu bilden sich im Bildungswesen vermehrt auch regionale europäische und internationale Einheiten aus, zum Beispiel auf hochschulpolitischer Ebene die „Bodenseehochschule“.

Die Verfahren und Instrumente der Systemebene betreffen in erster Linie die ausgearbeiteten Bildungsstandards als Orientierungsrahmen, sodann die Lehrpläne und Curricula. Entscheidend ist, mit welchen

Verfahren und mit welchen begleitenden Interpretationsangeboten die Rückmeldungen an die Schulen gestaltet werden. Wichtigstes Prinzip ist die Fairness aller Vergleiche unter Berücksichtigung bedeutsamer Eingangs- und Kontextvariablen. Entscheidend ist auch, ob und wie die Rückmeldungen mit Beratung, Fortbildung und Förderung verknüpft werden – wie weit den Schulen geholfen wird, gemessenen Output in wirkungsvolleren Input und in verbesserte Lehr-Lernprozesse rückzuverwandeln.

Dies setzt eine nicht triviale Lehr-Lerndiagnostik voraus, die die Lehrpersonen jedoch nicht leisten können und die aufgrund der Aussagekraft allein von Bildungsmonitoring-Daten auch nicht befriedigend zu leisten ist. Deshalb müssen die Lehrpersonen die Funktion unterschiedlicher Tests verstehen und es muss ihnen Hilfe angeboten werden, damit sie aus Evaluationen mit anderen Mitteln didaktischen und individualdiagnostischen Nutzen für ihre tägliche Arbeit ziehen können. Daraus folgt, dass es neue Formen der Beratung braucht (z.B. Evaluationsberatung). Ebenfalls muss sich die Schulaufsicht den neuen Gegebenheiten anpassen. Wichtig ist schließlich auch, dass die neuen Medien, die neuen Informations- und Kommunikationstechnologien für einen systemweiten Erfahrungsaustausch über Bildungsstandards genutzt werden und nicht nur zur administrativen Kommunikation und Steuerung.

5.1.1 Ausgearbeitete Bildungsstandards und darauf bezogene Tests

Bildungsstandards und darauf bezogene Testinstrumente bilden aus der Sicht der Politik *das* zentrale Werkzeug der Reform. Dessen Kernfunktion ist die „Orientierung und ErgebnISRückmeldung für das staatliche Schulwesen“ (Klieme/Avenarius/Blum et al. 2003, S. 123), auf das sich die Politik verständigt hat. Mit der Qualität ausgearbeiteter Standards steht und fällt die neue Politik einer ertragsorientierten Steuerung des Bildungswesens. Im deutschen Sprachraum hat die Entwicklung von Standards gerade erst begonnen. Bei ihrer Formulierung handelt es sich auf der technischen Ebene um die anspruchsvolle Aufgabe von Experten; auf der Inhalts- und Zielebene setzt ihre verbindliche Festlegung einen demokratisch legitimierten Diskurs voraus. Was bislang vorliegt, ist in fachöffentlichen Diskursen noch kaum breit abgestützt und befriedigt vor allem fachdidaktisch und psychometrisch noch längst nicht.

Es herrscht großer Entwicklungs- und ebenso beträchtlicher Forschungsbedarf. Vor allem die noch weitgehend fehlende Kompetenzmodellierung und die Frage, inwieweit der formulierte Anspruch in den Fächern eingelöst werden kann, stellen ein Problem dar.

Ein weiteres Problem, das sich nicht nur für die Schweiz stellt, betrifft ein einheitliches Verständnis der Standards über Sprach- und regionale Kulturgrenzen hinaus: Sowohl für den internationalen Vergleich als auch für die Entwicklung von Bildungsstandards in einem mehrsprachigen Land sollte ein übergreifendes Verständnis von Standards und Kompetenzen bestehen. Das heißt, Begriffe müssen definiert und übersetzt werden, so dass alle Beteiligten darunter das Gleiche verstehen und die Kompetenzen auch in allen Sprachen entsprechend gemessen und ausgewiesen werden können. Die Kantone, aber auch die Sprachregionen, ja sogar die Landesgrenzen sollten irrelevante Kategorien sein, wenn es um die Beurteilung von Kernkompetenzen im Sinne elementarer fachlicher Allgemeinbildung und um die Schaffung von Transparenz geht. Die Auftragsländer sollten Anstrengungen zu einem gemeinsamen Verständnis von Grundbegriffen unternehmen.

Bildungsstandards

Bildungsstandards stellen ein Element eines umfassenden Systems der Qualitätssicherung dar, welches Informationen für die Weiterentwicklung des Bildungssystems bereitstellen soll. Sie geben eine Antwort auf die Frage, wie Leistungserwartungen in verbindlicher Weise im Bildungssystem kommuniziert und bezüglich ihrer Geltung überprüft werden können. Für die Schulen stellt die Überprüfung von Bildungsstandards eine externe Evaluation dar, die sich durch klare Maßstäbe (die einen Vergleich innerhalb bestimmter Alterskohorten erlauben) und standardisierte Leistungsrückmeldungen von aller bisherigen Praxis unterscheidet. Die Entwicklung von Bildungsstandards wird in den deutschsprachigen Ländern maßgeblich von PISA beeinflusst. Die Methodik des internationalen Schulleistungsvergleichs bildet vor allem für das Vorgehen in einem föderalistischen, mehrsprachigen Land eine wichtige Grundlage für die Entwicklung und die Überprüfung von Standards. Die bis heute öffentlichen Bildungsstandards im deutschsprachigen Raum sind im Anschluss an PISA mehrheitlich in den letzten vier bis fünf Jahren entwickelt und festgelegt worden. In Deutschland sind dies beispielsweise die Bildungsstandards für den Primarbereich (Jahrgangsstufe 4), für den Hauptschulabschluss (Jahrgangs-

stufe 9) und für den Mittleren Schulabschluss (Jahrgangsstufe 10) der Kultusministerkonferenz (KMK) oder die Bildungsstandards und Niveaue Konkretisierungen im Rahmen der Bildungsplanreform in Baden-Württemberg, in Österreich die nationalen Bildungsstandards des Bildungsministeriums für die 4. und 8. Klasse. Die erwähnten Beispiele sind, in verschiedenen Publikationen ausführlich diskutiert und kritisiert worden.¹⁵⁹

Vielen Kritikpunkten ist dabei wenig entgegenzuhalten; beispielsweise der Kritik, dass die Bildungsstandards der KMK in ihrer ersten Fassung als Listen von umgeschriebenen Lehrplanzielen erscheinen und dass die für Bildungsstandards als unabdingbar erachtete Kompetenzmodellierung entweder fehlt oder nur zum Teil in überzeugender Weise gegeben ist (vgl. zum „state of the art“ im Schulfach Mathematik z.B. Ehmke/Leiss/Blum/Prenzel 2006). Würden die bisherigen Standards nicht weiterentwickelt, so ergäbe sich für die Testentwicklung das Problem, dass Standards getestet würden, welche die selbst gesetzten Qualitätsansprüche gar nicht erfüllen und zu deren Aufbau im Unterricht die Schulen bisher keine Gelegenheit hatten.

Da Bildungsstandards nicht „einfach eine neue Vorgabe, die man ohne Sanktionen ignorieren kann“, darstellen, sondern definitiv überprüft werden sollen, ist auf jeden Fall von ihrer Wirksamkeit auszugehen. Das heißt, Lehrpersonen werden sich nur sehr schwer der Notwendigkeit entziehen können, ihren Unterricht nach besten Kräften so zu gestalten, dass ihre Schülerinnen und Schüler die als verbindliche Leistungsaufträge kommunizierten Kompetenzerwartungen erfüllen und in den Tests gut abschneiden. Wie *produktiv* der Steuerungseinfluss von Standards und Tests für Lernen und Leistung von Schülerinnen und Schülern und für das Schulklima sein wird, hängt deshalb (neben vielen weiteren Faktoren, von denen in dieser Expertise ausführlich die Rede ist) in substantieller Weise davon ab, inwiefern sich die Standards in für Lehrpersonen einsehbarer Weise auf Kernbereiche und Grundkompetenzen eines Fachs beziehen und inwiefern sie ihnen – aber auch den Eltern und Schülerinnen und Schülern – eine Orientierung und eine deutliche

¹⁵⁹ Für Deutschland: *Die Deutsche Schule*, 8. Beiheft 2004, Bildung und Standards. Zur Kritik der „Instandardsetzung“ des deutschen Bildungswesens. Für Österreich: *Erziehung und Unterricht*, 7-8, 2004. Themenschwerpunkt Bildungsstandards.

Vorstellung darüber vermitteln, was die Erwartungen an das Lernen und seine Ergebnisse sind.

Dies setzt eine hohe *inhaltliche Qualität der Bildungsstandards und der darauf bezogenen Tests* voraus. Das heißt, nur knapp und verständlich formulierte Standards, die sich auf begründbare und nachvollziehbare Kernbereiche und Grundkompetenzen eines Fachs beziehen, hohen fachdidaktischen und methodischen Qualitätskriterien genügen, an Beispielen und Aufgaben illustrierbar sind sowie pädagogisch wertvoll und umsetzbar erscheinen, werden eine Chance haben, von der Profession als verbindlicher Referenzrahmen auch langfristig internalisiert zu werden.

Dass solche inhaltliche und methodische Qualität bislang in keinem Schulfach und Auftragsland als vorzeigbare „best practice“ verwirklicht ist, hängt damit zusammen, dass die Aufgabe neu ist und es sich sowohl bei der Ausarbeitung von Standards als auch bei der Entwicklung von Tests, welche tatsächlich das messen, was sie messen sollen, um eine anspruchsvolle wissenschaftliche Aufgabe handelt, die neben fachdidaktischer und kognitionspsychologischer auch methodologische Kompetenz erfordert. Damit ein solcher Prozess möglich ist, braucht es eine entsprechende Infrastruktur. In Deutschland wurde dazu mit der Gründung des bundesweit tätigen „Instituts zur Qualitätsentwicklung im Bildungswesen“ (IQB) als einer wissenschaftlich qualifizierten Entwicklungs- und Koordinationseinrichtung ein wesentlicher Schritt getan. Zu den Kernaufgaben des IQB gehören nach dessen Leiter (Köller 2005) die Weiterentwicklung der von der KMK verabschiedeten Bildungsstandards, deren methodische Präzisierung sowie die Überprüfung ihrer Erreichung durch Schülerinnen und Schüler. Konkret geht es um die:

- Erstellung großer Aufgabensammlungen,
- Normierung und Überprüfung der Bildungsstandards mittels empirischer Studien,
- Bereitstellung von Aufgaben für die Länderprogramme der flächendeckenden Vergleichsarbeiten,
- Bereitstellung von Aufgaben für Schulen zum Zweck der internen Evaluation,
- wissenschaftliche Begleitung der Implementation der Standards.

In den anderen Auftragsländern besteht, so sie den von der Klieme-Expertise (Klieme/Avenarius/Blum et al. 2003) gesetzten und vom

IQB adoptierten Anforderungen nachkommen wollen, noch deutlicher Handlungsbedarf. Zum Beispiel fehlt für die mehrsprachige Schweiz, die mit dem HarmoS-Projekt im Prinzip eine ähnliche Richtung wie in Deutschland einschlägt, bislang die Infrastruktur für die Standard- und Testentwicklung.

Die entscheidenden Anforderungen und Gütekriterien, die an die Entwicklung von Standards und die darauf bezogenen Tests zu richten sind, werden in der Klieme-Expertise (Klieme/Avenarius/Blum et al. 2003, S. 119ff.) wie folgt festgehalten:

„Der Lernbereich bzw. das Fach, für das ein Standard entwickelt werden soll, muss bildungstheoretisch verankert werden. Die Grundkonzepte der Bezugsdisziplin, die zentralen Orientierungen der jeweiligen ‚Weltsicht‘ müssen herausgearbeitet werden. Das System relevanter Kompetenzen mit ihren Teilaspekten, Stufen und Entwicklungslinien stellt den Kern in der Bestimmung der Bildungsstandards dar“ (ebd., S. 121).

Als „pièce de résistance“ bei der Entwicklung von kompetenzorientierten Standards und als Basis von Kompetenzmessungen hat sich dabei die in einem psychologisch-theoretischen Sinne verstandene *Kompetenzmodellierung* erwiesen. Jeder Kompetenzmessung mit Hilfe von Testaufgaben liegt eine Strukturannahme der zu messenden Kompetenz zugrunde. Das heißt, die einer bestimmten Kompetenz innewohnenden Dimensionen müssen aufgrund theoretischer Annahmen und Überlegungen ermittelt werden.¹⁶⁰ Kompetenzmodelle beschreiben – idealiter – als Ergebnisse von Analysen domänenspezifischer Anforderungen und Aufgabenlösungsprozessen die spezifischen Dimensionen und Operationen, welche in einem festgelegten Bereich eines Leistungskontinuums eine Rolle spielen. Weiter können sie genutzt werden, um Mindest- oder Regelstandards zu definieren, welche von den Lernenden erreicht werden sollen. Auf diese Weise sollen standardbasierte Tests

¹⁶⁰ Von theoretischen Überlegungen inspirierte Kompetenzmodelle hat es schon früher gegeben. Den wohl bekanntesten Versuch, das Spektrum der geistigen Fähigkeiten des Menschen strukturell zu beschreiben und in eine kognitionspsychologisch begründete Stufung zu bringen, stellen die auf Gagné und Bloom zurückgehenden, taxonomischen Lern(ziel)hierarchien dar (vgl. Leutner 2006).

nicht einfach eine allgemeine „Messlatte“ anlegen, sondern die Unterscheidung von in Stufen zunehmender Schwierigkeit und Komplexität darstellbaren Kompetenzdimensionen ermöglichen, die es erlauben, individuelle Profilbildungen von Schülerleistungen bezüglich theoretisch und/oder empirisch bedeutsamer Dimensionen vorzunehmen.

Die Klieme-Expertise hält zur kompetenzorientierten Formulierung von Bildungsstandards fest:

„Bildungsstandards, wie sie hier vorgeschlagen werden, stützen sich auf Kompetenzmodelle, die in Zusammenarbeit von Pädagogik, Psychologie und Fachdidaktik entwickelt werden müssen. Ein solches Kompetenzmodell unterscheidet Teildimensionen innerhalb einer Domäne (also z.B. Rezeption und Produktion von Texten, mündlichen und schriftlichen Sprachgebrauch), und es beschreibt jeweils unterschiedliche Niveaustufen auf solchen Dimensionen. Jede Kompetenzstufe ist durch kognitive Prozesse und Handlungen von bestimmter Qualität spezifiziert, die Schüler auf dieser Stufe bewältigen können, nicht aber Schüler auf niedrigeren Stufen. Zum Bildungsstandard gehört, dass für einzelne Jahrgänge festgelegt wird, welche Stufen die Schülerinnen und Schüler erreichen sollen“ (ebd., S. 22).

Im Gegensatz zu PISA – wo ebenfalls von einer kompetenzorientierten Grundbildung ausgegangen wird, „Kompetenzskalen“ („described proficiency scales“) jedoch post hoc aus Itemkennwerten rekonstruiert werden –, sollte bei Bildungsstandards im Sinne eines *deduktiven und systematischen* Vorgehens das hierarchisch organisierte Kompetenzmodell vorgängig festgelegt werden. Kompetenzmodellierungen erfordern streng genommen eine theoretisch fundierte Bestimmung mehr oder weniger komplexer domänenspezifischer Kompetenzbereiche, welche durch noch spezifischere Teilkompetenzen aufgeschlüsselt werden. Die Aktivierung und der kumulative Aufbau dieser Teilkompetenzen erfolgt sodann über Aufgaben, welche sich auf fachspezifische Leitideen beziehen, die das Fachcurriculum spiralförmig durchziehen (vgl. Rupp/Leucht/Hartung 2006).

Wie weit dieser kompetenzdiagnostische Anspruch der Dimensionalisierung und Stufung von (Teil-)Fähigkeiten nicht nur testtheoretisch, sondern auch auf der Basis denkpsychologischer und fachinhaltlich-didaktischer Prozessvorstellungen eingelöst werden kann, muss als derzeitige Knackpunkt-Frage gelten, die man angesichts der großen Komplexität kognitiver Leistungen und damit einer Kompetenz – verstanden

als ein relationales Gefüge von Wissens-, Könnens-, Fähigkeits-, Erfahrungs- und Motivationskomponenten – auch skeptisch beantworten kann. Dies umso mehr, als es sich beim Konzept der Kompetenzen und ihrer Modellierung mit psychometrischen Mitteln um ein offenes Forschungsgebiet handelt, an dem international zwischen Kognitionspsychologie, Psychometrie und Erziehungswissenschaft/Fachdidaktik interdisziplinär gearbeitet wird (vgl. Klieme/Leutner 2006). Kommt dazu, dass man in der Diskussion um inhaltliche Kompetenzen für einen Bildungsbereich eigentlich Kompetenzentwicklungsmodelle¹⁶¹ benötigt, diese aber aus Mangel an längsschnittlichen Validierungsstudien nur in seltenen Fällen zur Verfügung stehen (Rost 2006). Als paradigmatischer Bereich dafür, was derzeit in Bezug auf eine dimensionale Kompetenzmodellierung und Kompetenzstufung möglich und in der praktischen Umsetzung auch machbar ist, kann die Rahmenkonzeption der Bildungsstandards in Mathematik und deren Umsetzung in Testverfahren gelten (Ehmke/Leiss/Blum/Prenzel 2006). Auch wenn die Arbeit an dieser Konzeption zu einem erfolgreichen Feldtest-Abschluss gekommen ist, bleibt vorerst offen, inwieweit es gelingen wird, sie in analoger Weise auch auf andere Fächer (Naturwissenschaften, Deutsch, Fremdsprachen) zu übertragen.

Ob sich der Anspruch im strengen Sinne einlösen lässt, ist eine nicht triviale Frage, setzt das Verfahren der Modellierung von Kompetenzabstufungen doch voraus, dass die zu modellierenden Kompetenzdimensionen in eindeutiger Weise durch voneinander abgrenzbare, in definierten Abstufungen fassbare Aspekte geprägt sind. Für die inhaltlich sehr vielfältigen Schulfächer scheint dies aber nur begrenzt plausibel, da eine Vielzahl von zu erwerbenden Konzepten und Fertigkeiten bei der Beherrschung einer Fachkompetenz eine Rolle spielen. Ob bzw. wie weit eine Fachkompetenz beherrscht wird, dürfte sodann auch davon abhängen, ob die zugrundeliegenden Teilkonzepte und Fertigkeiten im Unterricht behandelt wurden. Deren Behandlungsabfolge kann zwischen Schulen/Regionen variieren, so dass kein für alle (oder zumindest viele) Personen einheitliches dimensionales Kompetenzmuster entsteht. Anders als etwa im Falle des allgemein psychologischen und relativ in-

¹⁶¹ Es handelt sich dabei um Modelle, in denen „den Komponenten einer Kompetenzstruktur eine zeitliche Abfolge in ihrem Erwerb beigemessen“ wird (Rost 2006, S. 8).

haltsarmen Konstrukts der Intelligenz ist das Kompetenzfeld der Mathematik – und mehr noch jenes der Naturwissenschaften und der (Fremd-)Sprache(n) – nicht nur an zum Teil heterogene fachliche Kulturen, sondern auch eng an Vorgaben der Institution Schule geknüpft.¹⁶² Schon aus Gründen der Forschungsökonomie und der Vielfalt von fachkulturell und institutionell geprägten Kompetenzdimensionen wird es kaum möglich sein, zu allen wünschbaren Dimensionen kontextarme Stufenmodelle – etwa im Sinne von Piaget – zu formulieren (wobei sich bekanntlich Piagets Modelle mittlerweile ebenfalls als nicht kontextfrei erwiesen haben).

Empirisch vertragen sich „reine“ (kontextfrei gedachte) Stufenmodelle zudem schlecht mit in der Regel weitgehend kontinuierlichen Kompetenzdimensionen, wie sie empirisch in Tests nachgewiesen werden. Es gilt auf jeden Fall die Umkehrung: Zu einem kompetenztheoretisch umschriebenen, bestimmten Prozess lassen sich Aufgaben mit einem Schwierigkeitsgrad finden, die weit über die entsprechende Kompetenzstufe der Beherrschung des Prozesses hinausgehen. Zum Beispiel lassen sich zum simplen Reproduzieren von Faktenwissen Aufgaben mit beliebiger Schwierigkeit konstruieren, indem „exotisches“ Wissen abgefragt wird. Das widerspricht zwar nicht der Stufendefinition. Man wird jedoch deshalb in der Praxis eine Verteilung von Aufgaben auf der Schwierigkeitsdimension finden, die nur bedingt der theoretisch postulierten Stufenabfolge entspricht.

Unabhängig von der Wahl des Kompetenzmodells muss auf Konzeptionsebene entschieden werden, in welcher „Detailtiefe Rückmeldungen an Schülerinnen und Schüler gewünscht sind“ (Rupp, Leucht, Hartung 2006, S. 212). Entsprechend muss im Vorfeld der Einführung von Tests klar bestimmt werden, „welche Aussagen über welche Form der Kompetenzstruktur mit den Ergebnissen der jeweiligen Untersuchung ermöglicht werden sollen“ (ebd., S. 213).

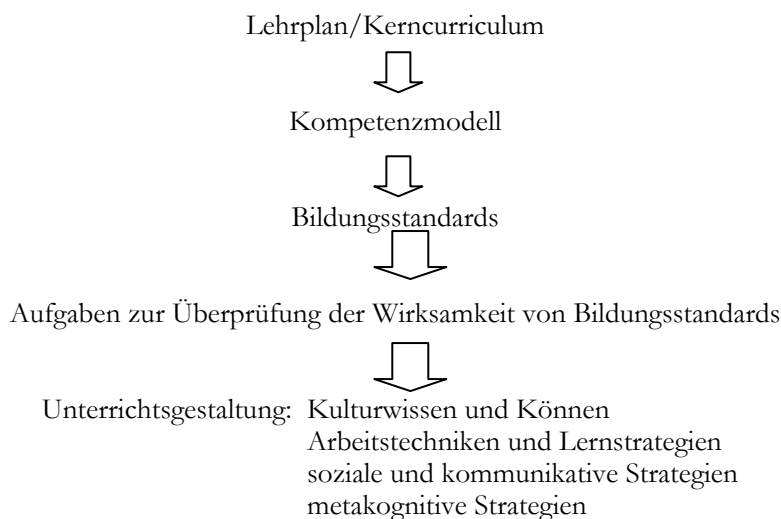
Im DFG-Schwerpunktprogramm „Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen“ (Klieme/Leutner 2006) soll deshalb eine Kooperation zwi-

¹⁶² Nun könnte man einwenden, dass die variantenreiche Behandlungsabfolge bzw. die Aufbaustruktur von fachlichen Kernkonzepten gerade eingeschränkt werden müsste im Sinne von weniger Beliebigkeit. Damit würden aber in nicht ohne Weiteres zumutbarer Weise die didaktischen Freiheiten tangiert bzw. es würde zu einer unerwünschten Standardisierung der Unterrichtsabläufe beigetragen.

schen Experten, in Bezug auf die domänenspezifische Seite der Kompetenzmodellierung und Experten auf dem Gebiet des Messens und Testens angestrebt werden. Dies geschieht mit dem Ziel, valide und faire Instrumente zu schaffen, welche sowohl der Kompetenzdiagnostik als auch dem Assessment von Bildungssystemen dienen. Ergebnisse dazu stehen noch aus, doch dürfte mit Hilfe dieses Ansatzes die in Bezug auf die Schaffung leistungsfähiger Kompetenzmodelle notwendige Verbindung zwischen theoretischen und domänenspezifischen Ansprüchen einerseits und messtheoretischen Ansprüchen andererseits realisierbar werden.

Zusammenfassend kann gesagt werden, dass Bildungsstandards und ihre Implementation in ihrem Gesamtzusammenhang betrachtet werden müssen (vgl. Abbildung 7). Ausgehend von einem Kerncurriculum bzw. Lehrplan muss die Entwicklung von Kompetenzmodellen erfolgen, welche ihrerseits die Basis für praktikable Standards und darauf abgestimmte Testaufgaben bilden, und die den Blick auf längerfristige, kumulativ angelegte Lernprozesse fördern. Daran schließen sich verschiedene, die Ebene des Unterrichtens betreffende Aspekte an, die in den folgenden Abschnitten ausführlich thematisiert werden.

Abbildung 7: Bildungsstandards in ihrem Gesamtzusammenhang



Quelle: Dubs (2005b).

Tests

Obgleich Standard- und Testentwicklung miteinander verschränkte Prozesse sind, ist eine ausgearbeitete theoretische Rahmenkonzeption für Bildungsstandards, welche eine *Modellierung der Struktur, der Stufung und der Entwicklung von Kompetenzen erlaubt*, das eine; kriteriumsorientierte Tests, welche im Sinne der ihnen eigenen psychometrischen Gütekriterien zuverlässig diese *Kompetenzen in konkrete empirische Messverfahren übertragen und tatsächlich jene Leistungen erfassen, die gemeint sind*, sind das andere. Dass sich die modellierungs- und testtheoretischen Kriterien einlösen lassen, ist aufgrund der Anforderungen an die Komplexität des Gegenstandes und die zyklischen Prozesse der Modellierung und empirischen Erprobung in Tests keineswegs selbstverständlich und muss hier als zentrale Aufgabe und als Gütekriterium des Anspruchs an gute Bildungsstandards und ihre Testung nochmals betont werden.

Da pädagogische Konzepte und Bildungsstandards u.a. auf tiefes Verstehen und Problemlösen zielen, ist es unabdingbar, dass auch die Tests solche Kompetenzfacetten valide und reliabel erfassen. Das stellt hohe Anforderungen an die Aufgabenentwicklung. Diese Forderung bedeutet nicht den Verzicht auf Multiple-Choice-Aufgaben. Das Format der Aufgaben bzw. die Mischung von Aufgaben unterschiedlicher Formate in einem Test muss vielmehr ein optimales Verhältnis von vollständigem Erfassen der Kompetenz, hoher Standardisierung, ökonomischer Durchführung und Unterrichtsnähe bzw. Kontext-Ange messenheit finden. Mit Multiple-Choice-Aufgaben lassen sich zwar Kreativität und kommunikative Kompetenzen nicht erfassen, hohe Denk- und Verstehensleistungen dagegen schon. Ökonomische, standardisierte Tests sind somit nicht grundsätzlich abzulehnen, im Gegenteil. Wo es die Art der geforderten Kompetenzen nötig macht, z.B. bei Denk- und Arbeitsweisen oder in der mündlichen Kommunikation sowie bei anderen „weichen“ Bildungszielen, müssen in gewissem Umfang jedoch (auch) aufwändigere Kontroll- und Testformen, z.B. authentische Performanztests (vgl. Shavelson/Ruiz-Primo 1999) eingesetzt werden bzw. ist eine Verständigung auf Indikatoren notwendig, die sich auch für Formen der kontinuierlichen Berichterstattung im Rahmen einer periodischen Rechenschaftslegung von Schulen eignen.¹⁶³

¹⁶³ Zum Beispiel: Wie viele Schüler/innen erlernen ein Musikinstrument, engagieren sich für andere Menschen [...] (aus: GEW 2004b, S. 217).

Eine weitere auf Tests bezogene Forderung ist diejenige nach einer Investition in deren Weiterentwicklung. Neuere Entwicklungen in der kognitiven Psychologie und der Messtheorie eröffnen in Verbindung mit der computergestützten Durchführung neue Perspektiven, wie gestützt auf ein Modell der kognitiven Prozesse und des Lernens komplexe Kompetenzen zunehmend besser erfasst werden können (National Research Council NCR 2001b). Die bereits erfolgte und künftige Entwicklung computergestützter interaktiver Medien macht es auch auf theoretisch weniger elaborierter Grundlage möglich, Tests mit vielfältigeren Aufgaben als heute üblich zu entwickeln und ökonomisch einzusetzen (hingewiesen sei hier auf die Förderinitiative des BMBF zur Entwicklung einer technischen Plattform für netzbasierte Kompetenzdiagnostik).

Des Weiteren ist zu beachten, dass es *sehr unterschiedliche Arten und Funktionen von Tests* gibt, worüber auch Lehrpersonen ein Basiswissen vermittelt werden sollte. Die Spannweite reicht dabei von internationalen Vergleichsuntersuchungen über nationale oder länderspezifische Vergleichsarbeiten bis zu von Lehrpersonen abrufbaren, im Unterricht selbst nutzbaren adaptiven Klassen- und Fachtests. All diesen Tests ist gemeinsam, dass sie sich an fachinhaltlichen Kriterien orientieren und dass ihre Interpretation ebenfalls kriteriumsorientiert erfolgt. Bei den verwendeten Begriffen „Lernstandserhebungen“, „Vergleichsarbeiten“, „Orientierungsarbeiten“, „Kompetenztests“ oder „Standardisierte Schulleistungstests“ handelt es sich dabei nicht um wissenschaftlich eindeutig definierte Fachbegriffe, sondern um Begriffe aus einer recht heterogenen Praxis der deutschen Bundesländer. Tests werden zudem für Forschungszwecke, für den Vergleich von Schulen, Schulklassen oder Schülerinnen und Schülern, für internationale Schulleistungsvergleiche, für diagnostische Zwecke usw. eingesetzt und sind je nach Zweck unterschiedlich konzipiert. Welche Begriffe für die Durchführung von Tests verwendet werden, hängt meist mit dem Ziel zusammen, das angestrebt wird.¹⁶⁴ Grundsätzlich gilt, dass sich je nach Art, Umfang und Anlage einer Untersuchung mit Hilfe von Tests unterschiedliche Arten

¹⁶⁴ Scheerens/Glas/Thomas (2003, S. 34ff.) unterscheiden bei Evaluationen, die auf der Leistungsmessung bei Schülerinnen und Schülern basieren, zwischen (1) National Assessment Programs, (2) International Assessment Programs, (3) School Performance Reporting, (4) Student Monitoring Systems, (5) Assessment-based School Self Evaluation und (6) Examinations.

von Informationen ergeben, die wiederum zu unterschiedlichen Steuerungsmöglichkeiten führen (Terhart 2002, S. 72). Flächendeckende Vergleichsarbeiten können gleich wie zentrale Abschluss- oder Aufnahmeprüfungen für Informationen im Rahmen eines Bildungsmonitorings, internationale Schulleistungsvergleiche oder Lernstandserhebungen wie PISA oder TIMSS auch für regionale Vergleiche oder für den Vergleich von Schulen genutzt werden (Deutsches PISA-Konsortium 2003). Grundsätzlich sind es die Stichprobe, das Testdesign und die Testtheorie, welche über den Nutzen und die Möglichkeiten der eingesetzten Tests entscheiden.

Aus der Sicht von Lehrkräften und Schülerinnen und Schülern sind Tests jedoch nicht einfach nur Leistungsmessungen, sondern selbst Lernobjekte, die von den Lernenden mehr oder weniger ernst genommen werden im Sinne eines produktiven oder eher unproduktiven Lehrens und Lernens „to the test“. Sowohl Lehrkräfte als auch Schülerinnen und Schüler richten sich kognitiv und motivational darauf ein, dass und wie getestet wird. Ohne klare Zweckbestimmung und ohne effektive Rückmeldesysteme besteht die Gefahr der Ritualisierung und des „Coping“ (des reinen Bewältigens und Hinter-sich-Bringens der Testsituation mit allen Mitteln), was kaum produktive Auswirkungen zeitigt. Wenn regelmäßige Tests die Qualitätsentwicklung von Schulen steuern sollen, dann müssen sie mit klaren inhaltlichen Anforderungen verknüpft sein, kognitive Ansprüche stellen und sich an der Leistungsfähigkeit und am Unterrichtsprogramm der Zielgruppe orientieren. Das verlangt:

- Sensibilität für den tatsächlich geleisteten Unterricht: Der Test muss erfassen, dass der gute Unterricht den Unterschied macht;
- hohe Qualität der technischen Seite des Testens;
- langfristige Anlage der Tests;
- klare und exklusive Ziele/Standards, keine vagen Additionen;
- Evidenz der Eignung des Testes für verschiedene Zwecke und jede Subgruppe von Schülerinnen und Schülern (Baker 2004b, S. 162).

Tests müssen somit in einem pädagogisch verantwortlichen Sinne sensitiv sein für die Zielgruppen, bei denen sie angewendet werden. Ohne Berücksichtigung des sozialen und kulturellen Hintergrundes der Schülerinnen und Schüler geht die Testkonstruktion an den Realitäten vorbei (so schon Porter 1993). Problematisch ist es, wenn in der Anlage von Tests von der Prämisse ausgegangen wird, dass es keine individuellen Unterschiede im Entwicklungsstand der Schülerinnen und Schüler gibt,

sondern dass alle das gleiche Ziel erreichen sollten, was nicht der Fall ist und sein kann (Lam 2004).

Fairness des Vergleichs ist deshalb eine zentrale Anforderung an alle Leistungstests (Nachtigall/Kröhne 2006). Die Kernfrage sollte auch aus pädagogischer Sicht nicht die sein, wie Tests vermieden, sondern wie sie in einem fairen Sinne durchgeführt werden können. Das Vertrauen in Tests wird nur dann wachsen, wenn Regeln der Fairness formuliert sind und strikt eingehalten werden. Amerikanische Wissenschaftsvereinigungen¹⁶⁵ haben *Standards for Educational and Psychological Testing* formuliert, die fairen Testen fördern und unintendierte Folgen verhindern sollen. Diese mutatis mutandis auch für die Auftragsländer gültigen Regeln umfassen:

- Any decision about a student's continued education, such as retention, tracking, or graduation, should not be based on the results of a single test, but should include other relevant and valid information.
- When test results substantially contribute to decisions made about student promotion or graduation, there should be evidence that the test addresses only the specific or generalized content and skills that students have had in opportunity to learn. For tests that will determine a student's eligibility for promotion to the next grade or for high school graduation, students should be granted, if needed, multiple opportunities to demonstrate mastery of materials through equivalent testing procedures.
- When a school district, state, or some other authority mandates a test, the ways in which the test results are intended to be used should be clearly described. It is also the responsibility of those who mandate the test to monitor its impact, particularly on racial and ethnic-minority students of lower socioeconomic status, and to identify and minimize potential negative consequences of such testing.
- In some cases, special accommodations for students with limited English proficiency may be necessary to obtain valid test scores. If students with limited English skills are to be tested in English, their test scores should be interpreted in the light of their limited English skills.

¹⁶⁵ American Psychological Association, American Educational Research Association und das National Council on Measurement in Education.

- Likewise, special accommodations may be needed to ensure that test scores are valid for students with disabilities (American Psychological Association 2001, S. 2).

Lehrpersonen müssen darüber hinaus in der Lage sein, Tests bezüglich ihres Aussagewerts verstehen und interpretieren zu können. Wie dargestellt, ist die Entwicklung standardbasierter Tests eine herausfordernde Aufgabe, deren Lösung laufend aktualisiertes wissenschaftliches Know-how erfordert. Auch die Interpretation von Testergebnissen stellt alles andere als eine triviale Aufgabe dar. Es ist davon auszugehen, dass insbesondere Lehrpersonen im Primar- und Grundschulbereich nur in seltenen Fällen über fundierte Kenntnisse in Statistik verfügen. Um den angestrebten Nutzen aus rückgemeldeten Testergebnissen ziehen zu können, müssen sie in die Lage versetzt werden, elementare statistische Informationen, wie Häufigkeitsverteilungen, Normalverteilung, Mittelwerte und Dispersionsmaße sowie Korrelationen in ihrer Funktionsweise zu verstehen und für ihre Zwecke sinnvoll anwenden zu können. Zudem müssen die Ergebnisrückmeldungen derart erfolgen, dass sie den Lehrpersonen viele und möglichst konkrete Anknüpfungspunkte an das normale unterrichtliche Handeln bieten (vgl. Humpert/Hauser/Nagl 2006; Hosenfeld/Schrader/Helmke 2006).

5.1.2 Bildungsmonitoring

Eine als Bestandteil der Steuerung des Bildungswesens etablierte systematische Beobachtung von Lernergebnissen im Kontext ihrer Rahmenbedingungen gehört heute zu den Merkmalen und Routinen der meisten erfolgreichen Bildungssysteme. Die in PISA überwiegend europäischen Länder etwa verfügen mehrheitlich über Systeme, in denen professionelle nationale Agenturen für die Evaluation verantwortlich zeichnen und aufgrund dieser Rückmeldungen Ressourcenzuweisungen in Verbindung mit gezielter Unterstützung der im Bildungsprozess involvierten Akteure vornehmen.¹⁶⁶ So etwa dienen zentrale Evaluationen mit

¹⁶⁶ Klieme/Avenarius/Döbert et al. (2003) halten in ihrer vertiefenden Analyse der Bildungssysteme ausgewählter PISA-Teilnehmerstaaten fest: „Feststellbar ist, dass in keinem der untersuchten Staaten am Primat von Input-Steuerung und Selbstreferenzialität festgehalten wurde. In den Referenzstaaten ging die Übertragung von operativen Steuerungskompetenzen auf die Einzelschule stets mit der Etablierung von Strukturen der Output-

daran anschließenden nationalen Untersuchungen in Schweden dazu, Chancengerechtigkeit im Bildungssystem zu verbessern respektive zu erhalten. Aufgrund der Ergebnisse sollen diejenigen Schülerinnen und Schüler identifiziert werden, welche die national gesteckten Ziele nicht erreichen, um gegebenenfalls notwendige Fördermaßnahmen zu veranlassen. Die Vorgabe von Kompetenzzielen über verbindlich gesetzte Bildungsstandards wird nur dann Auswirkungen auf die Schulpraxis und die Erreichung dieser Ziele haben, wenn diese im Rahmen institutionell eingeführter Bestandesaufnahmen der Leistungen des Systems auch systematisch überprüft und an die Steuerungsebene rückgemeldet werden. Über die Teilnahme an internationalen Assessments (TIMSS, PISA) hinaus streben die Auftragsländer deshalb ein der Systembeobachtung dienendes nationales Bildungsmonitoring an. Unter dem Begriff „Bildungsmonitoring“ wird die systematische, mittels objektiver Verfahren wie Tests, Fragebögen und statistischen Auswertungen erfolgende und auf Dauer angelegte Beschaffung und Aufbereitung von Informationen über das Bildungssystem und dessen Umfeld verstanden mit dem Ziel, Ersteres unter Qualitätsgesichtspunkten deutlicher sichtbar und damit besser kontrollier- und führbar zu machen. Die als Indikatoren von Bildungsqualität dienenden Informationen können dabei auf den unterschiedlichen Ebenen des Bildungssystems (Land, Region, Schule, Klasse, Schüler) erhoben werden. Das Bildungsmonitoring dient primär der Orientierung und Ergebnismeldung auf der Systemebene und gibt datengestützte Auskunft über die Leistungsfähigkeit des Bildungssystems. Sodann erlaubt es Trendanalysen in Bezug auf die längerfristige leistungsbezogene Entwicklung des Bildungssystems. Bildungs- oder System-Monitoring umfasst den Bereich der Bildungsberichterstattung ebenso wie die Überprüfung von Schulleistungen in nationalen und internationalen Vergleichsstudien. Nach Scheerens/Glas/Thomas (2003) speist sich ein ausgereiftes Bildungsmonitoring aus drei Quellen:

- bildungspolitischen Daten, welche über ein leistungsfähiges System erfasst und ausgewertet werden können,

Steuerung einher: der Vorgabe von Leistungszielen/-standards durch die zentralen Instanzen und ihrer Überprüfung in externen und unabhängigen Qualitätskontrollen sowie der Verpflichtung der Schulen zur Rechenschaftslegung“ (S. 262).

- Daten von Schulinspektionen, Selbstevaluationen von Schulen und zentralen Abschlussprüfungen und
- Testdaten aus internationalen, nationalen und regionalen Erhebungen.

Alle diese Daten haben evaluatorischen Charakter und können als diagnostische Grundlage für Bildungsplanung und bildungspolitische Entscheide sowie für die Rechenschaftslegung und die öffentliche Diskussion dienen.¹⁶⁷

Specht/Freudenthaler (2004) umreißen die diagnostische Funktion von Bildungsstandards auf der Systemebene der Bildungssteuerung wie folgt:

„Die Rückmeldungen an die Steuerungsebene (Bildungspolitik und Verwaltung) haben in erster Linie die Funktion sichtbar zu machen, wo – bezogen auf das Schulsystem als Ganzes – Stärken und Schwächen liegen (also Schülerkompetenzen befriedigend oder zu schwach ausgebildet sind); in welchen inhaltlichen Bereichen oder organisatorischen Sektoren besonderer Handlungsbedarf sichtbar wird und Entwicklungsschwerpunkte zu setzen sind; in welchen Bereichen sich über die Zeit Veränderungen in wünschenswerter oder unerwünschter Richtung zeigen. Hier geht es insbesondere darum, internationale Vergleiche herzustellen sowie sektorale und regionale Unterschiede und Entwicklungen über die Zeit abzubilden. Dabei muss vermittelt werden, dass dort, wo bestimmte Stärken oder Schwächen systemweit sichtbar sind, die Logik der Intervention auf der Ebene von Politik und Verwaltung liegt und nicht einfach den Einzelschulen übertragen werden kann“ (S. 624).

Während Bildungsmonitorings als periodisch eingesetzte Instrumente der Orientierung und der Selbstvergewisserung in Bezug auf die eigene Bildungsqualität in vielen Industrieländern, darunter den USA, eine mehr oder minder lange Tradition haben, sind sie für die deutschsprachigen Länder neu. Zu einem ausgebauten Bildungsmonitoring würden „Erhebungen in drei bis vier Jahrgangsstufen gehören (z.B. Jahrgangsstufen 3, (6), 9 und 12). Zur Information über den ‚output‘ des Bildungssystems reicht es völlig aus, alle drei bis vier Jahre eine repräsen-

¹⁶⁷ Vgl. für die Schweiz: www.skbf-csre.ch/bimo/bimo_de.html;
www.edk.ch/d/EDK/Geschaeft/Monitoring (Stand 12.09.06).

tative Erhebung mit jeweils neuen Stichproben durchzuführen“ (Klieme/Avenarius/Blum et al. 2003, S. 126). In der initialen Phase der Implementation von Bildungsstandards stellt ein nationales Bildungsmonitoring ein wichtiges Werkzeug dar, mit welchem geprüft werden kann, ob und in welchen Bandbreiten die gesetzten Standards von den Schülerinnen und Schülern auch wirklich erreicht werden.

Die Klieme-Expertise benennt eine Reihe von Fragen, die sich auf die Freiheitsgrade in der strukturellen Ausgestaltung eines Bildungsmonitorings beziehen (Klieme/Avenarius/Blum et al. 2003, S. 102) und welche bei einer Einführung des Instrumentes geklärt werden müssen:

- Werden nur Stichproben untersucht oder werden alle Schulen in das Monitoring-Programm einbezogen?
- Ist die Teilnahme für Schulen bzw. für einzelne Schülerinnen und Schüler und Lehrpersonen freiwillig oder verpflichtend?
- In welchem Rhythmus werden Erhebungen durchgeführt und auf welche Altersstufen und Fächer sind sie bezogen?
- Wie wird mit den Ergebnissen umgegangen? Werden sie veröffentlicht? Erhält die Schulaufsicht Einsicht in die Daten oder werden sie nur der Schule bzw. den Lehrkräften als Rückmeldungen zur Verfügung gestellt?
- Werden die Schulergebnisse als Rohwerte mitgeteilt oder an die Eingangs- und Rahmenbedingungen der Schulen „adjustiert“?
- Welche Art von Tests wird eingesetzt? Sind es reine Wissensprüfungen oder solche, die auch tiefes Verstehen und Problemlösen, ev. sogar Gruppenarbeiten und handlungsorientierte Aufgaben einschließen?
- In welcher Arbeits- und Verantwortungsteilung zwischen den Schulen selbst, den Behörden und dem wissenschaftlichen Personal werden die Tests geplant, durchgeführt und ausgewertet?¹⁶⁸

In Deutschland hat im Juni 2006 bezüglich einer Reihe dieser offenen Fragen insofern eine Klärung stattgefunden, als die Kultusministerkon-

¹⁶⁸ „Für die Forschung ist es wichtig, dass alle Daten der wissenschaftlichen Überprüfung zugänglich sind und bleiben und dass transparent gehalten wird, was wie und mit welcher Genauigkeit gemessen wird“ (Ramseier 2007).

ferenz einer Gesamtstrategie zum Bildungsmonitoring (KMK 2006b) zugestimmt hat, in der die folgenden Instrumente vorgesehen sind:

- Teilnahme an den nächsten Erhebungsrunden folgender internationaler Schulleistungsuntersuchungen: PISA, PIRLS/IGLU und TIMSS.
- Zentrale Überprüfung des Erreichens der KMK-Bildungsstandards in einem Ländervergleich: In mehrjährigem Rhythmus sollen repräsentative Stichproben von Schülern jeweils ein Jahr vor Abschluss des entsprechenden Bildungsgangs mittels Leistungstests geprüft werden: 3. Klasse (Primarbereich), 8. Klasse (Hauptschulabschluss) und 9. Klasse (Mittlerer Schulabschluss). Basis hierfür sind strukturierte Aufgabenpools, die vom Institut zur Qualitätsentwicklung im Bildungswesen (IQB) bis 2010 für die verschiedenen Fächer entwickelt und normiert werden. Organisatorisch sollen die Überprüfungen mit wechselnden fachlichen Schwerpunkten an die Durchführung der internationalen Studien PIRLS/IGLU (Grundschulbereich) und PISA (Sekundarbereich) gekoppelt werden; eine inhaltliche Anbindung über sog. Ankeritems ist ebenfalls geplant.
- Kriteriumsorientierte Vergleichsarbeiten in Anbindung oder Ankopplung an die Bildungsstandards zur landesweiten Überprüfung der Leistungsfähigkeit einzelner Schulen: Es ist vorgesehen, länderspezifische und länderübergreifende Vergleichsarbeiten künftig an den KMK-Bildungsstandards auszurichten.
- Gemeinsame Bildungsberichterstattung von Bund und Ländern: Ein indikatorengestützter Bildungsbericht soll regelmäßig differenzierte Informationen über zentrale Merkmale von Bildungsprozessen – wenn möglich im internationalen Vergleich und aufgeschlüsselt nach Ländern – liefern. Darüber hinaus soll er Anregungen für vertiefende Analysen bieten.

Beim Aufbau eines Bildungsmonitorings geht es nicht in erster Linie um die Entwicklung einzelner Tests, sondern um eine durchdachte Vernetzung im Sinne eines ganzen Assessment-Systems, das an die Bildungsstandards nicht nur verbal anschließt. Zu diesem Zweck müssen

die Tests auch zeitlich vernetzt sein und langfristige Trends müssen mit ihnen gemessen werden können (Ramseier 2007).¹⁶⁹

Nationale Bildungsmonitorings sind für die Akteure auf allen Ebenen des Bildungssystems interessant, weil sie Informationen über das durchschnittlich erreichte Leistungsniveau im Schulsystem als Ganzem bzw. in seinen sektoralen und regionalen Gliedern liefern. So lassen sich Fragen beantworten nach der Erreichung oder Nicht-Erreichung von Mindeststandards durch verschiedene Schülergruppen oder nach der erwartungskonformen Verteilung der Schülerinnen und Schüler über die Kompetenzstufen. Bildungsstandards leisten hingegen keinen Beitrag zu einer aus der Sicht der Lehrpersonen wünschbaren Individual- und Förderdiagnostik. Ohne Zusatzaufwand bei der Stichprobenbildung und bei der Testgestaltung ist es nicht möglich, einzelnen Schulen und ihren Schülerinnen und Schülern individuelle Rückmeldungen zu geben bzw. ihnen zu sagen, „wo sie stehen“. Von den auf der Systemebene erhobenen Daten lässt sich nicht auf die Individualebene schließen. Auch Leistungsstudien wie PISA, TIMSS, IGLU und DESI sind nicht dafür gemacht, Leistungsdiagnosen für einzelne Schülerinnen und Schüler zu erstellen.¹⁷⁰ Um verlässliche individuelle Leistungsprofile bzw. Schulprofile (Stärken-Schwächen-Analysen) zu erhalten, bedarf es fokussierterer und auch wesentlich aufwändigerer Instrumente wie z.B. länderbezogener oder interkantonaler Vergleichs- und Orientierungsarbeiten, die noch zusätzlich an die Bedürfnisse und Profile individueller Schulen angepasst werden müssen. Da Bildungsmonitorings die Standardsicherung zudem in der Regel an den Endpunkten wichtiger Abschnitte der Schulkarriere (in der Schweiz: 2., 6. und 9. Klasse) zu überprüfen beabsichtigen, wären dies auch nicht die geeigneten Zeitpunkte für individual- und förderdiagnostisch angelegte Evaluationen.

¹⁶⁹ Wie Ramseier weiter ausführt, betrifft diese Vernetzung auch weitere Entwicklungen: So wird in der Bildungsstatistik (zumindest in der Schweiz) für Schülerinnen und Schüler ein persönlicher Identifikator eingeführt. Bezieht man diesen in die Leistungsmessung ein, gewinnt das Monitoring mit wenig Aufwand eine stark erhöhte Aussagekraft, da longitudinale Aussagen möglich werden. Dafür müssten jetzt die rechtlichen Grundlagen geschaffen werden (Ramseier 2007).

¹⁷⁰ Groß angelegte (inter)nationale Vergleichsstudien müssen, um wissenschaftlich haltbar zu sein, im Gegenteil alles zu eliminieren suchen, was von individuellen Eingangsvoraussetzungen abhängig ist.

Herausforderungen an Infrastruktur und Nachhaltigkeit eines Bildungsmonitorings

In der oben skizzierten deutschen Gesamtstrategie spielt das IQB als wissenschaftliche Einrichtung dadurch eine bedeutsame Rolle, dass es in den verschiedenen Bundesländern empirische Studien durchführt, welche den interessierten Bildungsadministrations Steuerungsinformationen in Form von Schulleistungsindikatoren zur Verfügung stellen. Dagegen fehlen in Österreich und in der Schweiz gleichrangige Entwicklungs- und Koordinationsträger. In der Schweiz wurde bisher in Bezug auf PISA eine dezentrale Organisationsstruktur angewendet. Die Erhebung, aber auch die Datenauswertung und Berichterstattung werden von einer politischen Steuergruppe strategisch geplant. Die Koordination der operativen Tätigkeit liegt beim Bundesamt für Statistik. Die operative Tätigkeit selbst wird von vier regionalen Zentren ausgeführt, die sich weitgehend auf ausländische Forschung verlassen müssen. Zwei Zentren befinden sich in der Deutschschweiz, je eines in der italienisch- und französischsprachigen Schweiz. Übergreifende und dauerhafte wissenschaftliche Strukturen, welche auch Forschungskompetenzen in sich vereinigen, fehlen bislang in der Schweiz. Eine Herausforderung für ein gesamtschweizerisches Bildungsmonitoring liegt in der Schweiz zudem darin, dass Bildungsstandards und vor allem entsprechende Leistungstests in drei Sprachen übersetzt werden müssen. Nur so kann sichergestellt werden, dass zuverlässig nationale und sprachregionale Aussagen über das Einhalten der Bildungsstandards gemacht werden können. Allerdings wird man in der Schweiz nicht ohne sprachregionale Tests auskommen. Die Entwicklungen zur Qualitätssicherung mittels Tests in den drei Sprachregionen sind – abgesehen von der Beteiligung an PISA – bis heute unkoordiniert. Die Schwierigkeit wird in Zukunft darin bestehen, Tests zu entwickeln, die einen empirisch nachgewiesenen Bezug zu den Kompetenzmodellen und somit zu den Bildungsstandards haben. Was es für diese Aufgabe in der Schweiz braucht, ist ein wissenschaftliches Kompetenzzentrum mit Hochschulanbindung und zugeordnetem Konsortium, das im Rahmen der Bildungsstandards für Kompetenzmodellierung, Testentwicklung, Durchführung und Auswertung sowie Rückmeldung und Berichterstattung zuständig ist.

Eine zentrale Anforderung an eine nachhaltige und positive Wirkung eines nationalen Monitorings ist sodann seine längerfristige Nachhaltigkeit. Durch die Funktion des IQB dürfte eine Langzeitperspektive in Deutschland weitgehend sichergestellt sein, nicht jedoch in den anderen

Ländern, wo eine solche Perspektive noch fehlt. Bei einem nationalen Bildungsmonitoring kann es nicht darum gehen, ab und an isolierte, vergleichende Tests durchzuführen, besteht doch ein zentrales Monitoring-Bedürfnis gerade darin, zeitliche Trends in der Erreichung von Bildungsstandards zu ermitteln. Das heißt, dass zweckdienliche Tests einerseits der Komplexität der in den Bildungsstandards verlangten Kompetenzen gerecht werden und diese andererseits in eine langfristige Perspektive im Sinne von Trendanalysen bzw. Vergleichen im zeitlichen Verlauf stellen müssen. Eine entscheidende Vorbedingung dazu ist eine Testaufgaben-Datenbank, die aufgebaut, überprüft, laufend erweitert und gepflegt wird und auf deren Grundlage jeweils die jährlichen Tests zusammengestellt werden. PISA leistet dies bereits in einem rudimentären Rahmen. Bei einer stark ausgebauten Aufgaben-Datenbank wäre zudem denkbar und zu klären, ob nicht aus der Vielzahl von Aufgaben verschiedene Tests und lerndiagnostische Online-Hilfsmittel zusammengestellt werden können, die unterschiedlichen Bedürfnissen genügen. Eine Grenze würde allerdings dadurch gesetzt, dass solchermaßen verwendete Aufgaben, um ihre Funktion nicht zu verlieren, vertraulich bleiben müssten.

5.1.3 Rückmeldung von Ergebnissen

Ein sehr wichtiges Werkzeug einer Implementation von Bildungsstandards sind Rückmeldungen an Schulen und Lehrpersonen. Nach Darling-Hammond (1994) bemisst sich die Güte von Testsystemen und Evaluationen weniger an technischen Aspekten, als daran, wie deren Ergebnisse sich nach ihrer Rückmeldung für die Weiterentwicklung eines Schulsystems nutzen lassen – als summative Ergebnisinformationen für Politik und Öffentlichkeit und/oder als kommentierte, diagnostisch gehaltvolle Feedbacks für Lehrpersonen, Schülerinnen und Schüler und Eltern. Bezüglich einer Rückmeldung von Evaluationsergebnissen ist zunächst zu unterscheiden zwischen Resultaten aus einem Bildungsmonitoring, länderspezifischen oder kantonalen Leistungstests und Schulevaluationen. Je nachdem, ob das Gesamtsystem evaluiert wird, eine Schule oder die Leistungen einzelner Schülerinnen und Schüler, oder ob Ergebnisse aus einem Bildungsmonitoring auf einzelne Schulen (oder Klassen, nicht jedoch auf den Einzelschüler im Sinne einer Individualdiagnostik oder -bewertung) heruntergebrochen werden, sind die Daten verschieden aufzubereiten (Aggregation, Adjustierung) und rückzu-

melden. Ein kritischer Punkt ergibt sich dabei aus der Forderung nach Rechenschaft und Transparenz einerseits und den negativen Wirkungen und entsprechenden Ängsten, die mit der Möglichkeit von Schulrankings verbunden werden, andererseits. Hinter die inzwischen in politischer Hinsicht etablierten Transparenzerwartungen kann jedoch nicht mehr mit guten Argumenten zurückgegangen werden. Die Erwartungen müssen mit einer sinnvollen wissenschaftlichen Aufbereitung der Daten umgesetzt werden.

Damit Bildungsstandards produktive Wirkungen auf allen Ebenen des Bildungssystems entfalten, bedarf es *fairer und intelligenter, auf die Verantwortungen der jeweiligen Adressaten abgestimmter Rückmelde-systeme*. In einer der Akzeptanz förderlichen Rückmeldekultur ist darauf zu achten, „dass man nur für etwas folgenreich evaluiert werden kann, wofür man auch verantwortlich ist“ (Fend 2005, S. 25). So ergibt es für eine Lehrperson und ihre Klasse wenig Sinn, den Leistungsmittelwert der eigenen Schule in einem Fach zu kennen, wenn die Daten mehrerer Klassen und Lehrpersonen an der Schule in den Test eingeflossen sind. Sodann erscheint es bei Leistungsrückmeldungen aus schulübergreifenden, gegebenenfalls systemweiten Monitoring-Tests unabdingbar, den Leistungsstand einer Schule um die Effekte wichtiger Rahmenbedingungen und soziokultureller Schüler-Eingangsvoraussetzungen zu bereinigen,

„denn nur so kann die eigentliche Leistung der Schule, d.h. der von ihr erarbeitete Mehrwert gegenüber den Eingangs- und Rahmenbedingungen fair evaluiert werden. Die Verwendung solcher statistisch bereinigter (value-added) Kennwerte gehört heute – angesichts manch unerwünschter Vorkommnisse, die vor allem in Großbritannien nach der Veröffentlichung unkorrigierter Testergebnisse in ‚league tables‘ zu verzeichnen waren – zum Standardrepertoire der professionellen Schulevaluation“ (Klieme 2005a, S. 52).

Zu bestätigen ist somit die Forderung nach einer Vergleichbarkeit vor allem von Schulen mit ähnlichen Eingangsbedingungen, aber auch nach dem Vergleich im zeitlichen Verlauf, um auf diese Weise unterschiedliche Ausgangsbedingungen berücksichtigen zu können. Idealerweise sollten sich Rückmeldungen auf längsschnittlich erhobene Daten beziehen, weil nur so verlässliche Aussagen auch über die Entwicklung von Schulen und Schülerleistungen möglich sind. Zudem sollten sich Rückmeldungen auf mehr als einen einzigen Test stützen, der zufolge der Auswahl von oftmals nur wenigen Aufgaben nicht selten wenig facet-

tenreich ist. Das heißt auch, dass Schulen und Lehrpersonen nicht nur vor undifferenzierten, nicht kontextuierten Rückmeldungen, sondern auch vor überhöhten Ansprüchen an ihre diagnostischen Möglichkeiten zu schützen sind. Linn (2000, S. 15), Coffmann (1993) zitierend, ist jedenfalls zuzustimmen, wenn er zum Schluß kommt, „comparing averages among schools, systems, or states on any tests is inherently unfair because it is not possible to separate school effects from effects resulting from nonschool factors.“ Unter anderem schlägt er vor: „Don't put all of the weight on a single test. Instead seek multiple indicators. (...) Place more emphasis on comparisons of performance from year to year than from school to school. (...) Consider both value added and status in the system. (...)“

Für einen wirkungsvollen Umgang mit Leistungstests gilt es, den Stellenwert der Leistungsmessung richtig einzuschätzen und die Ergebnisse auch zu nutzen. Leistungsmessungen führen zu Transparenz bei ausgewählten Kompetenzen der Schülerinnen und Schüler. Die Ergebnisse können als Anlass zur Förderung sowie zur Optimierung des Unterrichts genutzt werden. Sie dienen zudem der Orientierung der Schülerinnen und Schüler, der Eltern, der die Tests durchführenden Schulen, der Behörden und der Öffentlichkeit.

Weil Testergebnisse – wie auch die Ergebnisse anderer Beurteilungsinstrumente – immer mit einem Messfehler behaftet sind, ist es von Vorteil, die Beurteilung der Schülerinnen und Schüler auf der Basis von verschiedenen Datenquellen vorzunehmen. Ein einzelnes Ergebnis darf nie abschließend interpretiert werden. Weisen allerdings mehrere Ergebnisse in die gleiche Richtung, dann wird eine zuverlässige Interpretation möglich; dies immer unter der Voraussetzung der Verwendung von um schulleistungsrelevante Input-Variablen statistisch korrigierten Testergebnissen.

In Bezug auf den Zusammenhang zwischen Testqualität und dem diagnostischen Potenzial der Ergebnisse gilt: Je besser ein Test bereits in seiner Anlage nicht nur auf eine schulübergreifende Vergleichspopulation, sondern auch auf einzelschulische Profile abgestimmt ist, d.h. je kontextsensibler er ist, desto höher ist sein diagnostischer Wert zu veranschlagen – und desto besser die Akzeptanz an der Basis. Merkmale einzelschulischer Profile können sich beispielsweise auf die Gewichtung und Sequenzierung der Stoffbehandlung, auf verwendete Lehrmittel, aber auch auf die Klassenzusammensetzung oder die Lektionenzahl beziehen.

Idealerweise erfüllt die detaillierte Rückmeldung von Ergebnissen von Tests und Schulleistungstudien an die Verantwortungsebenen und Adressaten unterhalb der Systemebene folgende *Funktionen*:

- An Bildungsstandards orientierte, kontextsensible Tests verhelfen *Schülerinnen und Schülern* zu *unterrichtsunabhängigen und dennoch lernzielorientierten* Kompetenzprofilen, aus denen Stärken und Schwächen hervorgehen. Diese Standortbestimmung dient sowohl Lehrpersonen wie Schülerinnen und Schülern zur Ausrichtung des Lehr-Lern-Prozesses.
- Durch den Vergleich aggregierter Individualergebnisse beziehungsweise von Klassenergebnissen mit den Ergebnissen anderer Klassen erhalten die Lehrpersonen eine *objektivierte Fremdbeurteilung des Leistungsstandes ihrer Klasse*. Wie schon oft nachgewiesen, fällt es Lehrpersonen schwer, die Leistungen ihrer Schülerinnen und Schüler realistisch einzuschätzen. Mit externem Feedback verbundene Standortbestimmungen können damit zur Objektivierung der schulischen Beurteilungspraxis beitragen und den Lehrpersonen zur Optimierung der Schülerwahrnehmung dienen.
- Sofern die Leistungstests in mehreren Klassen einer Schule durchgeführt werden (Parallelarbeiten), lassen sich die Ergebnisse auch zwischen den Klassen innerhalb einer Schule vergleichen. Diese Standortbestimmung kann einer Schule zu einer *verbesserten Wahrnehmung der Schulqualität* dienen; vor allem dann, wenn die Ergebnisse im Kollegium analysiert werden und der Erfolg der getroffenen Maßnahmen gemeinsam überprüft wird.
- Falls die Teilnahme an der Leistungsmessung nicht freiwillig erfolgt, sondern aufgrund der Zugehörigkeit zu einer wissenschaftlich gebildeten Stichprobe, dann können die Ergebnisse auch national oder international sinnvoll verglichen werden. Diese Standortbestimmung kann für Anpassungen des Bildungssystems genutzt werden.
- Detaillierte Rückmeldungen, welche die Kommunikation von Leistungsprofilen der einzelnen Schülerinnen und Schüler einschließen, können verbunden mit verständlichen Interpretationshilfen *Ausgangspunkte zur Optimierung des eigenen Unterrichts* darstellen.

Bislang gibt es nur wenig Forschung zum Umgang und zur konkreten Nutzung von Schulen und Lehrpersonen mit bzw. von Ergebnismeldungen. Die Befragungsdaten, die aus der Nutzung standardisierter Schulleistungstests zum Beispiel in der Schweiz (vgl. Kap. 5.3.2) und

aus deutschen Studien vorliegen, zeigen zwar eine recht hohe Akzeptanz von abrufbaren Leistungstests und eine demgegenüber eher gemischte Rezeption von extern verordneten Lernstandserhebungen, sind aber in der Frage des realen Nutzens nicht sehr aussagekräftig. Dennoch wird davon ausgegangen, dass differenzierte und verständliche Ergebnisrückmeldungen an Schulen und Lehrpersonen eine wichtige Rolle im Prozess der Etablierung einer stärker an Wirkungen orientierten Qualitätssicherung, welche auch die Unterrichtsebene erreicht, zukommt.

Wenn mit Specht/Freudenthaler (2004) im Folgenden von intelligent gestalteten Rückmeldungen gesprochen wird, dann stellt sich die Frage nach den Kriterien für eine optimale Nutzung der rückgemeldeten Informationen durch die Empfänger des Feedbacks. Das heißt, um Ergebnisse von Evaluationen und von Schulleistungstests für faire Vergleiche und als Input für die Schulentwicklung nutzen zu können, muss eine Reihe von Bedingungen erfüllt sein. Wir formulieren diese hier in Anlehnung an ein Gutachten von Schneewind (2006, S. 20ff.) und an Specht/Freudenthaler (2004) als *12 Gütekriterien von Rückmeldungen von Ergebnissen in standardisierten Leistungstests*:

- (1) Die Rückmeldung hat einen vertrauenswürdigen Absender.
- (2) Definition des Instruments und Zweck der Untersuchung werden klar kommuniziert.
- (3) Die Rückmeldung ist sowohl für Lehrpersonen als auch für Schülerinnen und Schüler einfach zu erhalten, z.B. online über einen Lernstandsserver (vgl. Lucyshyn 2006b; Peek/Dobbelstein 2006).
- (4) Die Rückmeldung erfolgt zeitnah, so dass sich die Erkenntnisse bei den einbezogenen Schülerinnen und Schülern umsetzen lassen.
- (5) Die Rückmeldung ist verständlich und in der Fachsprache der Zielgruppe abgefasst. Wichtige Interpretationen von Daten und Vergleichen werden mit den Daten mitgeliefert.
- (6) Die Rückmeldung erfolgt primär kriteriumsorientiert (und nicht sozial vergleichsorientiert) unter Berücksichtigung der empirisch gegebenen Leistungsbandbreite, auf der sich die Klassen/Schülerinnen und Schüler situieren können.
- (7) Die Rückmeldung benennt über die Kommunikation von Mittelwerten hinaus individuelle Kompetenzprofile bzw. Schulprofile (Stärken-Schwächenanalysen); sie knüpft an Stärken an, während

Schwächen im Rahmen von differenzierten Profilen abgebildet werden; sie vermeidet es, bedrohlich zu sein.

- (8) Die Rückmeldung muss fair sein, d.h. sie orientiert sich an Referenzwerten von Teilpopulationen, deren Eingangsvoraussetzungen und Rahmenbedingungen denen der Adressaten der Rückmeldung ähnlich sind; Bezüge zwischen Schülerleistungen und Kontextvariablen werden nach Möglichkeit hergestellt.
- (9) Die Beschreibung der Ergebnisse soll das Herstellen von Bezügen zu Formen und Merkmalen der Unterrichtsgestaltung ermöglichen; beeinflussbare, veränderbare Aspekte – und damit Orte der Verantwortung und Räume für positive Veränderungen – hebt die Rückmeldung hervor.
- (10) Die Rückmeldung liefert für Ergebnisse, die unterhalb definierter Schwellenwerte (Niveaus von Mindeststandards) liegen, Anregungen zu Förder- und Unterstützungsmassnahmen im Sinne von Entwicklungshinweisen und sie macht Vorschläge, in welchen Kompetenzbereichen vermehrte Eigenanstrengungen primär ansetzen sollten.¹⁷¹
- (11) Im Falle ungünstiger Rückmeldung benennt eine gute Rückmeldung einen Ansprechpartner für Moderation und Unterstützung, um Demotivation und Abwehrreaktionen zu vermeiden sowie Anregungen zur Verarbeitung der Rückmeldung und zur Maßnahmenplanung zu geben.
- (12) Die Rückmeldung richtet sich in konstruktiver Weise insbesondere auch an die Schulleitung, weil eine positive Einstellung der Schulleitung gegenüber externen Rückmeldungen maßgeblich zur Akzeptanz beim Lehrerkollegium beiträgt.

Kontextsensible, informationsreiche und adressatenfreundliche Rückmeldungen sind das eine, daraus angemessene Folgerungen für das Handeln bzw. die Weiterentwicklung des eigenen Unterrichts zu ziehen, das andere. Bei der pädagogischen Nutzung von Leistungsrückmeldungen handelt es sich, wie mittlerweile vielfach belegt werden kann, weder um einen automatisch ablaufenden noch um einen linearen und einstu-

¹⁷¹ „Sanktionen sind bestenfalls dort angebracht, wo einzelne Schulen oder Lehrpersonen trotz entsprechender Rückmeldungen über längere Zeiträume unterhalb der Erwartungen bleiben und Bemühungen um Verbesserung nicht glaubhaft dokumentieren können“ (Specht/Freudenthaler 2004, S. 626).

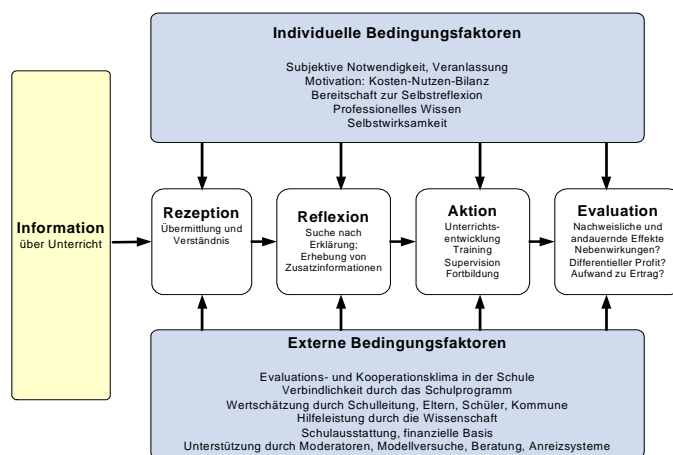
figen Prozess (vgl. z.B. Helmke 2004a; Kuper/Schneewind 2006; Schneewind 2006). Nach Schneewind (2006) ist von einem mehrschrittigen Ablauf auszugehen:

- Wahrnehmung einer Rückmeldung;
- Analyse der Information;
- Entwicklung von Maßnahmen;
- Umsetzung der Maßnahmen;
- Überprüfung der Wirkung der Maßnahmen.

Nach Helmke/Hosenfeld (2005; zitiert aus Schneewind 2006) handelt es sich um einen von der Evaluation zur Innovation verlaufenden zyklischen Prozess (vgl. Abb. 8), der von der Rückmeldungsinformation ausgeht und über deren Rezeption und Analyse zur Entwicklung und Umsetzung von Maßnahmen führt, deren Wirkungen wiederum evaluiert werden. Dass es sich dabei um keinen einfachen Prozess handelt, lässt sich bei Helmke (2004a) wie folgt nachlesen:

„Von der Evaluation zur Innovation verläuft kein einfacher, direkter und linearer Weg. Soll das Innovationspotenzial der schulischen Leistungsmessung im Kontext von Vergleichs-, Orientierungs- und Parallelarbeiten sowie Lernstandserhebungen ausgeschöpft werden, dann bedarf dies eines breiten Bündels sich wechselseitig stützender Maßnahmen über einen längeren Zeitraum. Voraussetzung ist eine sorgfältige Diagnose förderlicher oder hinderlicher individueller (Lehrpersonen) und kontextueller Bedingungen“ (S. 92).

Abbildung 8: Von der Evaluation zur Innovation: Zyklusmodell



Quelle: Helmke/Hosenfeld 2004, S. 100

Empirische Befunde zum Umgang mit Schulrückmeldungen liegen aus Deutschland u.a. mit der WALZER-Studie (Schrader/Helmke 2003) und mit der VERA-Studie (Vergleichsarbeiten in der Grundschule; Groß Ophoff/Koch/Hosenfeld/Helmke 2006) vor. Die Rückmeldungen von VERA beziehen sich sowohl auf Informationen auf der Ebene des Landes und der Schule als auch auf den Referenzrahmen der Klasse und auf die Ebene individueller Lernender und sogar Aufgaben. Insgesamt zielen sie auf eine Anregung von Entwicklungsprozessen an den beteiligten Schulen. Eine begleitende Rezeptionsstudie macht deutlich, dass es nicht einfach ist, zu vertieften Ergebnisinterpretationen anzuregen, die auch Rückhalt im Kollegium einer Schule finden. Insgesamt lässt sich nach Sichtung vorliegender Erfahrungen sagen, dass Schulleitungen den Rückmeldungen in der Regel positiv gegenüberstehen, während die Reaktionen bei Lehrpersonen eher gemischt sind. Einig sind sich Schulleitungen und Lehrkräfte in der grundsätzlichen Erwartung, dass Rückmeldungen erfolgen sollten, und ebenso einig sind sie sich in der Ablehnung der Veröffentlichung von Leistungsergebnissen und Schulrankings („naming and blaming“). Rankings – im Sinne eindimensionaler Rangordnungen – bilden nur einen eingeschränkten Ausschnitt des schulischen Geschehens ab und sind daher bloß minimal informativ. Als Momentaufnahmen sind sie nicht geeignet, Aussagen darüber

zu machen, weshalb Schulen an einer bestimmten Position stehen. Im ungünstigsten Falle wird Ungleichheit zwischen Schulen vergrößert, „wenn schlechte Schulen stigmatisiert werden und sich Eltern von ihnen abwenden“ (Rolff 2007, S. 49). Was sich sagen lässt ist, dass Rückmeldungen durchaus dazu anregen, den eigenen Unterricht zu überdenken (Funktion der Reflexion). Es scheint aber ein weiter und kurvenreicher Weg zu sein, bis daraus auch nach außen hin sichtbare Handlungsfolgen erwachsen und unterrichtsbezogene Maßnahmen ergriffen werden. Das heißt auch, dass die Akzeptanz differenzierter Rückmeldungen zwar groß, deren Handlungsnutzen jedoch begrenzt ist (Ditton/Merz/Edelhäusser 2002; Schneewind 2006). Ein weiteres Problem ist ferner, dass von Rückmeldungen vor allem jene Aussagen aufgenommen werden, welche die eigenen Erwartungen bestätigen. Neue Einsichten entwickeln sich in der Regel nur zögerlich, auch wenn die Informationsgrundlage durch die Rückmeldung dazu gegeben wäre. Wenn eine Maßnahme von außen vorgeschlagen wird, heißt es nicht selten: „Genau das machen wir doch schon seit langem.“ Zudem bleibt zu beachten, dass Lehrpersonen mit der Interpretation wissenschaftlicher Daten Neuland betreten und dass sie möglicherweise andere Informationsansprüche haben als die übrigen an der Evaluation beteiligten Akteure. So zeigten Auswertungen der Evaluationsstudien VERA, WÄLZER und MARKUS, dass Lehrpersonen Interpretationsschwierigkeiten im Umgang mit adjustierten Leistungsdaten haben, welche die Herkunft der Schülerinnen und Schüler berücksichtigen (Hosenfeld/Schrader/Helmke 2006). Auch der zeitliche Abstand zwischen der erfolgten Leistungsmessung und der Rückmeldung bedarf der Ausbalancierung verschiedener Interessen: Um einen möglichst hohen Aktualitätswert zu erhalten, soll eine Rückmeldung rasch erfolgen können. Dies kann jedoch zulasten einer maximal hohen Qualität und Differenziertheit der Feedbacks gehen.

Nach den vorliegenden Studien kann es keinen Zweifel darüber geben, dass kontextsensitive (adjustierte) Rückmeldungen aus Tests und Evaluationen eine notwendige Bedingung und ein wichtiges Instrument für eine pädagogische Qualitätsentwicklung, die auch die Unterrichtsebene erreicht, darstellen. Dennoch genügen Rückmeldungen – genügt noch so intelligent aufbereitete Outputinformation – nicht. Diese informiert wohl darüber, *dass* etwas – und allenfalls *was* – an bestimmten Ergebnisqualitäten nicht stimmt. Durch welche Prozesse die Ergebnisse zustande gekommen sind, darüber liefern Outputdaten keine Informa-

tion. Insbesondere zur Ermittlung sinnvoller Verbesserungsvorschläge bedarf es Informationen auch über die Input- und Prozessseite des Unterrichts. Was die Interpretation von Outputinformation erschwert ist, dass Unterricht als Prozess eine andere Handlungslogik hat als die wissenschaftliche Analyse von leistungsbezogenen Ergebnissen, eine Logik zudem, die nicht ohne Weiteres verfügbar ist (Messner 2006). Vor allem dort, wo die Resultate problematisch sind, muss die Rückmeldung durch Interpretationshilfen und Supportangebote ergänzt werden, die den Schulen helfen, die ungünstigen Ergebnisse ohne Demotivationsfolgen, und möglichst unter Durchbrechung selbstreferentieller Abwehrstrategien, in institutionelle und personenbezogene Lernprozesse zu übersetzen. In Österreich wird diese Aufgabe „offensiv“ angegangen, indem im Sinne eines „Einstiegs in eine Testkultur“ ein Pool von Lehrpersonen für das Rückmeldeverfahren an den Schulen ausgebildet wird (Lucyshyn 2006a, 2006b).

Schulrückmeldungen als Kernelement der Rechenschaftsfunktion von Testprogrammen können Anlass sein und einen Beitrag leisten für eine dateninduzierte Schulentwicklung. Nicht erfüllen können sie hingegen den Anspruch, eine mehrperspektivische und multikriteriale Schulevaluation zu ersetzen oder tiefergreifende Prozesse der professionellen Entwicklung quasi automatisch in Gang zu setzen und in produktive Bahnen zu lenken (Watermann et al. 2003; Rolff 2002). Dazu bedarf es über „intelligente Rückmeldungen“ hinaus weiterer Werkzeuge der Fortbildung und der Schul- und Unterrichtsentwicklung.

5.1.4 Rahmenlehrpläne, Kerncurricula, Lehrmittel

Lehrpläne werden nicht, wie in manchen Publikationen angenommen, durch Bildungsstandards überflüssig. Wir unterstreichen diesen Punkt. Lehrpläne ordnen das didaktische Feld und definieren den Zusammenhang der verschiedenen Fächer oder Lernbereiche im schulischen Curriculum. Lehrpersonen und Schulleitungen benötigen sie zur thematischen Orientierung sowie zur systematischen und zeitlichen Gliederung des Unterrichts. Zudem geben sie den staatlichen Rahmen vor und sind ein wichtiges Mittel der Legitimation des schulischen Angebots. Wenn Lehrpläne nur begrenzt den Unterricht steuern, dann ist das kein grundsätzlicher Einwand gegen sie (Criblez 2006).

Für die Unverzichtbarkeit von Lehrplänen gibt es ausreichend historische Evidenz (Goodson/Hopmann/Riquarts 1999). Ein Schulfach wird zu einem „Fach“, indem und soweit es in einem Lehrplan Berück-

sichtigung erfährt. Das gilt für die Benennung und die Funktion des Faches, für seine sachlichen Essentials, für die innere Gliederung und das Verhältnis zu den anderen Fächern, für die mit dem Rang eines Faches verbundenen Ressourcen, kurz: für seine Stelle in der *Topik* des Lehrplans (Künzli 1986). Wenn heute von „Kerncurricula“ gesprochen wird, dann ist das eine Reaktion auf additive Lehrpläne ohne Prioritäten, die allein schon aufgrund ihres Umfangs nicht mehr handhabbar waren. Der ordnenden Funktion von Lehrplänen tut auch dies keinen Abbruch.

Lehrpläne sagen über die tatsächliche Nutzung im Schulalltag nichts aus. Wer Lehrpläne so betrachtet, kann sie eigentlich nur entwerten, weil ihre Funktion und ihre Reichweite dadurch überdehnt werden. Die heute üblichen Lehrpläne ersetzen weder Lehrmittel noch geben sie die Inhalte des Unterrichts vor, der dann tatsächlich erteilt wird. Vielmehr zeigen sie einen Rahmen auf, der im Einzelfall sehr dehnbar gehalten werden muss, wenn Unterricht zustande kommen soll. Unterricht ist Anpassung an die gegebene Situation, nicht an den Lehrplan oder nur durch diesen vermittelt. Zwischen Lehrplan und Unterricht gibt es so immer eine Distanz, die mehr oder weniger groß sein kann.

Der Expertise von Klieme/Avenarius/Blum et al. (2003, S. 95ff.) ist aber darin zuzustimmen, als dass eine ausgewogene Mischung von Bildungsstandards, Kerncurricula und Kanonorientierung auf der Input-Seite nach wie vor nötig ist. Andererseits sind Bildungsstandards nicht einfach Ableitungen aus Curricula und Lehrplänen. Standards bestimmen „konkrete Kompetenzanforderungen“ und beinhalten somit „einen Maßstab zur Bewertung von Lernergebnissen“ (ebd., S. 99). Standards decken nicht den gesamten Bereich eines Fachs ab. Lehrpläne gehen darüber hinaus. Sie beschreiben und erfassen den Gesamtzusammenhang eines Curriculums.

Lehrpläne und vor allem Lehrmittel haben somit auch unabhängig von Kompetenzmodellen für Aufgaben und Tests eine zentrale Bedeutung für das, was die Schülerinnen und Schüler in der Schule lernen und leisten sollen. Wenn aber das Ziel einer Steuerung des Bildungssystems über den Output und die Rückübersetzung der Ergebnisse von Leistungstests in den Unterricht erreicht werden soll, dann müssen die Leitfunktion von Bildungsstandards und die Orientierungsfunktion von Lehrplänen systematisch verkoppelt werden. Andernfalls droht die Gefahr, dass die Arbeitsfelder auseinander fallen, in diejenigen von Lehr-

plankommissionen und in diejenigen von Testpsychologien und Fachdidaktikern.

In der Schweiz gibt es eine große Anzahl kantonaler Lehrpläne, die wohl in gegenseitiger Annäherung entwickelt wurden, aber keinem einheitlichen Schema unterliegen und immer noch große Unterschiede markieren. Das gilt auch für Kernfächer wie Mathematik oder die Landessprachen. Bildungsstandards konkretisieren die Bildungsziele in Form von Kompetenzmodellen, die dem Schweizer Projekt *HarmoS* und so der Harmonisierung des curricularen Angebots zugrunde liegen werden. Wir gehen darauf in 6. Kapitel näher ein. An dieser Stelle genügt der Hinweis, dass sich der Prozess der Harmonisierung der obligatorischen Schule auch auf die Lehrpläne beziehen muss, wenn nicht gegenläufige Entwicklungen in Kauf genommen werden sollen. Geplant ist daher, bis zum Jahr 2011 einen einheitlichen Deutschschweizer Lehrplan für die Volksschule zu entwickeln.

Die Trennung der Entwicklung von Standards auf der einen Seite und Leistungstests auf der anderen bringt die Gefahr mit sich, dass die im Unterricht verfolgten Standards nicht mit den durch Leistungstests überprüften Standards übereinstimmen. Eine mangelnde Koordination zwischen Lehrplänen und Standards läuft ihrerseits Gefahr, dass sich Lehrpersonen nur an Standards oder den damit verbundenen Testaufgaben oder nur an Lehrplänen orientieren. Wenn mit Tests selektive Funktionen verbunden sind, werden Lehrpläne auch in ihrer Ordnungsleistung marginal. Um das zu verhindern, sollte eine möglichst große Kongruenz zwischen Bildungszielen, Bildungsstandards, Unterrichtsentwicklung und Leistungstests bestehen.

Das erlaubt eine Schlussfolgerung: *Allein* durch die Entwicklung von Bildungsstandards, unabhängig von der Entwicklung von Lehrplänen, Tests, Unterricht sowie der Aus- und Weiterbildung der Lehrpersonen, werden die mit den Standards verbundenen Hoffnungen kaum erfüllt werden können. Die Ergänzung der Lehrpläne durch Standards ist vor allem dann wirksam, wenn sich die Standards und die damit verbundenen Kompetenzbeschreibungen in Leistungstests, aber auch im Unterricht, wiedererkennen lassen (ebd., S. 50). Eine isolierte Beurteilung von Bildungsstandards ergibt deshalb keinen Sinn. Zudem muss immer wieder darauf verwiesen werden, dass nicht in allen Fächern getestet wird und dass Tests nur zu bestimmten Zeitpunkten vorgesehen sind. Der weitaus größere Teil des Unterrichts wird daher wie bislang durch den Lehrplan geordnet sein.

Lehrmittel und Schulbücher sind für Lehrpersonen häufig weit wichtiger als Lehrpläne. Das lässt sich an einem Beispiel zeigen. Im Kanton Zürich wurde vor zehn Jahren ein neuer Lehrplan der Volksschule in Kraft gesetzt, vergleichbar den 2004 vorgelegten neuen Bildungsplänen in Baden-Württemberg. An der Erstellung des Zürcher Volksschullehrplans waren viele sehr engagierte Lehrkräfte beteiligt. Der Lehrplan stieß in der Vernehmlassung¹⁷² auf große Zustimmung und wurde am Ende einhellig begrüßt. Man sah sich nach einer aufwändigen Erprobung in der Lehrplanarbeit auf einem guten Weg und größere Probleme bestanden nicht.

Wenige Jahre später wurde eine kantonsweite Evaluation des Nutzungsverhaltens durchgeführt. Ihr zentrales Ergebnis war, dass auch dieser neue – an sich schlanke – Lehrplan in den Schulen nicht systematisch verwendet wurde und eher nur geringen Einfluss auf den Unterricht hatte (Landert/Stamm/Trachsler 1998). Routine und Repertoire der Lehrkräfte waren davon nicht wirklich berührt, und die Unterrichtspraxis wurde nicht auf eine neue Basis gestellt, wie bei der Einführung erwartet wurde. Lehrpläne wollen Ideale sein, die gleichwohl Bindungscharakter haben sollen (Fries 1998), ohne dadurch mehr zu sein als ein Rahmen, der den tatsächlichen Unterricht nicht wirklich vorschreiben kann. Die Lehrkräfte unterrichten nicht strikt „nach Lehrplan“, sondern benutzen Lehrpläne in verschiedenen Funktionen, darunter auch derjenigen der Legitimation ihres Unterrichts.

Die Lehrmittel müssen einen Bezug zu den Lehrplänen haben, der aber keineswegs immer gegeben ist. Die Lehrpläne ihrerseits dürfen einen bestimmten Umfang nicht überschreiten, während sie in den vergangenen Jahrzehnten ständig angewachsen sind und immer neue Komponenten aufgenommen haben. Das Leitmedium für den Unterricht sind die Lehrmittel. Für die Lehrkräfte sind Lehrpläne dann handhabbar, wenn sie knapp, übersichtlich und informativ die Struktur des Faches und den Zusammenhang des Curriculums wiedergeben. Die Lehrpläne wirken nicht direkt auf den Unterricht, sie werden nur in bestimmten Funktionen genutzt und sollten darauf eingestellt sein.

Das spricht dafür, Lehrpläne frei zu halten von detaillierten Stoffverteilungen und die Konkretisierung den staatlichen Lehrmitteln zu überlassen. Stoffverteilungen haben den Nachteil, dass sie kaum etwas

¹⁷² Befragung aller relevanten Gruppen vor Verabschiedung.

begründet ausschließen können und deshalb die Form der Addition von Beispielen wählen müssen, die sich beliebig verlängern lässt. Beispiele bilden keinen didaktischen Zusammenhang, wie dies in den Lehrmitteln der Fall sein muss. Lehrmittel organisieren ein zeitliches Nacheinander, was allein schon für Beschränkung sorgt. Zudem ist der Umfang aufgrund des Nutzungsverhaltens der Schülerinnen und Schüler limitiert.

Auch das – wie oben dargestellt – sehr durchstrukturierte National Curriculum in England ist in den Details knapp gehalten. Es wirkt nur deswegen so umfangreich, weil es für alle Stufen und alle Fächer konstruiert worden ist. Entscheidend ist aber keine Stoffverteilung, die schnell ins Uferlose führt, sondern entscheidend sind vielmehr der konsequente Aufbau der Fächer über die Key Stages hinweg, die acht Kompetenzstufen für jedes Fach und die Zielvorgaben. Es gibt daneben Schulbücher, Unterrichtsmaterialien und auch ein Online-Curriculum, das bis auf Lektionenstufe hinabreicht und einen nützlichen Wegweiser durch die unüberblickbare Literatur darstellt.¹⁷³

Ein Gegenbeispiel sind die neuen Lehrpläne für die allgemeinbildenden Schulen im deutschen Bundesland Baden-Württemberg. Die *Bildungspläne 2004* sind als „grundlegender Paradigmenwechsel“ verstanden worden. Offiziell heißt es: „Während frühere Bildungsplangenerationen vorrangig auswiesen, was zu unterrichten ist, schreiben die neuen Bildungspläne vor, welche Kompetenzen Kinder und Jugendliche erwerben müssen. Hiermit wird ein Wechsel von einer Input- zu einer Outputsteuerung vollzogen.“ Die Etappen des Kompetenzerwerbs sind durch die Ausweisung von Bildungsstandards gekennzeichnet.

„Diese beschreiben fachliche, personale, soziale und methodische Kompetenzen der Schülerinnen und Schüler. Diesen Kompetenzen sind in Form eines Kerncurriculums Inhalte zugeordnet, die so ausgewählt sind, dass sie in rund zwei Dritteln der verfügbaren Unterrichtszeit erarbeitet werden können. Sie sind Grundlage der zentralen Prüfungen. Ansonsten wird das Erreichen der Bildungsstandards in den weiterführenden Schulen unter anderem mit zentral gestellten Vergleichsarbeiten auf der Basis des Kerncurriculums überprüft. In der Grundschule geschieht dies über Diagnosearbeiten, die den Lern- und Entwicklungsstand feststellen und damit die Grundlage für weitere, gezielte Förderplanung darstellen.“

¹⁷³ <http://www.curriculumonline.gov.uk/> (Stand 15.09.06).

Neben dem *Kerncurriculum* steht ein *Schulcurriculum*, das in den Schulen selbst erarbeitet wird und das verbleibende Drittel der Unterrichtszeit umfasst.

„Das Schulcurriculum enthält die Curricula der einzelnen Fächer beziehungsweise Fachverbände, ergänzende und vertiefende sowie fächerverbindende und profilbildende Elemente. Darüber hinaus hat es die Funktion, einzelne Fächer und Fächerverbände miteinander im Sinne des fächerverbindenden Unterrichts zu vernetzen. Erst durch das Zusammenspiel von Kerncurriculum und Schulcurriculum werden die in den Bildungsstandards geforderten Kompetenzen erreicht.“

Was „Bildungsstandards“ genannt wird, ist in den Bildungsplänen beschrieben. Diese gliedern sich in drei Ebenen mit unterschiedlichem Verbindlichkeitsgrad:

- Die erste Ebene (Bildungsstandards) legt die staatlichen Vorgaben gemäß Schulgesetz für die Schulen fest, sie liegt als Druckfassung vor.
- Die zweite Ebene (Niveaunkretisierung) konkretisiert die erste anhand ausgewählter Beispiele. Sie ist verbindlich, was das dargestellte Niveau betrifft – nicht jedoch bezüglich der ausgewählten Beispiele.
- Die dritte Ebene (Umsetzungsbeispiele) ist nicht verbindlich. Auf ihr werden gelungene Beispiele zur Implementierung des Bildungsplanes 2004 vorgestellt.¹⁷⁴

Zugleich sind allgemeine und verbindliche „Maßstäbe“ formuliert worden, die in der gemeinsamen Einführung zu jedem Einzellehrplan erläutert werden. Auch diese Maßstäbe sollen sich überprüfen lassen. Dazu, heißt es in der Einführung, sei Konzentration erforderlich. Die Schülerinnen und Schüler sollen in Zukunft nicht mehr „enzyklopädisch“ ausgebildet werden, auch nicht einfach „über etwas“ Bescheid wissen, sondern sie sollen „Grundtatbestände“ lernen. Gemeint ist eine, wie es heißt, „Konfiguration von wenigen, aber grundlegenden Kompetenzen“. Definiert wird dieser Zusammenhang so:

¹⁷⁴ http://www.bildung-staerkt-menschen.de/schule_2004/bildungsplan (Stand 15.09.06).

„Eine Kompetenz ist eine komplexe Fähigkeit, die sich aus richtigem Wahrnehmen, Urteilen und Handeln können zusammensetzt und darum notwendig das Verstehen der wichtigsten Sachverhalte voraussetzt“ (Bildungsplan 2004, S. 8).

Eine Überprüfung der Kompetenzerreichung ist tatsächlich nur möglich, wenn Maßstäbe vorhanden sind. Aber die allgemeinen Maßstäbe kommen in den anschließenden Fachlehrplänen gar nicht vor. Alle einzelnen Lehrpläne formulieren „Leitgedanken zum Kompetenzerwerb“, aber kein einziger greift auf den zitierten Begriff der Kompetenz zurück. Es gibt so in materieller Hinsicht kein allgemeines, sondern nur ein fach- oder domänenspezifisches Konzept von „Kompetenz“. Die einzelnen Lehrpläne formulieren für jede einzelne Klassenstufe und Schulart nicht nur Leitgedanken, sondern bestimmen auch Kompetenzen und Inhalte, die wiederum nichts mit den allgemeinen „Maßstäben“ zu tun haben. Es sind Lehrpläne für ganz unterschiedliche Fächer oder Lernbereiche, die weder eine einheitliche Begrifflichkeit verwenden noch aus einer bestimmten Theorie heraus verfasst sind.

Die Auflistungen in den einzelnen Lehrplänen scheinen abschließend gemeint zu sein. Prioritäten werden jedoch nicht formuliert und die Bildungspläne verfügen auch nicht über eine Ausschlussregel. Der Detaillierungsgrad ist in den verschiedenen Lernbereichen und Schulfächern ganz unterschiedlich. Und nirgendwo steht, wie das Ganze den Unterricht erreichen soll. Es gibt offenbar keine entwickelte Implementationsstrategie. Man kann auf das Landesportal Baden-Württemberg gehen, um sich Umsetzungsbeispiele anzusehen,¹⁷⁵ die wiederum ganz verschiedene Themen berühren und bislang weitgehend unsortiert sind. Mit den allgemeinen Maßstäben hat auch das nichts zu tun. Generell ist zu sagen, dass unter dem Ausdruck „Bildungsstandards“ nicht verstanden wird, was international üblich ist, nämlich ein gestufter Erreichungsgrad. Kompetenzen sind wie früher Fertigkeiten und Fähigkeiten, die von allen Schülerinnen und Schülern unterschiedslos erwartet werden. Es ist unmöglich, dass alle Schülerinnen und Schüler gleiche Kompetenzen erreichen, doch nur das sieht der Lehrplan vor. Graduierte Kompetenzen sind für die Förderung unerlässlich und sie schaffen auch einen realistischen Erwartungshorizont.

175 <http://www.bildung-staerkt-menschen.de/unterstuetzung> (Stand 15.09.06).

Zur Illustration des Problems dienen zwei Beispiele, ohne dass wir damit eine Gesamtanalyse des Bildungsplans zu präsentieren beanspruchen. Der Unterricht in Französisch als erster Fremdsprache 6. Klasse Realschule soll erstens kommunikative Fähigkeiten vermitteln, soll zweitens dazu führen, dass die sprachlichen Mittel beherrscht werden – der Lehrplan spricht von phonologischer, lexikalischer und grammatischer Kompetenz –, drittens soll der Umgang mit Texten zu bestimmten Fertigkeiten führen, viertens soll kulturelle Kompetenz entstehen und fünftens soll auch noch Methodenkompetenz ausgebildet werden, zu der Lern- und Arbeitstechniken sowie Sprachlernkompetenz gehören. Der Kompetenzbegriff wird dadurch überdehnt, was genau erworben werden kann, wird dadurch nicht klarer. Es ist kein Zufall, dass der Detaillierungsgrad bei der „grammatischen Kompetenz“ hoch und bei der „interkulturellen Kompetenz“ nur gering ist. Das eine wird unterrichtet und das andere nicht (ebd., S. 85ff.).

Zweites Beispiel: Geschichte Realschule Klassen 6, 8 und 10. „Kompetenz“ bedeutet hier historisches Verstehen und zum Fach gehörende methodische Fertigkeiten. Interdisziplinäre Verknüpfungen werden erwähnt, aber nicht sehr stark gemacht. Der Geschichtsunterricht vermittelt Kenntnisse und Reflexionsformen des Faches. Die Themenbereiche sind auf grundlegende historische Probleme ausgerichtet und verfahren in sich chronologisch. Der Lehrplan sagt nichts darüber aus, ob „Grundtatbestände“ zum Verstehen der Welt, wie dies in den allgemeinen Maßstäben gefordert wird, angeeignet werden können oder nicht. Es ist einfach ein Lehrplan für das Fach Geschichte, wie er auch schon vorher verbreitet war. Gesteuert wird die Lehrplanentwicklung durch die Fachdidaktik und nicht durch allgemeine „Maßstäbe“, die auf das Fach gar nicht anwendbar sind. Eine solche Deduktion unterschätzt die Eigenlogik von Schulfächern (ebd., S. 104ff.).

„Bildungsstandards“ sind in Baden-Württemberg klar *curriculare* Standards, die über Aufgabenbeispiele konkretisiert werden. Zwischen fachlichen und überfachlichen Standards wird dem Sinn nach unterschieden, aber der Konkretisierungsgrad ist bei den überfachlichen Standards weit geringer als bei den fachlichen. Was erreicht werden soll, sind fachliche Kompetenzen, die in Lehrplänen festgelegt sind. Die Menge des Stoffes ist reduziert worden, dafür hat die Verbindlichkeit zugenommen. Kontrolliert wird die Zielerreichung mit zentralen Prüfungen sowie durch Vergleichs- und Diagnosearbeiten. Tests und externe Evaluationen, die sich auf Standards beziehen, sind offenbar kein

Thema. Auch über die einschlägigen Lehrmittel wird nichts gesagt. Und das hat vermutlich einen systematischen Grund.

Die Frage der Lehrmittel ist nämlich in den Auftragsländern sehr unterschiedlich gelöst. Während in der Schweiz staatliche Lehrmittelverlage Aufträge der Erziehungs- und Bildungsdirektionen ausführen, herrscht in Deutschland das Marktprinzip. Entscheidend ist hier, wer eine staatliche Zulassung erhält. Der Staat selbst produziert keine Lehrmittel und gibt auch keine in Auftrag. Aber auf diese Weise dürfte eine enge Bindung zwischen Lehrplänen, Schulbüchern und Unterrichtsmaterialien schwerlich zu erreichen sein. Solange nur curriculare Standards erwartet werden, mag das hingehen. Wenn aber Kompetenzmodelle die Zielerreichung wenigstens in Teilen bestimmen, müssen die einschlägigen Lehrmittel darauf eingestellt werden und das verlangt vermutlich staatliche Steuerung.

Es ist unklar, wie der Schulbuchmarkt auf die Entwicklung von Bildungsstandards und Leistungstests reagieren und inwiefern er helfen wird, die Qualität zu sichern. Der Markt ist bereits heute unübersichtlich und kennt, abgesehen von Auszeichnungen einzelner Lehrwerke und Preisverleihungen an Autoren, kaum Qualitätskontrollen. In verschiedenen Kantonen der Schweiz sind wohl die Lehrkräfte an der Beurteilung der Lehrmittel beteiligt, auch gibt es in aller Regel Erprobungsphasen, jedoch keine empirische Überprüfung der Wirksamkeit. Die Bildungspläne 2004 in Baden-Württemberg sind durch verschiedene Anhörungen gegangen, bevor sie veröffentlicht wurden, was für die Lehrmittel nicht gilt. Sie werden staatlich zugelassen oder nicht, was ohne Kontrolle nicht möglich wäre, aber es ist unklar, welche Rolle dabei in Zukunft die Bildungsstandards spielen werden.

Schulbücher und Unterrichtsmaterialien sind ein *Hauptobjekt* der Implementation von Bildungsstandards. Trotz der hohen Bedeutung für den Unterricht sind die Lehrmittel politisch kaum beachtet worden. Eine durchdachte Implementationsstrategie muss bei ihnen enden, wofür es in den vier Auftragsländern bislang kaum Vorarbeiten gibt. Ein positives Beispiel sind Fremdsprachenportfolios, die etwa für das Fach Englisch mit gestuften Kompetenzen konstruiert worden sind und inzwischen die Lehrmittel erreicht haben. Einen Service wie das englische Online-Curriculum gibt es allerdings nirgendwo. Bildungsportale wie in Baden-Württemberg sind auch in anderen Ländern vorhanden oder befinden sich im Aufbau, ohne dass davon aber die heutigen Lehrmittel betroffen wären.

Der direkte Internetzugriff der Lehrkräfte auf Musterlektionen oder Aufgabenbeispiele wird rasch an Bedeutung gewinnen und bereits heute konkurrieren viele Lehrmittel im Nutzungsverhalten der Schülerinnen und Schüler mit „Wikipedia“ und ähnlichen Angeboten. Auch das wird bei der Implementation von Bildungsstandards zu beachten sein. Lehrmittel müssen nicht mehr Buchform annehmen, Alternativen zu dem, was staatlich anerkannt ist, sind jederzeit möglich, und die tatsächlich erteilten Lektionen werden sich aus verschiedenen Quellen zusammensetzen. Gleichwohl werden zusammenhängende Lehrmittel, die einen bestimmten und folgerichtigen Aufbau von Kompetenzen gewährleisten, auch in Zukunft ihren Rang behalten. Umso wichtiger ist dann die Frage, wie sie auf Bildungsstandards eingestellt werden können.

Die Autoren von Lehrmitteln müssen für diese Art von Qualitätssicherung gewonnen werden, ohne ausschließlich Tests oder Kerncurricula vor Augen zu haben. Bei den Lehrmittelzulassungen ist darauf zu achten, dass keine curriculare Verengung vorgenommen wird, in der nur noch die durch Bildungsstandards abgedeckten Kernbereiche erscheinen. Lehrmittel stellen Prioritäten dar, nicht zuletzt aus der Sicht der Schülerinnen und Schüler bestimmen sie, was relevant ist und was nicht. Wenn die Lehrmittel sich auf Bildungsstandards beziehen sollen, dann muss sehr genau bestimmt werden, was sie ausmacht, wo sie zum Einsatz kommen und was sie ausschließen.

Grundsätzlich sollten Bildungsstandards wirklich *outputorientiert* sein. Sie sollten kurz gefasst sein und den Schulen alle Freiheiten überlassen, auf welchem Weg sie erreicht werden können. Das gilt auch in der Zuordnung von fachlichen Inhalten zu Kompetenzzielen. Die Ziele und Stufen sind vorgegeben und verbindlich; die Wahl passender Inhalte liegt bei den Lehrkräften oder es werden ihnen zumindest große Freiheiten gestattet, was auch für die Auswahl und Gestaltung der Methoden des Unterrichts gilt. Anders gäbe es keinen Fortschritt gegenüber den alten Plänen der Stoffverteilung und keinen Paradigmenwechsel hin zum Aufbau von fachlichen und überfachlichen Kompetenzen. Dann aber dürfen Bildungsstandards keine umgeschriebenen Lehrpläne sein.

Bildungsstandards müssen auf einen Kern hin ausgerichtet sein. Lehrpläne dagegen sollen nach wie vor Freiheiten der Stoffauswahl und -gestaltung ermöglichen. Beide, Bildungsstandards und Lehrpläne, dürfen nicht irgendwie deckungsgleich verstanden werden. Bildungsstandards sind kompetenzorientiert, die Lehrpläne sind nach wie vor inhalts- und stofforientiert, auch wenn sie zeigen sollten, wie Inhalte mit

Kompetenzen zusammenhängen. Zudem stellen sie eine Zeit- und Ressourcenverteilung dar, mit der der Rahmen des Unterrichts bestimmt wird. All das muss von Bildungsstandards unterschieden werden, die aber vermutlich Auswirkungen haben auf den Rang der Fächer im Lehrplan.

Grundsätzlich sind Lehrpläne *inputorientiert*. Aber damit muss mehr gemeint sein als eine Auflistung möglicher Inhalte und deren Illustration durch Beispiele. Bisherige Lehrpläne beschreiben nicht, was die Schülerinnen und Schüler am Ende eines Lehrgangs können müssen, sondern nur, womit sie sich beschäftigen sollen. Das aber ist zu wenig. Andererseits, wie die Beispiele aus Baden-Württemberg zeigen, verbessert sich die Situation nicht, wenn einfach nur Kompetenzen aneinander gereiht werden. Das englische National Curriculum zeigt, dass sich alle Fächer mit Kompetenzstufen abbilden lassen, die sich im Grad ihrer Erreichung unterscheiden. Der Input lässt sich so mit dem erwarteten Output verbinden.

Ein Problem ist auch, der Magie des Wortes „Standards“ zu unterliegen und unter dem Eindruck der heutigen Bildungspolitik *alles* so zu nennen, um möglichst viel vom Bisherigen bewahren zu können. Wer 102 Standards allein für den Hauptschulabschluss Deutsch verlangt, zeigt, dass die zentrale Botschaft nicht verstanden oder überhört worden ist. Standards sind nicht dasselbe wie die früheren Lehrplaninhalte. Standards formulieren einen Kompetenzkern und nicht alles. Das Beispiel zeigt auch, wie wenig die Realität der Überprüfung der Zielerreichung die Optik bestimmt hat. 102 Standards kann niemand überprüfen, und wenn das nicht die Absicht ist, dann muss der Begriff „Standards“ falsch verwendet worden sein. Der Kern eines Lernfelds kann nicht hundertfach bestimmt sein. Die Devise sollte lauten: Weniger ist mehr, aber nicht alles.

Man kann auch ganz einfach unterscheiden zwischen Themen und Objekten, die im Unterricht zu *behandeln* sind, und Zielen, die auf verschiedenen Niveaus *erreicht* werden können oder sollen. Aber dann müssten tatsächlich Mindeststandards formuliert werden, wie dies im Schweizer Projekt HarmoS angestrebt wird. Auch dieser Entscheid ist unter den Auftragsländern nicht einheitlich getroffen worden, so dass schon aus diesem Grunde der Ausdruck „Bildungsstandards“ materialiter sehr Verschiedenes bezeichnet. Oder anders gesagt: Die Outputsteuerung wird nicht einheitlich verwirklicht und der „Paradigmenwechsel“ findet unterschiedlich statt. Die Probleme aber stellen sich mindestens in bestimmten Hinsichten durchaus gleich.

5.1.5 ICT-Werkzeuge, Schul- und Bildungsserver

Informations- und Kommunikationstechnologien (ICT) bieten vielfältige Werkzeuge, um eine an Bildungsstandards orientierte Schul- und Unterrichtsreform zu unterstützen. Je nachdem, ob die Lehrpersonenbildung, die Unterrichtsentwicklung von Lehrpersonen oder das Lernen von Schülerinnen und Schülern im Zentrum steht, ergeben sich unterschiedliche Zielgruppen und Ansätze. Insbesondere das Internet bietet für Schulverwaltungen, Schulleitungen, Lehrpersonen, Eltern sowie Schülerinnen und Schüler je eigene Potenziale: vom Austausch von Materialien und Informationen zur Entwicklung der Bildungsstandards und ihrer Implementation über die internetbasierte Durchführung von (adaptiven) Fachtests (Wainer 2000) bis hin zur Rückmeldung von Leistungsergebnissen an Schülerinnen und Schüler und Lehrkräfte in netzbasierter Form.

Erfolgversprechende Modelle

Es kann davon ausgegangen werden, dass die Implementation von Bildungsstandards für Lehrpersonen und andere Akteure des Bildungswesens zu einem erhöhten Orientierungsbedarf führen wird, der mehrere Ebenen betrifft. Zu jeder dieser Ebenen existieren Modelle, die Anregungen für die Weiterentwicklung bestehender Angebote in den Auftragsländern bieten können.

Präsentation von Schulprofilen und Schulorganisation: Den Schulen werden größere Autonomiespielräume und wird gleichzeitig eine größere Verantwortung (*accountability*) für das Erreichen von Standards übertragen. Zur internen und externen Kommunikation und zur wirksamen Umsetzung ihres Profils benötigen Schulen Ideen und Instrumente. Vor allem aus den USA, aber auch aus europäischen Ländern, sind Ansätze bekannt, wie diese Anforderungen mit ICT unterstützt werden können:

- Praktisch in jedem Bundesstaat der USA hat sich der Gebrauch einer einheitlichen Schulverwaltungssoftware etabliert, mit Hilfe derer sich relevante Daten prozessorientiert erfassen und lokal wie zentral auswerten lassen (vgl. Farnsworth 2002). Diese integrierte Form der Datenerhebung schärft das Bewusstsein für Ausgangslagen und für die Verantwortlichkeit von Einzelschulen und deren Bildungsangebot.

- Von erfolgreichen Schulen werden „Good-practice“-Portraits erstellt. Schulen können ihren gelungenen Ansatz einerseits selbst im Internet vorstellen;¹⁷⁶ Good-practice-Modelle können aber auch redaktionell zusammengetragen werden, unterstützt durch Ausschreibungen, Wettbewerbe und Prämierungen.¹⁷⁷ Besonders anregend ist auch die Vermittlung von Modellen mittels Video (z.B. „Treibhäuser der Zukunft“: Kahl 2004). Solche Portraits können mit Möglichkeiten der Kontaktaufnahme versehen werden und zum Gegenstand von Diskussionen von Schulleitungen und Lehrpersonen werden.
- Aus wissenschaftlichen Befunden werden durch führende Institutionen praxisorientierte „What-works“-Leitlinien synthetisiert. Diese werden über zentrale Internetseiten zur Verfügung gestellt.¹⁷⁸ Bei ihrer Umsetzung ist auf die Bedeutsamkeit der Adaption in bestehende Schulkulturen zu achten.

Videoserver des Unterrichts: Um innovative Modelle anspruchsvollen Unterrichts anschaulich zu machen oder um über alltäglichen Unterricht und seine Probleme zu reflektieren, wird verstärkt auf das Medium Video zurückgegriffen (vgl. Brophy 2004; Reusser 2005b). Videos können mittlerweile problemlos über das Internet angeboten werden. Folgende zwei Beispiele illustrieren die Spannweite der Verwendung von Unterrichtsvideos:

- <http://www.teacher.tv> ist ein Internetangebot, das durch das britische Bildungsministerium (DfES) getragen wird. Hier finden sich über 700 Videobeispiele von „good practice“ zu den wichtigsten curricularen Themen der Primar- und Sekundarschule. Die Videos haben den Charakter von Unterrichtsreportagen, welche mit Kommentaren der Lehrpersonen sowie weiteren redaktionellen Beiträgen versehen sind.

¹⁷⁶ Z.B. <http://www.schulweb.de/de/deutschland/index.html>, <http://www.schulweb.de>; <http://enis.eun.org> (Stand 3.07.07).

¹⁷⁷ Z.B. <http://www.digionline.de/dyn/1818.htm> der Bertelsmann-Stiftung oder <http://schoolportraits.eun.org> des ERNIST-Projektes, die sich über eine Datenbank nach verschiedenen Themen gezielt abfragen lassen (Stand 03.07.07).

¹⁷⁸ Z.B. <http://www.whatworks.ed.gov> des US Department of Education oder <http://www.standards.dfes.gov.uk> des britischen Department for Education and Skills (Stand 3.07.07).

- Das Videoportal des Pädagogischen Instituts der Universität Zürich (<http://www.didac.uzh.ch/videoportal>) bietet demgegenüber eine Reihe von Unterrichtsvideos, die in einem wissenschaftlichen Kontext produziert wurden und die weitgehend ungeschnitten präsentiert werden. Diese Art von Videos ist vor allem dann sinnvoll, wenn sie in problemorientierte Aufgaben eingebettet werden und im Rahmen der Lehrpersonenaus- und -weiterbildung diskutiert werden.

Unterrichtsmaterialien: Über das Internet können Unterrichtsmaterialien, die Standards in besonderer Weise unterstützen, zur Verfügung gestellt werden. Elektronisch aufbereitete Unterrichtsmaterialien haben sich als Schlüsselfaktor einer verstärkten Nutzung bestehender Computernetzwerke in Bildungskontexten herausgestellt (z.B. Wiggenhorn/Vorndran 2002). Lehrpersonen suchen auf Bildungsservern vor allem nach unmittelbar einsetzbaren Unterrichtshilfen (Kos/Lehmann/Brenstein/ Holtsch 2005). Onlinematerialien lassen sich in Datenbanken nach sinnvollen Kriterien ordnen.

- Anregungen für einen derart strukturierten Distributionskanal kann die „lehrplanbasierte Navigation“ auf dem Zentralschweizer Bildungsserver geben, die sich angesichts drastisch steigender Nutzerzahlen seit ihrer Einführung als großer Erfolg herausgestellt hat.¹⁷⁹ In ähnlicher Weise wäre auch eine „standardbasierte Navigation“ denkbar. Der Erfolg des Portals ist wahrscheinlich nicht nur auf die intuitive Navigation zurückzuführen, sondern auch auf die strenge Qualitätskontrolle der auf diesem Bildungsserver abgelegten Lernmaterialien durch die kantonalen Fachberater.
- Das BLK-Programm SINUS-Transfer, in dessen Rahmen Lehrpersonen in wählbaren Modulen lernen, wie sie ihre Professionalität steigern und ihren Unterricht weiterentwickeln können, bietet über eine Datenbank u.a. Unterrichtsmaterialien und Aufgabensammlungen zu den naturwissenschaftlichen und mathematischen Fächern an.¹⁸⁰

¹⁷⁹ <http://www.zebis.ch/zebis11.php> (Stand 11.09.06).

¹⁸⁰ <http://sinus-transfer.uni-bayreuth.de/startseite.html> (Stand 28.08.06).

- Im Zuge von SINUS-Transfer werden in den beteiligten Bundesländern verschiedene Bildungsserver betrieben. So besteht etwa für Nordrhein-Westfalen eine Online-Schnittstelle zur Kommunikation über Inhalte und Umsetzung der in SINUS-Modulen entwickelten Unterrichtsmaterialien. Darin werden neben einem umfangreichen Materialpool (u.a. mit Manualen zu standardbasierten Kernlehrplänen) Beispielaufgaben mit Hinweisen zu Aufgabenstellung und möglichen Schwierigkeiten zur Einsicht und Weiterverwendung angeboten.¹⁸¹ Ähnlich Möglichkeiten existieren für Hamburg und Schleswig-Holstein. Die Ergebnisse des praxisnahen Modellversuchsprogramms sollen möglichst breit gestreut werden, damit alle Lehrpersonen beispielsweise Zugang zu erarbeiteten Aufgabenstellungen haben. In ähnlicher Weise könnten Internetplattformen für Materialpools und Netzwerke verwendet werden, wenn es um die Erstellung, Erprobung, Verwendung und Verbesserung von sinnvollen, standardbasierten Unterrichtsmaterialien geht.

Internetbasierte Testdurchführung: Das Internet ist grundsätzlich auch für Bildungsmonitoring und flächendeckende Tests (PISA als Beispiel) geeignet und sein diesbezüglicher Einsatz sollte erwogen werden. Um die Standards zu überprüfen, aber auch als Vorbereitung für die obligatorischen Tests, können Lehrpersonen formative Formen der Standortbestimmung zur Verfügung gestellt werden. Diese lassen sich entweder über das Internet verteilen oder effizient direkt im Rahmen von Lernplattformen gestalten und durchführen.

- Das Projekt VERA (Vergleichsarbeiten an der Grundschule), an dem mehrere deutsche Bundesländer beteiligt sind, bietet Vergleichsarbeiten in den Fächern Deutsch und Mathematik für die 4. Jahrgangsstufe der Grundschule. Die Lehrpersonen können die Testaufgaben vom Server der Projektleitung herunterladen.¹⁸² Neben Informationen zum zeitlichen Ablauf, dem Aufgabenpool und Durchführungshinweisen sind auch Korrekturanweisungen, didaktische Erläuterungen und Lehrplanbezüge für die jeweiligen Bundesländer sowie Beschreibungen der Fähigkeitsniveaus einsehbar.

¹⁸¹ <http://www.learn-line.nrw.de/angebote/sinus/grundschule/material/material.html> (Stand 27.08.06).

¹⁸² <http://www.uni-landau.de/vera/> (Stand 11.09.06).

Die Lehrpersonen korrigieren die Schülerarbeiten selbst und können die Ergebnisse über einen geschützten Bereich der Website eingeben.

- In der Schweiz wurden in verschiedenen Kantonen formative Instrumente zur Leistungsüberprüfung entwickelt, mittels derer Lehrpersonen Hinweise darauf erhalten können, wo die Leistung ihrer Schülerinnen und Schüler einzuordnen ist.¹⁸³ Solche Tests können im Idealfall über das Internet bezogen und ausgewertet werden.
- In Österreich wurde die Lernplattform „Moodle“ als Instrument genutzt, über das Lehrpersonen, Eltern und Schülerinnen und Schüler sich mit freigegebenen PISA-Aufgaben auseinander setzen können.¹⁸⁴

Netzbasierte Formen der Rückmeldung

Neben dem Aufbau eines System-Monitorings können auch Lernstandserhebungen und Vergleichsarbeiten einen Beitrag zur Schul- und Unterrichtsentwicklung leisten. Bei VERA wird beispielsweise der Verständlichkeit und Nützlichkeit der Ergebnissrückmeldungen an die Lehrpersonen großer Wert beigemessen. Die über das Web erfolgende Rückmeldung sieht verschiedene Bezugssysteme vor. So sollen neben einem an der Zentralstichprobe vorgenommenen, verteilungsbezogenen Vergleich auch ein kriterialer sowie ein Vergleich über den zeitlichen Verlauf ermöglicht werden. Die Ergebnissrückmeldung bildet die Grundlage für Unterrichtsentwicklungsschritte, welche durch die Lehrpersonen angeregt und unternommen werden sollen. Nach Angaben der Lehrpersonen wird der Nutzen der Ergebnissrückmeldungen vor allem als Ausgangspunkt für Förderaktivitäten im Unterricht gesehen und weniger als Basis zur Reflexion des eigenen Unterrichts. Die Eltern können ihrerseits über einen geschützten Passwortbereich Informationen über die Leistungen ihrer Kinder abrufen (Groß Ophoff/Koch/Hosenfeld/Helmke 2006). Dieses Projekt zeigt beispielhaft den Einbezug der Lehrpersonen in ein umfangreiches System-Monitoring. Gleichwohl ist für standardbasiertes Lehren und Lernen darauf zu achten, dass netzbasierte Ergebnissrückmeldungen nicht nur die Grundlage für prag-

¹⁸³ Z.B. <http://www.klassenscockpit.ch> (Stand 9.09.06).

¹⁸⁴ <http://www.classroom.at> (Stand 9.09.06).

matische Fördermaßnahmen bilden, sondern darüber hinaus Anlass bieten für individuelle oder kooperative Formen der Unterrichtsreflexion.

In Österreich wurden im Rahmen der Einführung von Bildungsstandards ebenfalls Anstrengungen unternommen, webbasierte Rückmeldeverfahren zu entwickeln. Über eine Internetplattform mit dem Namen „Gemeinsam lernen“¹⁸⁵ werden u.a. Informationen zu Kompetenzen und Standards, zu deren Bezügen zum Lehrplan und zur Unterrichtsentwicklung angeboten. „Gemeinsam lernen“ soll gemäß Absicht der Autoren ein Forum bilden, innerhalb dessen Projektdokumente einer größeren Öffentlichkeit zugänglich gemacht werden. Neben einer Ausbildung für Lehrpersonen, welche das Rückmeldeverfahren an ihren Schulen betreuen sollen, bilden individuell einsehbare Leistungsrückmeldungen mit Stärke-/Schwächeprofilen für die Schülerinnen und Schüler einen Teil der Rückmeldekultur.¹⁸⁶ Ebenso ist eine internetbasierte Datenaufbereitung für Lehrpersonen sowie für weitere Akteure des Schulsystems in Vorbereitung.

Kommunikation mit Lernenden und Eltern: Die Einführung von Bildungsstandards für Schülerinnen und Schüler macht Bildungserfolg sichtbar und vergleichbar, sei es auf individueller Ebene (soweit dies angestrebt wird) oder auf Schulebene. Lernende und ihre Eltern haben üblicherweise ein ausgeprägtes Interesse an Orientierungshilfen. Für Lernende kann die Transparenz zu verstärkter Selbstwirksamkeits- und Kontrollüberzeugung führen. Eltern, die sich regelmäßig darüber informieren können, was in der Schule geschieht, können ihre Kinder aktiver unterstützen. Gerade die Einbindung von Eltern ist ein wesentlicher Faktor des Schulerfolges von Kindern (vgl. Redding 2000). Dies ist mit ICT in vielfältiger Form möglich und sinnvoll.

- Viele Schulen kommunizieren bereits heute mit Schülerinnen und Schülern und Eltern per E-Mail. Andere stellen nicht-vertrauliche Informationen auf Schulhomepages zur Verfügung (vgl. Penuel/Kim/Michalchik et al. 2002). Diese Wege der Kommunikation sind zu fördern. Kindern und Eltern ohne privaten Internetzugang sollten entsprechende Möglichkeiten in der Schule angeboten werden.
- Im Rahmen von Vergleichsarbeiten und Lernstandserhebungen kommen – wie schon gezeigt wurde – internetbasierte Rückmelde-

¹⁸⁵ <http://www.gemeinsamlernen.at/> (Stand 29.08.06).

¹⁸⁶ <http://www.bildung-standards.at/> (Stand 3.09.06).

systeme zum Einsatz, die über geschlossene Bereiche auch Informationen zu Ergebnissen und der Leistungsentwicklung der Kinder übermitteln. Hier ist darauf zu achten, dass Verantwortliche (Testadministration, Lehrpersonen) beratende Funktionen wahrnehmen können.

- Schülerinformationssysteme (*SIS*; engl.: *student information systems*) gehen hier einen Schritt weiter – sie bringen das Klassenbuch (bzw. das Klassenheft, die Schulkontrolle, das Aufgabenheft) ins Internet, so dass sich Lernende und Eltern jederzeit informieren können (vgl. Petko 2005). Sie bieten Lesezugang zu den jeweiligen persönlichen Daten der Schulverwaltungssoftware und erweitern diese durch diverse Kommunikationsfunktionen. Dies kann zu breiterer Informiertheit führen, aber auch zu einer Erfahrung von Selbstwirksamkeit bei Lernenden und ihren Eltern sowie zu Gesprächsanlässen mit Lehrpersonen.
- Die Nutzung von Lernplattformen kann zu einem wichtigen Katalysator einer erweiterten Schulkultur werden, in der Lernenden größere Autonomie und zugleich individuellere Betreuung zuteil wird. Die Nutzung solcher kollaborativer Systeme kann außerdem zu einer Öffnung des Klassenraumes führen, wodurch Lernende und Klassen innerhalb einer Schule im Sinne einer „Knowledge Building Community“ in verstärktem Maße kooperieren (Scardamalia/Bereiter 1994), aber auch Kooperationen außerhalb der Schule verstärkt möglich werden. Durch die Nutzung von Lernplattformen lassen sich viele der angesprochenen Potenziale in sinnvoller Weise unterstützen.

5.1.6 Schulaufsicht und Schulinspektion

Mit der Zubilligung von mehr Selbstständigkeit und erweiterten Kompetenzen an die Einzelschule verändert sich auch die fach- und dienstrechtliche Rolle der Schulaufsicht. Es wäre widersprüchlich, den Schulen einerseits größere Autonomie zuzugestehen, andererseits aber das Netz dienstrechtlicher Einbindungen und die Interventionsdichte der Schulaufsichtsbehörden auf dem heutigen Stand zu belassen. Dies umso mehr, als mit der Einsetzung von Schulleitungen in Zukunft häufiger dienstrechtliche Funktionen auch an diese übertragen werden. Insgesamt ist die Diskussion über die Ausgestaltung und die künftigen hoheitlichen und dialogisch-beraterischen Aufgaben von Schulaufsichts-

behörden in den Auftragsländern im Fluss. Die Diskussionen bewegen sich dabei grundsätzlich zwischen zwei Arten einer periodischen, dienstaufsichtsbezogenen Evaluation/Inspektion: Zum einen werden die Möglichkeiten externer Meta-Evaluationen ausgelotet, welche sich im Wesentlichen darauf beschränken, die Güte der Selbstevaluation bzw. der schuleigenen Verfahren der Durchsetzung der Standards und der Qualitätssicherung in dialogischen Verfahren zu überprüfen. Zum anderen stehen Verfahren einer inhaltlichen Inspektion zur Debatte, welche direkt auf die Qualität des Unterrichts, der Schulleitung und des schuleigenen pädagogischen Profils abzielen und zu Urteilen über die Güte des Inputs und der Prozesse zu gelangen suchen. Was sich jedenfalls nicht abzeichnet, ist eine Gleichsetzung der Erweiterung schulischer Eigenständigkeit mit einem Laisser-faire ohne staatliche Kontrollen. Die systematische Überprüfung der tatsächlichen Schülerleistungen auf der Basis von Bildungsstandards dürfte in dieser Hinsicht künftig die wichtigste Maßnahme einer vermehrt über den Output erfolgenden Steuerung des Bildungswesens darstellen.

Wie in früheren Abschnitten gezeigt wurde, repräsentieren lokale Schulaufsichten bzw. Inspektorate wichtige Elemente einer demokratisch-föderalistischen Entwicklung. Traditionell hatten und haben diese Organe eine Evaluationsfunktion inne. In England etwa regeln seit 1902 die *Local Education Authorities* (LEA) das Bildungswesen auf lokaler Ebene. In den vergangenen Jahren verlor diese regionale Schulaufsicht im Zuge der Realisierung lokaler Schulmanagements ihre Bedeutung, zumal die Realisierung und Überwachung der seit geraumer Zeit eingeführten Leistungsstandards eine zentrale Aufsichtsbehörde, das nationale *Office for Standards in Education* (OFSTED) notwendig machte. Insgesamt kann also eine Veränderung der Aufsichtsbefugnis festgestellt werden, welche die Macht der lokalen Schulbehörden schwächt und diejenige der staatlichen Aufsichtsorgane stärkt.

Für die Niederlande gilt – wie an früherer Stelle ausführlich beschrieben – eine Kopplung zwischen zentraler Systemsteuerung und dezentraler Schulverwaltung. Während im Rahmen einer umfassenden Deregulierung und Dezentralisierung den Schulen ein hoher Grad an Autonomie ermöglicht wurde, die beispielsweise auch die Gestaltung schulinterner Curricula vorsieht, wurden auf der anderen Seite Qualitätssicherung und Evaluation durch das Schulinspektorat ebenfalls ausgebaut. Diese Stellen erfassen im Rahmen eines vierjährigen Inspektionszyklus relevante Daten über eine Schule auf einer Qualifikationskarte, welche im Internet veröffentlicht wird. Insgesamt zeigt sich hier

die Verbindung zwischen einem starken und differenzierten zentralistischen Aufsichtssystem und einer gleichzeitig weitgehenden Autonomie der jeweiligen Schule.

In Schweden wiederum wurden gute Erfahrungen mit einer Verlagerung der Inspektionsaktivitäten weg von einer inhaltlichen zu einer stärker verfahrensmäßigen Überprüfung gemacht, indem dialogisch angelegte Verfahren aufgewertet wurden, die sich auf die Prüfung der Qualität der Methoden der Selbstevaluation der Schulen beziehen.

Auch in Deutschland hat in den vergangenen Jahren eine Neuausrichtung der Aufgabentätigkeit der lokalen Schulbehörden stattgefunden (Brockmeyer 1997). Unter den Vorgaben „Hierarchieabbau, Dezentralisierung und Kundenorientierung“ (ebd., S. 126) soll unter anderem vermehrt ein partnerschaftlicher und dialogischer Umgang mit den Schulen angestrebt und sollen diese auf dem Weg zu erweiterter Selbstständigkeit auch beraten und unterstützt werden. Entsprechend wurde die Schulverwaltung und -aufsicht in den einzelnen Bundesländern in den vergangenen Jahren zum Teil umfassend in Richtung des Ausbaus von Beratungsfunktionen umstrukturiert. In Nordrhein-Westfalen besteht für Beamte der Schulaufsicht eine Pflicht zu „Dialog-Gesprächen“, die der Auswertung der zwischen Schule und Schulaufsicht geschlossenen Zielvereinbarungen im Prozess der Schulentwicklung und der Schulprogrammarbeit dienen. In Sachsen erhielt die Schulaufsicht eine stärker beraterische und stützende Funktion. In mehreren Ländern wurden vor allem „in Modellversuchen positive Erfahrungen mit Formen der zyklischen Evaluation der einzelnen Schulen durch Beratungsteams der Schulaufsicht bei gleichzeitiger kontinuierlicher Beaufsichtigung der Schule durch einen zugeordneten Schulaufsichtsbeamten im Schulamt gesammelt. Insgesamt zieht sich die Schulaufsicht tendenziell aus der personenbezogenen Betreuung zurück und übernimmt stärker Aufgaben der Unterstützung und Beaufsichtigung der Entwicklungsprozesse der Schule als eines Gesamtsystems“ (Avenarius/Ditton/Döbert et al. 2003, S. 162). Für das gesamte Bundesgebiet lässt sich somit ein Trend beschreiben, der für die lokalen Schulaufsichtsorgane einen Funktionswandel in Richtung Beratung und Unterstützung und für die Schulen eine gleichzeitige Erweiterung ihrer Eigenständigkeit vorsieht, wobei dieser Trend aber je nach Bundesland unterschiedlich klar akzentuiert ist.

In der Schweiz betrafen die in den vergangenen Jahren vorgenommenen Schulreformen unter anderem ebenfalls die Aufgabe und Rolle der

Schulbehörden (Rhy 1997). Unter dem Ansatz des „New Public Management“ (NPM) wurde in den meisten Kantonen eine Reorganisation der öffentlichen Verwaltung angestrebt, welche neben einer Effizienzsteigerung auch eine Neuregelung bestimmter Zuständigkeiten anstrebte. Kennzeichnend dafür sind etwa folgende Aspekte:

- Unterscheidung zwischen operativer und strategischer Führung,
- öffentliche Verwaltung als Dienstleistungs- statt als staatlicher Vollzugsapparat,
- Abbau von Bürokratie durch Delegation der Entscheidungskompetenzen nach unten sowie
- Output- und Kundenorientierung der Verwaltung (ebd., S. 166).

Gleichzeitig ist ein Trend zur Teilautonomie festzustellen, welche die pädagogische Leitung durch professionelle Schulleitungen vorsieht. Die lokalen Schulbehörden, deren Mitglieder in der Regel Laien sind, haben neben der Aufsicht in stärkerem Maße die Förderung und Beratung der jeweiligen Schulen zu realisieren. Die Aufsichtsfunktion wurde im Zuge des NPM vielerorts neu gestaltet, indem die örtlichen Schulpflegen kriteriengeleitete Mitarbeiterbeurteilungen vorzunehmen hatten. Eine Evaluation dieser in den neunziger Jahren eingeführten Maßnahme ergab für den Kanton Zürich, dass die Schulbehörden den Lehrkräften durchgängig gute bis überdurchschnittliche Leistungen attestieren und dass die Lehrpersonen die lohnwirksame Beurteilung mehrheitlich ablehnen (Wehner/Legler/Siegrist 2003). Die für die Bundesrepublik skizzierte Veränderung des Rollenverständnisses ist auch in der Schweiz zu beobachten, wobei im Zeichen von NPM gleichzeitig ein Macht- und Bedeutungszuwachs der Gemeindeschulpflegen möglich wird.

In Österreich ist die Schulaufsicht auf der Ebene der Bundesländer organisiert (Schratz 1997). Auch hier ist der Übergang von der klassischen Aufsichtsbehörde, welche den Unterricht inspiziert, hin zu einem prozessbegleitenden Gremium im Rahmen von Teilautonomie und Schulentwicklung zu beobachten. Nicht zuletzt angesichts der hohen Zahl von durch die Schulinspektoren zu betreuenden Lehrpersonen ist eine flächendeckende und kriteriengeleitete Erfassung der Unterrichtsqualität aus zeitlichen Gründen kaum möglich (ebd.). In der Folge wurde in den vergangenen Jahren eine professionelle Qualifizierung der Schulaufsichtspersonen realisiert, welche u.a. die Vermittlung von Kompetenzen in Systemberatung, Projektmanagement und Qualitätssi-

cherung anstrebt (ebd., S. 230ff.). Auch hier lässt sich der beschriebene Funktionswandel feststellen.

Die mit einer Teilautonomie einhergehende Aufwertung der Bedeutung der Einzelschule verlangt eine stärkere Selbst- und Fremdevaluation von Schule und Unterricht. Damit ändert sich auch die Rolle der Schulaufsicht. Sie kann nicht mehr als Kontrollmechanismus in einem streng hierarchisch geordneten System fungieren, sondern muss sich in den Dienst der Qualitätssicherung und -entwicklung stellen können. Eine gute Schulaufsicht unterstützt die Einzelschulen in ihrer Arbeit und ermöglicht ihnen die Verarbeitung von Rückmeldungen sowie Spiegelungen ihrer Arbeit von außen. Wie das Beispiel der Niederlande zeigt, muss dies nicht zu einem Bedeutungsverlust der Schulbehörden insgesamt führen, sondern kann mit einer Kompetenzerweiterung einhergehen. Voraussetzung dafür ist aber, dass eine Qualifizierung und Professionalisierung von Schulaufsichtspersonen stattfindet. Schratz (1996, S. 229) nennt in Anlehnung an bestehende Schulleitungsausbildungen fünf zentrale Bausteine zur berufsbegleitenden Professionalisierung von Schulaufsichtsorganen:

- Führungsmanagement und Systemberatung,
- Umgang mit Problemsituationen und systematisches Konfliktmanagement,
- Projektmanagement und Innovationstransfer,
- Qualitätssicherung und Evaluation von Schulen,
- Führungskräfteauswahl und -entwicklung.

Entsprechende Anforderungsprofile dürften allerdings die Möglichkeiten kleinerer Gemeinwesen übersteigen, was zu einer Schwächung lokaler Gremien bei gleichzeitigem Machtzuwachs staatlicher und zentralistischer Aufsicht führen dürfte.

In jedem Fall muss vermieden werden, dass der Schulaufsicht eine widersprüchliche Doppelrolle auferlegt wird, welche sowohl lohn- und qualifikationswirksame Maßnahmen als auch partnerschaftliche Beratung und Begleitung vereinigt. Nicht zuletzt im Hinblick auf die Akzeptanz des Inspektorats durch die Lehrpersonen ist in dieser Hinsicht zumindest auf eine personelle Entflechtung dieser unterschiedlichen Funktionen zu achten.

Bezogen auf die Implementierung von Bildungsstandards und abgeleitet von den Ausführungen zu Funktion und Stellung von Schulaufsichtsor-

ganen in den verschiedenen Ländern lassen sich zusammenfassend verschiedene Gelingensbedingungen von Schulaufsichtsarbeit ableiten:

- Ein hoher Grad an Qualifizierung und Professionalisierung des Inspektionspersonals,
- klar definierte Kompetenzen und Funktionen der Schulinspektoren,
- ein entwicklungs- und innovationsfreundliches Schulklima,
- eine hohe Bereitschaft der Schulaufsicht, die jeweilige Schule bei der Realisierung der Standards zu unterstützen,
- ein soziales Verhältnis, das professionelle Distanz und Vertrauensbildung berücksichtigt und sinnvoll ausbalanciert,
- eine etablierte und institutionalisierte Evaluationskultur, welche zwischen interner und externer Evaluation unterscheidet und eine maßgeschneiderte Schulentwicklung ermöglicht und vorsieht sowie
- ausgezeichnete und allseitige Kenntnisse der zu realisierenden Standards und des Entwicklungsbedarfs, der sich für die jeweilige Schule ergibt.

Damit wird auch zum Ausdruck gebracht, dass die Arbeit von Schulaufsichtsorganen ebenfalls an Standards ausgerichtet werden soll. Entsprechende Standards könnten sich an den oben aufgeführten Bedingungen orientieren und überprüfbar, der Schulentwicklung dienende Tätigkeiten der Inspektorate zum Gegenstand haben. So könnte beispielsweise die Qualität der Rückmeldungen nach Evaluationen erfasst und beurteilt werden. Grundsätzlich kann eine Evaluation durch die Schulaufsicht nur dann Sinn ergeben, wenn sie zu konkreten, auf Veränderungen abzielenden Handlungen führt. Solcherlei Qualität der Feedbacks wäre von den Betroffenen, also den Lehrpersonen, aufgrund zu erarbeitender Kriterien zu beurteilen und könnte insgesamt zu einer innovations- und entwicklungsfreundlichen Schulkultur beitragen.

Schritte in eine entsprechende Richtung wurden etwa bei der Schaffung des Referenzrahmens Schulqualität in Hessen unternommen. Ausgehend von einem Input-Prozess-Output-Modell zu Schul- und Unterrichtsqualität wurden sieben Qualitätsbereiche festgelegt. Den dazugehörigen Dimensionen wurden Kriterien zugeordnet, welche schulförmenspezifisch variiert oder ergänzt werden können. Ziel ist eine möglichst zutreffende Abbildung und Erfassung von Schulqualität über die Identifizierung von Stärken und Schwächen einer einzelnen Schule durch professionell arbeitende Schulinspektorinnen und Schulinspekto-

ren. In Zusammenarbeit mit Fachpersonen wird zudem laufend die Passung des Konzepts an die wissenschaftliche Befundlage sowie an die Bedürfnisse der praktischen Umsetzung geprüft. Die Schulinspektoren orientieren sich in ihrer Arbeit an den vom „Joint Committee on Standards for Educational Evaluation“ formulierten Prinzipien der Fairness, der Nützlichkeit, der Durchführbarkeit und der Genauigkeit. Ebenfalls sollen mit schriftlichen Arbeitsvereinbarungen Transparenz und Verbindlichkeit zwischen Schulen und Evaluatoren hergestellt werden. Die Datenerhebung umfasst Fragebögen, Beobachtungen, Interviews und die Analyse von Dokumenten (Hessisches Kultusministerium 2007).

5.1.7 Reformprojekte und Modellversuche

Regionale, länder- und landesbezogene *Reformprojekte* können als „Vehikel“ für die Einführung, Erprobung und Umsetzung von Bildungsstandards gesehen werden. Zwei Arten können unterschieden werden: *Modellversuchsprogramme* und *flächendeckende, politisch initiierte Reforminitiativen*. Gemeinsam ist ihnen, dass beide eine ganzheitliche Strategie der Implementation innovativer Ideen in die Breite sowie einen damit verbundenen situierten Transfer von komplexem Wissen verfolgen. Dies im Sinne einer „Research-based-whole-school-reform“-Strategie (im Gegensatz zu einer „High-stakes-accountability-testing“-Strategie) in der Beschreibung von Lam (2004). Ersterer werden deutlich positivere Effekte bezüglich schulbezogener Qualitätsentwicklung zugeschrieben. Während behördlich initiierte Flächenprojekte den verbindlichen Charakter politisch abgestützter Reformen besitzen, sind Modellversuchsprogramme stärker mit dem „Makel der Exklusivität“ bzw. der Nicht-Alltäglichkeit behaftet; sie bedürfen über die Abwicklung eines Grundprogramms hinaus zusätzlicher Anstrengungen zu ihrer systemweiten Verbreitung.

Unter den *Modellversuchsprogrammen* gehören die Programme der deutschen Bund-Länder-Kommission (BLK) zu den prominentesten. Eines der bekanntesten Programme ist das SINUS-Projekt zur „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“. Dessen Implementierung erfolgte als Reaktion auf das schlechte Abschneiden deutscher Schülerinnen und Schüler vor allem in der internationalen TIMSS-Schulleistungsstudie (Prenzel/Carstensen/Senkbeil/Ostermeier/Seidel 2005). Am zwischen 1998 und 2003 durchgeführten Grundprogramm nahmen 180 in Schulnetzwerken (Schulsets) organi-

sierte Schulen der Sekundarstufe I aus mehreren Bundesländern teil. Unter der Zielsetzung der Initiierung einer dauerhaften, unterrichtsbezogenen Qualitätssicherung und der Professionalisierung der Lehrpersonen erhielten die an SINUS teilnehmenden Schulen verschiedener Schultypen (Gymnasium, Real- und Hauptschulen sowie Integrierte Gesamtschulen und Schulen mit mehreren Bildungsgängen) die Gelegenheit, sich ausgehend von selbst identifizierten Problemstellungen wahlweise in elf verschiedenen thematischen Modulen problemorientiert weiterzubilden:

- (1) Weiterentwicklung der Aufgabenkultur im mathematisch-naturwissenschaftlichen Unterricht;
- (2) Naturwissenschaftliches Arbeiten;
- (3) Aus Fehlern lernen;
- (4) Sicherung von Basiswissen – Verständnisvolles Lernen auf unterschiedlichen Niveaus;
- (5) Zuwachs von Kompetenz erfahrbar machen: Kumulatives Lernen;
- (6) Fächergrenzen erfahrbar machen: Fächerübergreifendes und -verbindendes Arbeiten;
- (7) Förderung von Mädchen und Jungen;
- (8) Entwicklung von Aufgaben für die Kooperation von Schülern;
- (9) Verantwortung für das eigene Lernen stärken;
- (10) Prüfen: Erfassen und Rückmelden von Kompetenzzuwachs;
- (11) Qualitätssicherung: innerhalb der Schule und Entwicklung schulübergreifender Standards.

Die Teilnehmenden waren aufgefordert, an ihren Schulen Fachgruppen zu bilden, welche in Bezug auf die jeweiligen Modulziele kooperierten. Unterstützt wurden die gemeinsam an der Verbesserung ihres Unterrichts arbeitenden Lehrkräfte dabei von einem regionalen, mit dem Programmträger verbundenen Netzwerk von Koordinatoren und Beratern (z.B. Schulsets). Die Ergebnisse wurden in regionalen Netzwerken und auch länderübergreifend mit anderen Schulen ausgetauscht. Das Projekt wurde wissenschaftlich begleitet und als bisher einziges Modellversuchsprogramm auch anhand von Außenkriterien (und nicht nur auf der Basis subjektiver Einschätzungen und Meinungspapiere der beteiligten Akteure) evaluiert. Im Vergleich zu einer Teilstichprobe der deutschen PISA-Schulen zeigten sich verschiedene Unterschiede zugunsten der SINUS-Schulen. So schilderten die Schülerinnen und Schüler etwa ein stärkeres mathematisches Interesse und einen in verschiede-

nen Belangen veränderten Unterricht. Gleichzeitig erbrachten sie bessere Fachleistungen als ihre Kameradinnen und Kameraden aus der Vergleichsgruppe. Bemerkenswert erscheint, dass diese Unterschiede in den tieferen Schulniveaus stärker akzentuiert sind (ebd.).

Unter dem Begriff „SINUS-Transfer“ wurde das in der Evaluation seiner Implementation als erfolgreich bewertete Programm der Länder zur kooperativen Unterrichtsentwicklung in Mathematik und Naturwissenschaften als Disseminations- bzw. Anschlussprojekt mit mehr als 1.700 teilnehmenden Schulen weitergeführt mit dem Ziel, die Rezeption der im Grundprogramm gewonnenen Erfahrungen in die Breite zu fördern und diese auch an die Grundstufe zu adaptieren. Dabei wurde – durchaus erfolgreich – versucht, Lehrerfortbildungsinstitutionen und universitäre Lehrerbildungseinrichtungen an das Projekt zu binden und dessen Kernideen im System zu verankern. Weiter sind Abstimmungen mit Schulleitungen, Schulaufsichtsbehörden und Eltern Bestandteil der Transferstrategie. Was die Lehrkräfte anlangt, so lässt sich beobachten, dass je länger diese im Programm mitarbeiten, desto seltener Wünsche nach konkreter Anleitung und Zielvorgaben formuliert werden – was im Sinn einer größeren Autonomie beim Umgang mit den Spielräumen des Programms interpretiert werden kann.

Neben der Durchführung von Modellversuchsprojekten besteht eine weitere erfolgversprechende Strategie darin, die Einführung einer Innovation, und damit auch von Bildungsstandards, als *regionales oder länderübergreifendes, mehrjähriges Reformprojekt zu vertreten*. Mit einer Flächenstrategie der Implementation kann deutlich gemacht werden, dass es sich um eine sehr komplexe Aufgabe handelt, die einen langen Zeitraum in Anspruch nehmen wird. Erfahrungen mit dem von der Bertelsmann-Stiftung zusammen mit dem Ministerium für Schule und Weiterbildung des Bundeslandes Nordrhein-Westfalen realisierten Projekt „Selbstständige Schule“ zeigen, dass hier leicht Laufzeiten von acht und mehr Jahren erreicht werden können. Das 2002 gestartete und wissenschaftlich begleitete Projekt strebt qualitative Verbesserungen in der Schul- und Unterrichtsarbeit an. Erreicht werden sollen sie über qualitätsorientierte Selbststeuerung an Schulen und die Entwicklung regionaler Bildungslandschaften. Die 278 teilnehmenden Schulen werden in 19 Modellregionen zusammengefasst. Mit bedarfsgerechten Fortbildungen, an denen sich sämtlich Angehörige eines Schulteams zu beteiligen haben, soll primär Unterrichtsentwicklung angeregt werden. Daneben werden aber auch die Bereiche schulinternes Management, Personalentwicklung,

Ressourcenmanagement, Unterrichtsorganisation sowie Mitwirkung und Partizipation berücksichtigt (Selbständige Schule 2007). Die wissenschaftliche Begleitung erfasst zu verschiedenen Zeitpunkten Aussagen der Schulleitungen, Leistungstests der Schülerinnen und Schüler sowie Befragungsergebnisse der hauptsächlich betroffenen Akteure. Dem Zwischenbericht nach stehen sowohl Lehrpersonen als auch Schulleitungen dem Vorhaben recht positiv gegenüber und berichten u.a. positive Veränderungen in den Arbeitsfeldern Unterrichtsorganisation und -gestaltung, innere Organisation, Mitwirkung und Qualitätssicherung. Allerdings müssen das Ende des Projekts und die Schlussevaluation abgewartet werden, um schlüssige Aussagen darüber formulieren zu können, inwieweit es über den gesamten Zeitraum hinweg gelungen ist, bei den Lehrpersonen und Schulleitungen ein neues Verständnis von Schule und Unterricht und eine nachhaltige Akzeptanz des Reformvorhabens zu erreichen (Holtappels/Rolff 2005; Daschner 2006).

Zuvor waren im Projekt „Schule & Co“ („Stärkung von Schulen im kommunalen und regionalen Umfeld“; Laufzeit: 1997-2002), an dem anfangs 52 und am Schluss 89 Schulen teilgenommen hatten, bereits konkrete Erfahrungen mit Problemen der Schulentwicklung, darunter der konsequenten Verschränkung von Unterrichtsentwicklung und Schulentwicklungsmanagement (u.a. durch den Einsatz von Steuergruppen als innerschulisches Kooperationsinstrument) gesammelt worden (Rolff/Bastian 2002; Rolff 2006), die in das Projekt „Selbständige Schule“ einfließen und als Wissenstransfer zu verstehen sind. Wie die Hürden der Schulbürokratie überwunden werden können, musste im Verlaufe des Projektes oft mühsam gelernt werden, ebenso wie die Gestaltung des Umgangs mit den politischen Meinungsträgern im Land. Auch musste eine Strategie gefunden werden, was sich gegenüber Beharrungskräften tun lässt, die Innovationsstrategien abwarten und nach Ende des Projekts zu den alten Strategien zurückkehren wollen. Es ist nicht damit zu rechnen, dass Projekte wie dasjenige der Implementation von Bildungsstandards ohne Widerstand akzeptiert werden, und die wirkungsvollste Form von Widerstand ist die des Wartens auf Scheitern.

Ohne Folgeprojekt und somit eine Streckung der zeitlichen Dauer wären die Erfahrungen von „Schule & Co“ wohl versandet und die grundsätzlichen Widerstände hätten nicht überwunden werden können. Das wiederum bedeutet, Prozesse erfolgreichen Wandels, die in den Schulen tatsächlich etwas bewirken können, mit langen Laufzeiten und hoher Kontinuität zu versehen. Sonst werden auch erfolgreiche Pro-

jekte abgebrochen, ohne dass sie ihre Erfahrungen transferieren können.

Die Bedeutung der Erfolgsfaktoren *lange Dauer, hohe Kontinuität und politischer Rückhalt* lassen sich am Projekt „Selbstständige Schule“¹⁸⁷ belegen. Das Projekt geht vom Konzept der *Accountability* aus, also der stärkeren Verantwortung der einzelnen Einheit, die höhere Autonomie und andere Formen von Kontrolle voraussetzt. Das *Durham Board of Education* aus Ontario in Kanada¹⁸⁸ erhielt 1996 den mit € 150.000 dotierten Carl-Bertelsmann-Preis und war für die Entwicklung der Stiftungsprojekte im Bereich Qualitätsentwicklung in Schulen in verschiedener Hinsicht maßgebend. Das Beispiel zeigt auch, wie wirkungsvoll gute Beispiele aus dem Ausland sein können, wenn sie mit Sinn für die systemischen Unterschiede und ausreichend Zeit für die Entwicklung adaptiert werden. Übertragbar sind vor allem grundlegende Prinzipien und Verfahren der Qualitätssicherung.

Das Board in Durham verfolgt seit Jahrzehnten eine intensive Politik der regionalen Schulentwicklung, die darauf setzt, dass Bildung als herausragender Standortfaktor verstanden wird. Eingebunden sind alle Akteure, auch und gerade die Eltern, deren Ressourcen als Erfolgsfaktoren verstanden werden. Die Entwicklung wird durch das Board gesteuert, setzt aber eine hohe Autonomie der einzelnen Schulen voraus. Ein Erfolgsfaktor war zudem die Veränderung des Unterrichts durch neue Methoden, etwa solche des kooperativen Lernens. Derzeit gilt ein Systemplan der Entwicklung, der von 2004 bis 2009 reicht und klare Ziele in den Bereichen *Accountability* (Verantwortung der Schule), *Student Achievement* (Leistung im Unterricht) und *Equity* (Chancengleichheit in Schule und Umwelt) verfolgt (Durham Directions 2003).

Diese Zielsteuerung durch befristete Vorgaben, die nach einer bestimmten Zeit überprüft werden, ist für die meisten Schulen und Behörden im deutschen Sprachraum sehr ungewöhnlich. Pläne zur Systementwicklung über einen absehbaren Zeitraum liegen so gut wie nicht vor. Typisch ist die zeitlich unbestimmte Steuerung mit eher vagen Vorgaben wie allgemeinen Bildungszielen in Lehrplänen, die die Umsetzung offen lassen und keine Frist setzen, zu der das Erreichte überprüft wird. Bildungsstandards aber verlangen genau eine solche Organi-

¹⁸⁷ <http://www.selbststaendige-schule.nrw.de> (Stand 03.12.07).

¹⁸⁸ <http://ddsb.durham.edu.on.ca/DDSBMain.htm> (Stand 12.08.06).

sation, die Zielvorgaben an bestimmten Zeitpunkten überprüft. Das gilt auch für die Implementation von Bildungsstandards, die einen möglichst präzisen Zeitrahmen erhalten muss, der in der deutschen Expertise nur mit Hilfe von Verfahrensschritten angedeutet wird (Klieme/Avenarius/Blum et al. 2003, S. 121ff.).

Das Projekt „Selbstständige Schule“ nimmt entsprechend diesen Vorgaben und in Zusammenfassung der eigenen Erfahrungen sechs verschiedene Einflussfaktoren an, die sich auf Erfolg oder Misserfolg nachhaltig auswirken können. Unterschieden werden:

- Motivation und Kompetenzen für Innovation seitens der Akteure,
- gesetzliche Bestimmungen, die das Vorhaben absichern,
- Steuerung auf der Ebene der Einzelschule,
- regionale Vernetzung, Koordination und Steuerung,
- politischer Wille im Land und in der Region sowie
- neue und effektive Instrumente.

Auffällig an beiden Projekten, dem kanadischen wie dem deutschen, ist die stark in den Mittelpunkt gestellte Bedeutung der *Region*. Sie wird zu einer Handlungseinheit eigener Art, die auch bei der Implementation von Bildungsstandards als solche beachtet werden muss.

Während die Expertise von 2003 noch stark auf die Vorarbeiten und Entwicklungen in den einzelnen Bundesländern setzt (ebd., S. 131ff.), betont das größte Projekt regionaler Schulentwicklung in Deutschland die Vernetzung autonomer Schulen auf Kreis- oder Bezirksebene. Schulentwicklung hat offenbar dann größere Chancen der Innovation und Nachhaltigkeit, wenn nicht lediglich die einzelne Schule als Akteur und Adressat angesprochen wird, sondern regionale Netzwerke zur Verfügung stehen. Derzeit sind an dem Projekt 278 Schulen in 19 Regionen Nordrhein-Westfalens beteiligt, die Netzwerke bilden und in ihrer gemeinsamen Schulentwicklung unterstützt werden. In verschiedenen Regionen wurden „Korrespondenzschulen“ aufgebaut, denen eine besondere Modellfunktion zukommt und die vom Projekt speziell unterstützt werden.¹⁸⁹

¹⁸⁹ Etwa: „Korrespondenzschule“ Bezirksregierung Detmold für Ostwestfalen-Lippe:
<http://www.bezregdemold.nrw.de/Aufgaben/Organisation/Abteilung4/Bildungsregion/VO/Korrespondenzschule/index.html> (Stand 15.08.06).

Diese Netzwerke bestehen nicht nur aus Schulen, sondern stellen einen größeren Verbund dar, der Kommunen, Verbände, andere pädagogische Institutionen und auch einzelne Personen umfasst. Die Netzwerke müssen aufgebaut und konsolidiert werden, was dadurch erleichtert wird, dass Schulentwicklung und Qualitätssicherung als zentrale Standortfaktoren verstanden werden. Projekte der regionalen Entwicklung sind so immer auch Schul- und Bildungsprojekte, die enge Formen der Abstimmung und Kooperation verlangen. Ein Beispiel für die Organisation solcher regionaler Netzwerke stellt etwa die Akademie für Information und Management in Heilbronn¹⁹⁰ dar, die Schulentwicklung im Blick auf die örtlichen Schulen und ihr kommunales Umfeld anregt.

Für regionale Netzwerke spricht auch, dass nur so eine vertretbare Größe der Schulentwicklung erreicht werden kann. Die deutschen Flächenländer sind in aller Regel zu groß und zu heterogen, um als Handlungseinheiten und Träger der Entwicklung angesehen werden zu können. Zudem ist die Schuladministration kaum auf Netzwerke eingestellt, die sich nach örtlichen Faktoren ausrichten. Dazu gehören die Unterscheidung von Einzugsgebieten, die Siedlungsbewegungen in der Bevölkerung, die demografischen Entwicklungen, die Infrastruktur und die Verteilung der Angebote etwa im Bildungsbereich. „Standorte“ sind so nicht einfach Kreise oder Bezirke, sondern kooperative Einheiten, die einen bestimmten Raum entwickeln wollen. Das allgemeine Angebot an Bildungsmöglichkeiten und speziell das Angebot der Schulen in einer Region sind dabei in einer zunehmend mobilen Gesellschaft ausschlaggebende Faktoren.

Daher ist es kein Zufall, wenn im Projekt „Selbstständige Schule“ der Bertelsmann-Stiftung und des Ministeriums für Schule und Weiterbildung NRW der politische Wille *in der Region* als Erfolgsbedingung herausgestrichen wird. Neue Gesetze und neue behördliche Vorgaben auf Landesebene reichen mit Sicherheit nicht aus, um das Bildungssystem wirkungsvoll zu verändern. Vergleichbar dem Board in Durham müssen regionale Steuerungssysteme und Verantwortlichkeiten aufgebaut werden, die die Steuerung auf Landesebene und in den einzelnen Schulen wirkungsvoll unterstützen und ergänzen. Das gilt auch für die Implementation von Bildungsstandards, die nicht einfach den Verordnungsweg voraussetzen kann. Wenn die Schulentwicklung umso mehr Erfolg

¹⁹⁰ <http://www.aim-ihk.de> (Stand 15.08.06).

hat, je besser die regionalen Netze geknüpft und ausgebildet sind, dann gilt das auch für die Schulevaluation einschließlich des Einsatzes neuer Instrumente und Rückmeldeverfahren.¹⁹¹

Modellversuche und regionale Flächenprojekte wie das geschilderte bieten die Chance, schulnahe und an bestimmten Problemlagen orientierte Unterrichtsentwicklung „in großem Stil“ anzuregen. Bei der Einführung von Bildungsstandards könnten entsprechende Projekte wertvolle Dienste leisten, indem sie den Blick auf die fachbezogene Unterrichtsarbeit lenken und mit der Möglichkeit fachspezifischer Kooperation eine Reflexionsbasis für Unterrichtsarbeit schaffen. Zu den *strukturellen Erfolgsbedingungen*¹⁹² und *Merkmale der Implementation* von in die Breite zielenden Reformprojekten gehören – im Lichte von Erfahrungen, die in bisherigen Modellversuchen gesammelt wurden (vgl. Euler/Sloane 1998 sowie Nickolaus/Ziegler/Abel/Eccard/Aheimer 2006), sowie des internationalen Forschungsstandes zur Transfer- und Innovationsforschung (vgl. Gräsel/Jäger/Willke 2006):

- die politische und curriculare Absicherung der Reformziele,
- die schulnahe und problemorientierte Gestaltung und Durchführung,
- die Verfügbarkeit von Ressourcen, die sich insbesondere auf die Einrichtung von Unterstützungssystemen und die Ermöglichung von Kooperation beziehen,
- die Zuweisung einer aktiven Rolle an die Lehrkräfte bereits bei den Entwicklungsarbeiten,
- die Bereitstellung von Handreichungen, Materialien und darauf bezogenen Fortbildungsangeboten,
- eine Langzeitperspektive und ein kontinuierlicher politischer Willen während des gesamten Zeitraums – auch dann, wenn es Rückschläge gibt,
- ein qualitativ hochstehendes Kommunikations-, Informations- und Projektmanagement,

¹⁹¹ Schulevaluation „innerhalb einzelner Länder oder Regionen“ wird auch von der Expertise angedeutet (Klieme/Avenarius/Blum et al. 2003, S. 143).

¹⁹² Zu den systemisch-strukturellen Bedingungen hinzu treten einzelschulische (z.B. die Qualität von Schulleitungen und kollegialer Kooperation) und personale (Ausmaß von Problembewusstsein, Kompetenzerleben, Angstgefühlen) Bedingungen und Erfolgsfaktoren.

- eine aussagekräftige Selbst- und Fremdevaluation, die sich auch auf den „Output“ der Lernergebnisebene bezieht, sowie
- ihre wissenschaftliche Begleitung als Längsschnittprojekte.

Angesichts der Komplexität von Reformen und der Vielfalt offener Fragen in Bezug auf Kriterien und Merkmale eines erfolgreichen Designs von systemweiter Innovation einschließlich der Bestimmung ihrer Wirkungen, ist dem in der Expertise „Innovation und Transfer“ gemachten Vorschlag der Einrichtung eines Transferforschungsprogramms (Nickolaus/Ziegler/Abel/Eccard/Aheimer 2006) – gerade auch im Hinblick auf die produktive Begleitung und Unterstützung der Einführung von Bildungsstandards – nur zuzustimmen.

5.2 Schulebene

Im Rahmen der Entwicklung empirischer Schulforschung wurde in den siebziger und achtziger Jahren des vergangenen Jahrhunderts die Bedeutung der Einzelschule als pädagogischer Handlungseinheit herausgearbeitet. In den Arbeiten von Rutter (1983) und Rutter/Maughan/Mortimor/Ouston (1979) wurden die Leistungen der Schülerinnen und Schüler in verschiedenen englischen Gesamtschulen untersucht und deren Lernvoraussetzungen statistisch kontrolliert. Dabei ergaben sich beträchtliche Unterschiede zwischen den einzelnen Schulen, deren Ursache der Autor in den Kompetenzen und in der unterschiedlich ausgeprägten Verantwortungsbereitschaft der Kollegien identifizierte. Im deutschen Sprachraum waren es u.a. die Arbeiten von Fend (1987, 1998), in welchen die Wahrnehmungen verschiedener schulischer Prozesse (Arbeitszufriedenheit, Schulleben, Kollegialität, Konflikte, Schülerzuwendung und Gleichgültigkeit) in Beziehung zu beobachteten Wirkungen (Vertrauen zur Lehrerschaft, Schulfreude, Disziplinlosigkeit und Verschüchterung) gesetzt wurden. Dabei kam der Autor zum Schluss, dass sich gute und schlechte Schulen im Wesentlichen hinsichtlich der sozialen Integration des Lehrkörpers, der Reichhaltigkeit des Schullebens, der Arbeitszufriedenheit der Lehrkräfte und der pädagogischen Zuwendung zur Schülerschaft unterscheiden (Fend 1987). Neuere Ansätze nehmen die erwähnte Komplexität zum Ausgangspunkt, um mit Mehrebenen- und Längsschnittmodellen Aspekte sowohl der Schul- als auch der Unterrichtsqualität in Beziehung zu Wirkungen (Leistungen,

Haltungen und Einstellungen) zu setzen (Ditton/Merz/Edelhäusser 2002).

Aus den Arbeiten folgt ein mehrdimensionales Qualitätsmodell, das die Einzelschulebene fokussiert. Neben einem Verständnis von Qualität als Exzellenz oder als Perfektion, wird ein dynamischer und transformativer Qualitätsbegriff postuliert (Harvey/Green 2000). Hohe Qualität ist keine Eigenschaft, die man auf Dauer einfach „hat“, sondern etwas Prozesshaftes, das es permanent zu sichern und weiterzuentwickeln gilt. Schul- und Personalentwicklung im Rahmen der Pflicht zur (Selbst-) Evaluation spielen im heutigen Verständnis dessen, was eine „gute Schule“ ist, eine zentrale Rolle. Die Schule wird als „Lernende Organisation“ (Rolff 1998) gesehen, die sich als komplexes Gebilde nur beschränkt von außen steuern lässt. „Lernende Schulen“, die ebenfalls teilautonome Schulen sein sollen, müssen sich deshalb bis zu einem hohen Grad selber organisieren und steuern können, gekoppelt mit der Verantwortung, für die Erreichung und Weiterentwicklung eines bestimmten Niveaus an Ergebnis- bzw. Outputqualität zu sorgen. Diese Qualität soll u.a. mit Bildungsstandards, aber auch mit der Selbstevaluation sinnvoll ergänzender Fremdevaluation periodisch überprüft werden.

Für den Transformationsprozess von einer reinen Inputsteuerung zu einer vermehrt ergebnisorientierten Steuerung kommt den Einzelschulen deshalb eine zentrale Funktion zu.

Auf der Ebene der pädagogischen Arbeit der Einzelschulen bedeuten Bildungsstandards die Entwicklung von Modellen einer permanenten Sicherung und Weiterentwicklung der Qualität des eigenen professionellen Handelns. Schulentwicklung auf der Basis von Bildungsstandards erfordert eine kooperative Verständigung über die zu erreichenden Ziele, die Übernahme von Verantwortung für das Erreichen der gesetzten Ziele, die Verpflichtung zur Rechenschaftslegung über die erreichten Ergebnisse und die Fähigkeit, auf der Basis von Rückmeldungen und Diagnosen das eigene Handeln zu verbessern. Dies erfordert neben einer Veränderung des eigenen Aufgabenverständnisses und der Verfügbarkeit administrativer und professioneller Begleitung die Entwicklung sozialer und materialer Werkzeuge.

Sollen Bildungsstandards nicht primär unter dem Zweck einer administrativ verordneten Testkultur, sondern als im Dienst einer erfolgreichen Schulentwicklung im Kontext von einerseits teilautonomen, andererseits rechenschaftspflichtigen Schulen wahrgenommen werden, stellen sich eine Reihe von Fragen und Herausforderungen, die sich auf die

Instrumente und Verfahren der Durchsetzung der neuen Steuerungsphilosophie beziehen:

- Mit Hilfe welcher Verfahren und Werkzeuge können die den Einzelschulen im Zuge der Einführung von Bildungsstandards zuwachsenden neuen Aufgaben erfüllt werden? (Einstieg in eine neue Testkultur; Entwicklung von Schulprogrammen, die auf Kompetenzentwicklung bezogen sind; Formen der Qualitätssicherung auf Schul-, Klassen und Individualebene; Mitarbeit bei der Entwicklung und Gebrauch von Diagnoseinstrumentarien für diese unterschiedlichen Aufgaben; Etablierung von Rückmeldestrukturen sowie von Förder- und Unterstützungsangeboten für Schülerinnen und Schüler, welche minimale Standards nicht erreichen, usw.)
- Aufgaben der administrativen und fachlich-pädagogischen Selbstorganisation, Leitung und Profilgestaltung (z.B. durch die Nutzung schuleigener Budgets, die Arbeit an einem Schulprogramm) der bis zu einem gewissen Grad mit rechtlichen, finanziellen und personellen Befugnissen ausgestatteten, „teilautonomen“ und „rechnungspflichtigen“ Schule.
- Selbstorganisation in Bezug auf eine effiziente Nutzung von professionellen Ressourcen, wie sie die Einrichtungen der Schulaufsicht, der Landesinstitute, der Lehreraus- und -fortbildung sowie kollegiale Netzwerke der Schul- und Unterrichtsentwicklung zur Verfügung stellen, für die Erfüllung zentraler Aufgaben.

Insgesamt stellt sich für die im Zentrum der pädagogischen Qualitätssicherung stehenden Einzelschulen die Herausforderung, Bildungsstandards in ihrer Funktion als Zielvorgaben, Orientierungshilfen und im günstigsten Fall als Wegbeschreibungen und Handwerkszeug für die eigene Qualitätsentwicklung zu nutzen, so dass der gesamte Prozess als Unterstützung auf dem Weg zu besserer Leistung und mehr Professionalität und nicht als Bedrohung empfunden wird. Nur Schulen, welche Bereitschaft zeigen, Verantwortung bei Innovationen zu übernehmen und diese aktiv voranzutreiben, werden zu einer nachhaltigen Erneuerung fähig sein.

Vier Merkmale, durch die sich gemäß internationalem Erfahrungs- und Forschungsstand innovative, in Transfer- und Transformationsprozessen erfolgreiche Schulen auszeichnen, und welche als wichtige Gelingensbedingungen ihrer Weiterentwicklung gelten können, sind eine

unterstützende Schulleitung, die produktive Kooperation in Fachkonferenzen, Kollegien und schulübergreifenden fachlichen Netzwerken, die *Selbstidentifikation der Schule als professioneller Gemeinschaft* mit einem pädagogischen Profil sowie eine offene und vielfältig instrumentierte *Kultur der Selbst- und Fremdevaluation*.

5.2.1 Schulleitung

Mit der Tendenz, Schulen eine teilweise oder sogar volle Autonomie zu gewähren, ist eine Heraushebung und Aufwertung der Rolle von Schulleitungen, die nach anerkannten Regeln funktionieren, verbunden. Ursprünglich bestand diese überwiegend in der administrativen Organisation einer Schule – pädagogische Leitungsaufgaben wurden dagegen nachgeordnet wahrgenommen (Schratz 1996). Im bisherigen Verständnis verstand sich der Schulleiter als meist dem Kollegium entstammender „Primus inter pares“. Teilautonomie impliziert jedoch die Vorstellung, dass Schulen sich im Sinne „Lernender Organisationen“ weiterentwickeln (Rolff 1998) und Aufgaben übernehmen, die sie bisher nicht in diesem Umfang wahrgenommen haben. In Lernenden Organisationen werden Führungskräfte als zuständig für den Aufbau und die Weiterentwicklung einer Institution gesehen; sie sind insgesamt für die Strategie und das Lernen der Organisation verantwortlich (Senge 1996). Auch wenn die wenigsten Entwicklungsarbeiten von den Schulleitungen – die im Übrigen aus einer oder mehreren Personen bestehen können – selbst durchgeführt werden, sind diese für deren Initiierung, Organisation und Moderation und damit insgesamt für die *qualitätsorientierte Selbststeuerung* der Schulen verantwortlich. Diesem gewandelten Verständnis von Schulleitung entspricht auch ein Wandel in der Begrifflichkeit. So etwa spricht Dubs (2005a) von Leadership sowie normativem und strategischem Management.

Entsprechend vielfältig sind die Aufgabenbereiche, welche zu einem neuen Leitungsverständnis einer Schule gehören. Folgende Aspekte können in Anlehnung an Schratz (1996) genannt werden:

- Bürokratische Administration und Organisation,
- Führung und Leitung mit besonderer Berücksichtigung der internen und externen Kommunikation und Kooperation,
- schulinterne Erarbeitung von Schulprogramm und von Konzepten der Personalentwicklung und Fortbildung,
- Wahrnehmung von pädagogischen Beratungsaufgaben,

- Anregung und Inspiration der Schul- und Organisationsentwicklung,
- Pflege der Beziehungen zu den (lokalen) Schulaufsichtsbehörden,
- Vermittlung zwischen schulischem Umfeld, Elternschaft und der Schule,
- Außenvertretung der Schule und Ansprechstelle für außerschulische Partner und Netzwerke und
- interne Evaluation und Koordination von Verfahren der externen Qualitätskontrolle und der Verarbeitung von Rückmeldungen.

Diese Aufzählung entspricht in weiten Teilen auch den Rahmenbedingungen, welche immer mehr Gesetzgeber für Schulleitungen vorsehen. Beispielhaft sind nachfolgend die Aufgaben, die das neue Volksschulgesetz des Kantons Zürich vorsieht, aufgeführt (Kanton Zürich 2005, S. 11):

- Administrative und personelle Führung der Schule,
- Mitwirkung bei Personalgeschäften der Schulpflege,
- Durchführung von Mitarbeitergesprächen und Mitwirkung bei der Mitarbeiterbeurteilung,
- Zuteilung der Schülerinnen und Schüler zu den Klassen,
- Förderung und Koordination der Weiterbildung der Lehrpersonen,
- Verwaltung der der Schule zugeteilten Mittel,
- Leitung der Schulkonferenz.

Zudem und unter Mitwirkung der Schulkonferenz:

- Qualitätsentwicklung und -sicherung in der Schule,
- Festlegen von besonderen Unterrichts- und Organisationsformen, wie Projektwochen, Klassenlager, Exkursionen und
- Festlegen der Stundenpläne.

Den Schulleitungen kommt insgesamt also eine zentrale Rolle zu bei der Umstellung zu (teil-)autonomen Strukturen, zumal Veränderungen auch Angst und Unsicherheiten bei den Betroffenen auslösen können. Rolff (1993, S. 186) spricht deshalb von der „transformationalen Führung“, die versucht, eine kollegiale und professionelle Schulkultur aufzubauen, Lehrpersonen in ihrer Entwicklung zu fördern und die darauf achtet, dass die Problemlösekapazität des Kollegiums gefördert wird.

Auch der Dachverband Lehrerinnen und Lehrer Schweiz (LCH) umschreibt in einem Positionspapier (Dachverband Schweizer Lehrerinnen und Lehrer (LCH) 2005b) zu Schulleitungen deren Kernaufgabe wie folgt:

„Schulen brauchen eine verfasste Leitung, die nach anerkannten Regeln funktioniert, einen Auftrag und geklärte Zuständigkeiten hat und deshalb in ihrem Tun verantwortlich gemacht werden kann, verhandelbar und Veränderungen zugänglich ist“ (S. 1).

Das Papier (ebd.) nennt sodann fünf Kriterien für gutes Schulleitungshandeln. Gute Schulleitungen:

- (1) sind dem Bildungsauftrag und einem guten Arbeitsklima verpflichtet,
- (2) betrachten eine organische und verlässliche Schulentwicklung als Kernziel,
- (3) verstehen sich als Hüter von Vorschriften und als Vermittler zwischen Interessen,
- (4) nehmen in geklärten Zuständigkeiten balanciert Macht wahr,
- (5) wirken kompetent und entlastend dank ausreichender Ressourcen.

Schulleitungen sind dafür verantwortlich, dass Schulen (wieder) vermehrt als Gemeinschaft agieren und als pädagogische Handlungseinheiten Verantwortung auch für die Ergebnisqualität dieses Handelns übernehmen. Schulleitungen sind nicht selbst für die Ergebnisse bei Schulen, wohl aber für deren Professionalität verantwortlich, z.B. dafür, dass an Schulen an der Sicherung und Entwicklung der eigenen Qualität gearbeitet, in einem guten Arbeitsklima miteinander kooperiert und die fachliche Arbeit an geltenden Standards ausgerichtet wird.

Welche Wirkungen erzeugen Schulleitungen? Fend (1998) nennt im Rahmen eines Extremvergleichs Merkmale wie Verantwortungsbereitschaft, aktive Aufgabenbearbeitung, kooperative Arbeitsformen und eine pädagogische Grundhaltung als charakteristische Merkmale von Leitungen guter Schulen. Buholzer (2000) erhob das Innovationsklima im Teilautonomie-Projekt „Schulen mit Profil“. Eine hohe Innovationsbereitschaft war in hoch signifikanter Weise mit einer positiven Einschätzung der Schulleitung durch die Lehrpersonen gekoppelt – an tendenziell innovativen Schulen schätzten die Lehrerinnen und Lehrer die Arbeit der Leitung sowohl in administrativer, in unterstützender als

auch in sozialer Hinsicht als hoch ein. Schulleitungen nehmen in Schulen mit hoher Innovationsbereitschaft in starkem Maße Einfluss auf das Geschehen (ebd.).

Bonsen/von der Gathen/Iglhaut/Pfeiffer (2002) und Bonsen/von der Gathen/Pfeiffer (2002) identifizierten aufgrund verschiedener, über Lehrer-, Eltern- und Schülerdaten erhobener Schulqualitätsaspekte (z.B. Differenzieren und Fördern im Unterricht, pädagogische Innovationsbereitschaft im Kollegium, kollektive Selbstwirksamkeitserwartung, Lehrer-Schüler-Verhältnis) gute und verbesserungsbedürftige Schulen. Diese Ergebnisse wurden mit Skalen im Lehrerfragebogen verglichen, mit welchen das Schulleitungshandeln erfasst wurde. Dabei zeigten sich nicht nur höhere Mittelwerte der guten Schulen hinsichtlich sämtlicher verwendeter Skalen, sondern es zeigte sich auch eine gute Prädiktorqualität der Merkmale „Zielbezogene Führung“ und „Innovationsbereitschaft“ in Bezug auf die Vorhersage der Schulzugehörigkeit (ebd.). Auch die Faktoren „Partizipation in der Entscheidungsfindung“ und „Organisationskompetenz“ haben eine gewisse Erklärungsmacht hinsichtlich Schulqualität. In einem weiteren Schritt wurde der Einfluss von Schulleitungshandeln auf solche Variablen geprüft, welche als Bedingungen für guten Unterricht gewertet werden können. Im Strukturgleichungsmodell konnten sowohl direkte wie auch über das Kollegium vermittelte indirekte Effekte von Zielgerichtetheit auf Differenzieren und Fördern im Unterricht festgestellt werden. In einer angelsächsischen Studie haben Hallinger/Bickman/Davis (zit. aus Bonsen/von der Gathen/Iglhaut/Pfeiffer 2002) bereits 1996 die Wirkung von Schulleitungshandeln auf die Schülerleistungen untersucht. Dabei zeigt sich ein indirekter Zusammenhang zwischen dem unterrichtsbezogenen Führungshandeln der Leitung und den Leseleistungen der Lernenden. Die Leistungserwartung der Lehrpersonen als unmittelbarster Einfluss auf die Leistung wird maßgeblich durch die beiden Variablen „geeignete Lernanlässe“ und „klares Schulleitbild“ beeinflusst. Diese wiederum werden maßgeblich vom unterrichtsbezogenen Führungshandeln beeinflusst (ebd.).

Allerdings muss vor einer Rhetorik stabiler, allgemein erfolgversprechender Schulleitungsmerkmale gewarnt werden. Nach einer Forschungsübersicht über Ansätze des Leitungshandelns kommen Vroom/Jago (2007) zum Schluss, dass insofern von einem „leadership paradox“ gesprochen werden kann, als Stabilität in Bezug auf Führungshandeln immer situationsbezogen ist. Schulleitungshandeln lässt

sich nicht durch einfache und allgemeine Regeln beschreiben, sondern ist an bestimmte Situationstypen gebunden. Die empirische Forschung zu dieser Thematik ist jedoch noch wenig fortgeschritten; sie muss versuchen, Schlüsselsituationen zu identifizieren, die eine Differenzierung erfolgreichen und weniger erfolgreichen Schulleitungsverhaltens erlauben.

Angesichts der Befundlage kann davon ausgegangen werden, dass Schulleiterinnen und Schulleiter durchaus einen bedeutsamen Einfluss einerseits als Türöffner und Katalysatoren von Innovationen (Fullan 1991), andererseits auf den kooperativen Umgang sowie die Weiterentwicklung und Innovationsbereitschaft eines Kollegiums haben. In einigen Untersuchungen haben sie sich sogar als stärkster Prädiktor von Veränderungen erwiesen (Berends/Chune/Schuyler/Stockly/Briggs 2002; vgl. Desimone 2002, S. 449). Gerade diese Faktoren dürften bei der Implementation von Bildungsstandards eine begünstigende Wirkung haben, gilt es doch, die Lehrpersonen für diese neuen Werkzeuge zu gewinnen und eine interne Kultur zu schaffen, welche es den Betroffenen nicht nur erlaubt, ihren Unterricht gemeinsam zu reflektieren, sondern auch weiterzuentwickeln. Dies bedeutet, dass Schulleitungspersonen nicht nur über organisatorische und administrative Fähigkeiten und eine ausgewiesene Sozial- und Beratungskompetenz verfügen, sondern darüber hinaus pädagogisch innovativ sind und die Weiterbildungsbereitschaft ihres Kollegiums aktiv fördern.

Bisherige Befunde – vor allem im Rahmen von PISA 2003 – zeichnen diesbezüglich allerdings ein eher düsteres Bild. Dass Schulleitungen im Hinblick auf schulische Innovation zum Beispiel die Bildung fachdidaktischer Arbeitsgruppen oder die Thematisierung von Inhalten anstoßen, gilt lediglich in sehr beschränktem Maße (Senkbeil et al. 2004). Ebenso nachdenklich muss der Befund stimmen, dass lediglich ein Viertel der befragten deutschen Schulleiterinnen und Schulleiter mindestens einmal jährlich Unterrichtsbesuche vornimmt (ebd.). Schulleitungen müssten im Hinblick auf die Einführung von Standards Koordinationsleistungen erbringen, um die Fachkonferenzen ihrer Schulen auf eine adäquate und gehaltvolle Thematisierung standardorientierten Unterrichtens zu verpflichten. Köller/Köller (2006) verweisen in diesem Zusammenhang auf folgende Punkte, welche für die Neugestaltung der Schulleitungsrolle bedeutsam sein dürften:

- eigene Professionalisierung im Hinblick auf die Philosophie und die Inhalte der Standards,

- Auseinandersetzung mit Konzepten der dezentralen Lehrerfortbildung, wie sie etwa im Rahmen von SINUS bzw. SINUS-Transfer realisiert werden,
- Forcierung der Auseinandersetzungen mit Standards in den Fachschaften,
- Aufforderungen an die Lehrpersonen und Eltern, sich kompetent über Standards zu informieren sowie
- Beteiligung an der Planung der Umsetzung von Konzepten der Lehrprofessionalisierung mit entsprechenden Fachgruppen und politischer Administration.

Im Sinne transformationaler Führung muss aber auch die Erwartungshaltung relativiert werden, wonach einer guten Schulleitung die alleinige und entscheidende Scharnierfunktion zwischen dem Bildungssystem und den Lehrkräften zukomme – „gute Schulen ohne gute Schulleiter gibt es nicht; dennoch sind gute Schulleiter keine Garantie für gute Schulen“ (Rolff 1993, S. 183). Ein Forschungsdesiderat sind hier Längsschnittuntersuchungen, welche die Wirkungen unterschiedlicher Führungsstile auf Aspekte der schulischen Mesoebene und die Mikroebene des Unterrichts zum Thema haben. Ebenso muss die Einführung von Schulleitungen immer auch im Kontext umfassender Schulentwicklungsarbeit gesehen werden. Beispielhaft sei hier auf das Projekt GELVOS (geleitete Volksschulen im Schweizer Kanton Schwyz) verwiesen, das die Einführung von Schulleitungen an ein von den Schulen zu schaffendes Qualitätskonzept bindet. Darin müssen Input-, Prozess- und Outputqualitäten festgelegt und ihr Vorhandensein über interne und externe Evaluationen überprüft werden (Erziehungsdepartement Kanton Schwyz 2004). Schulleitungen sind im Hinblick auf die Einführung von Standards auf klare Rahmenbedingungen angewiesen, welche es ihnen ermöglichen, tatsächlich pädagogisch zu führen – im Sinne einer Sensibilisierung und Fokussierung ihres Personals auf innovativen und effektiven Unterricht.

Angesichts der Neuheit der Aufgabe ist schließlich auch ein Nachdenken notwendig über die Rekrutierung sowie die Ausbildung von Schulleitungen. Professionelles Handeln von Schulleiterinnen und Schulleitern bezieht sich auf ein Aufgabengebiet, das – wie weiter oben beschrieben – neben pädagogischen Kompetenzen solche der Organisation, der Personalführung und der Finanzverwaltung umfasst. Dies verlangt neben einer umfassenden theoretischen Ausbildung, welche

sowohl fachliche, als auch Selbst- und Sozialkompetenzen einschließt, auch eine Begleitung durch Fachpersonen in der Phase der Berufseinführung. Fragen, die sich dabei stellen, sind, ob es gerechtfertigt und wirkungsvoll ist, wenn Schulleiter und Schulleiterinnen ausschließlich aus der Lehrerschaft kommen. Weiter darf gefragt werden, ob Schulleitungspositionen nicht auch für außerschulische Quereinsteiger und Führungskräfte, insbesondere aus schulnahen Tätigkeitsbereichen, im Prinzip offen sein und in transparenten Ausschreibungsverfahren besetzt werden könnten. Fragen beziehen sich sodann auf die Definition der Wahlgremien, die Begrenzung von Amtszeiten und auf Wiederwahlzyklen. Schließlich stellt sich die wichtige Frage der Ausbildung. Im Sinne einer dezidierten Professionalisierungsstrategie ist anzustreben, dass die Ausbildung zum Schulleiter bzw. zur Schulleiterin neben der Qualifizierung für administrative Führungs- und fachpädagogische Entwicklungsaufgaben ebenfalls außerschulische Erfahrungen in Unternehmen und Organisationen einschließen sollte. Letzteres idealerweise im Sinne eigens dafür eingerichteter Fort- und Weiterbildung und als Qualitätskriterium für die Auswahl von Schulleitungspersonal.

5.2.2 Schulprogrammarbeit

Im Rahmen einer unter Bedingungen von Teilautonomie angestrebten Einführung von Bildungsstandards wird der Schulleitbild- oder Schulprogrammarbeit ebenfalls eine Bedeutung zuerkannt (vgl. zu Schulprogrammen als Instrumenten der Schulentwicklung Holtappels 2004). Verschiedene deutsche Bundesländer und Schweizer Kantone haben in den vergangenen Jahren Projekte auf den Weg gebracht oder gar gesetzliche Grundlagen geschaffen, welche die Entwicklung von Schulleitbildern vorsehen. So wurde beispielsweise im Jahr 2000 in Berlin das Pilotprojekt „Schulprogrammentwicklung und Evaluation“ initiiert, bei dem ca. 60 Schulen über einen Zeitraum von drei Jahren durch ein Beratersteam des Landesinstituts für Schule und Medien (LISUM) bei der Entstehung und Umsetzung eines Leitbildes unterstützt und begleitet werden.¹⁹³ In der Schweiz wurde im Kanton Luzern Mitte der neunziger Jahre ebenfalls das Projekt „Schulen mit Profil“ lanciert, welches neben Schulleitungen und institutioneller Elternmitwirkung ein von den Lehr-

¹⁹³ Berliner Bildungsserver (1998-2004): Forum Schulprogramm.
<http://bebis.cidsnet.de/fortbildung/unterstuetzung/forum/> (Stand 24.03.06).

kräften entwickeltes und von der Schulbehörde bewilligtes Leitbild vorsieht (Büeler/Buholzer/Kummer/Roos 2001). Dieses Projekt wurde bezüglich verschiedener Wirkungen wissenschaftlich begleitet und evaluiert (ebd.). Im Kanton Zürich werden die Schulen im Rahmen des neuen Volksschulgesetzes auf ein Schulprogramm verpflichtet, das von der Schulkonferenz, also dem Kollegium und der Schulleitung in gemeinsamer Arbeit entwickelt wird (Kanton Zürich 2005).

Wodurch lassen sich Schulprogramme charakterisieren? Schratz/Bauk-van Vugt (2000) definierten Qualitätsstandards für Schulprogramme und nennen folgende Kriterien, welche – unabhängig von den je schulspezifischen Besonderheiten – Basis sein müssen, um ein Leitbild hinsichtlich seiner Güte zu bewerten:

- Eine *klare gemeinsame Zielstellung*, welche die wichtigsten Entwicklungsziele der Schule schriftlich festlegt und sich auf ausgewählte Kernaufgaben innerhalb der wichtigsten Qualitätsbereiche (z.B. Lehren und Lernen) bezieht;
- *wirksame Maßnahmen* zur Umsetzung des Schulprogramms, welche pragmatisch, verbindlich und rechenschaftspflichtig sind;
- *Transparenz nach innen und außen*, indem die betroffenen Akteure (Lehrpersonen, Eltern, Schülerinnen und Schüler) Kenntnis über die Inhalte des Programms haben;
- ein *stimmiges internes Steuerungssystem der Entwicklungsprozesse* über eine eigens dafür eingesetzte Steuergruppe;
- eine *wirksame Qualitätssicherung*, welche die zu überprüfenden Ziele und Inhalte anhand klar definierter Items zu erfassen vermag und interne wie externe Evaluation vorsieht, sowie
- die *Verfügbarkeit und Einbindung von Ressourcen* personeller, materieller, räumlicher, finanzieller und zeitlicher Natur (ebd.).

Wendt (2002) orientiert sich bei seinem Orientierungsleitfaden zur Schulprogrammentwicklung an der Schulqualitäts- und Schulentwicklungsliteratur. Die Eckdaten und Strukturen einer Schule sind demnach ebenso zu berücksichtigen wie der bisherige Verlauf der Schulprogrammentwicklung, vorhandene Konflikte und die Planung und Durchführung interner als auch externer Evaluationen. Im Schlussbericht des Projektes „Schulprogrammentwicklung und Evaluation“ wird zudem auf die Notwendigkeit von gründlicher interner Evaluation nach Ein-

führung des Programms hingewiesen. Mit quantitativen und qualitativen Verfahren soll die Implementation des Leitbildes begleitet und kontrolliert werden. Eine eingehende Analyse und Interpretation der dabei gewonnenen Daten soll Grundlage allfälliger Korrekturen und Verbesserungen bilden. Die Autoren nennen auch die dabei auftretenden Probleme und Risiken: Widerstand im Kollegium, da eine gründliche Evaluation eines Schulleitbildes auch Unterrichtsbeobachtungen beinhaltet, und der Unmut darüber, dass dieses Werkzeug kaum zu Arbeitserleichterungen führt. Als hilfreiche Unterstützungsangebote haben sich praktische Methodenkoffer, die Beispiele anderer Pilotschulen, Einstiegs- und Prozessbegleitungen durch externe Beratung und von der Schulaufsicht durchgeführte Fachtagungen erwiesen (ebd.).

Auch eine umfangreiche Evaluation der Leitbildarbeit in Nordrhein-Westfalen brachte zum Ausdruck, dass mehr als die Hälfte der befragten Lehrpersonen die Schulprogrammentwicklung als große Belastung erlebten. In 75% der Schulen wurde eine Steuergruppe dafür ausgebildet und eingesetzt. Der Unterricht wird in Schulprogrammen bei drei Vierteln der 180 Schulen ganz oder teilweise als Kernpunkt der Programmentwicklung genannt, allerdings verneinen die befragten Lehrpersonen weitgehend eine Auswirkung des Schulprogramms auf den persönlichen Unterricht (ebd.). Weniger als 10% der Befragten stimmten dem Item „Das Nachdenken über gemeinsame Ziele hat den Unterricht der einzelnen Lehrer verändert“ zu (ebd., S. 18). Demgegenüber werden Schulleben, Erziehungsarbeit und die Öffnung von Schule stärker gewichtet (ebd.). Als Gelingensbedingungen der Schulprogrammarbeit nennen die Lehrkräfte vor allem das Vorhandensein pädagogischer Konzepte und Modelle, Kooperation und Teamarbeit sowie die Arbeit einer Schulprogrammgruppe. Der Nutzen des gesamten Prozesses wird zurückhaltend beurteilt. Nur ein knappes Fünftel erkennt in der Schulprogrammarbeit einen Ertrag für die eigene Schule, während über die Hälfte nur eine mittlere und gegen 30% eine geringe Wirkung feststellen. Im Detail wird vor allem der durch das Schulprogramm in Gang gesetzte Diskussionsprozess über gemeinsame Ziele in der Schule als positiver Nutzen erwähnt. Allerdings wird dieser Nutzen hinsichtlich der eigentlichen pädagogischen Arbeit als gering beurteilt und immerhin 28% der befragten Lehrpersonen halten das Programm gar für überflüssig. Als kritisch erwies sich ebenfalls, dass die Kollegien in weniger als 10% der Fälle ihre Arbeit selbst evaluierten (ebd.).

In der erwähnten Untersuchung in Luzern zu „Schulen mit Profil“ konnte festgestellt werden, dass die Eltern das Leitbild der Schule ihres

Kindes recht gut zu kennen scheinen, dass aber die Identifikation mit diesem Programm vor allen Dingen bei denjenigen Eltern am höchsten ist, welche eine allgemein gute Zusammenarbeit mit der Schule vermelden (Büeler/Buholzer/Kummer/Roos 2001).

Insgesamt werden auf sämtlichen Stufen und in den meisten deutschen Bundesländern und Schweizer Kantonen breit angelegte Bemühungen sichtbar, Einzelschulen programmatisch zu gestalten, um damit eine Vereinheitlichung von Zielen und deren Überprüfung zu gewährleisten, aber auch um dem Anspruch nach zunehmender Öffnung respektive Öffentlichkeitsarbeit gerecht zu werden. Schulprogramme bzw. -leitbilder sind häufig mit weiteren Schulentwicklungsmaßnahmen verbunden, wie etwa der Teilautonomie und der Implementation von Schulleitungen. Allerdings sind nur wenige empirische Untersuchungen zu Wirkungen verfügbar, zumal hierzu anspruchsvolle methodische Designs notwendig wären (Längsschnitte, Mehrebenen-Modellierungen). Auf das „Kerngeschäft“ der Unterrichtsarbeit scheinen die Wirkungen zudem eher gering zu sein. Dies kann – in Verbindung mit der von den Lehrpersonen wahrgenommenen Belastung – zu einer Gefährdung für die Leitbildarbeit werden, vor allem dann, wenn über einen längeren Zeitraum hinweg keine klar erkenntlichen positiven Effekte auszumachen sind und gleichzeitig Druck durch ähnlich angelegte Anliegen schulischer Reformen besteht.

5.2.3 Integrierte Selbst- und Fremdevaluation

Evaluation bedeutet erstens Informationen sammeln, die Auskunft geben, wie gut und zielbezogen ein System arbeitet; zweitens feststellen, welche Elemente nicht befriedigen; und drittens versuchen, die Situation zu verbessern. Nur eine Schule, welche sich ihrer Standards vergewissert und die daran interessiert ist, deren tatsächliche Erreichung periodisch zu überprüfen – d.h. eine Schule, für die Evaluation und Qualitätssicherung selbstverständliche Arbeitsbereiche des Kollegiums sind –, ist eine professionelle Schule. Professionen zeichnen sich gerade dadurch aus, dass sie sich ihrer Standards bewusst sind, sie explizit verfügbar halten und für ihre Anwendung sorgen. Während die Schulleiterinnen und Schulleiter im Rahmen ihrer strategischen Gesamtverantwortung für die Entwicklung einer Evaluationskultur verantwortlich sind, besteht der Auftrag des Kollegiums darin, sich Formen der inter-

nen und externen Evaluation zu stellen und den damit verbundenen Auftrag der pädagogischen Qualitätsentwicklung wahrzunehmen.

Auf der strukturellen Ebene des Bildungssystems besteht die Aufgabe in der Entwicklung eines integrierten und mehrdimensionalen, aus Instrumenten der Selbst- und Fremdevaluation bestehenden Systems der Qualitätssicherung und Qualitätsentwicklung. Bisher existieren solche Systeme, in denen die verschiedenen Ziele und Funktionen im Rahmen eines Gesamtkonzeptes aufeinander abgestimmt sind, in den Auftragsländern nicht oder nur in Umrissen. Eine zentrale Schwierigkeit besteht dabei darin, die verschiedenen Zwecke einer Evaluation – Kontrolle, Entwicklung, Forschung –, die immer auch unterschiedliche, im Feld vorhandene Kräfte und Ansprüche repräsentieren, in einem integrierten Gesamtkonzept zu vereinigen.

In ihrem Buch „Educational Evaluation, Assessment and Monitoring“ schlagen Jaap Scheerens, Cees Glas und Sally M. Thomas (2003) auf der Basis einer international ausgerichteten, zwei Jahrzehnte abdeckenden Analyse von Forschungsergebnissen zur Qualitätssicherung einen „systemischen Ansatz“ vor, welcher der Diskussion in den Auftragsländern wichtige Impulse zu geben vermag.¹⁹⁴ Evaluationsverfahren in einem Bildungssystem zu etablieren, bedeutet nach Scheerens/Glas/Thomas (2003, S. 15),

- dass die Verfahren nicht singulär, etwa für die Bewertung spezifischer Maßnahmen und Programme, eingesetzt werden, sondern als Bestandteil des Steuerungssystems institutionell eingeführt und laufend genutzt werden,
- dass Evaluation Entwicklungs- und Lernprozesse in Gang setzt und am Leben erhält, die dem Prinzip des feedbackgesteuerten Lernens genügen,
- dass verschiedene Datenquellen bzw. Arten der empirischen Erhebung (Testerhebungen im Rahmen (inter-)nationaler Leistungsstudien oder zentraler Prüfungen, amtliche administrative und statistische Daten, Daten aus Selbstevaluationen, Schulinspektion, Schulaufsicht oder dienstrechtlichen Bewertungen), die je spezifischen

¹⁹⁴ Wir orientieren uns in den folgenden Abschnitten was die deutsche Übersetzung und die Darstellung der Kernaussagen des Ansatzes von Scheerens/Glas/Thomas (2003) anlangt an Klieme (2005, S. 46ff.).

Zwecken dienen, gemeinsam betrachtet und damit Synergien genutzt werden und

- dass schließlich zur Systematisierung und Interpretation von Befunden ein theoretisches Rahmenmodell verwendet wird, welches Input-, Prozess- und Outputelemente miteinander in Beziehung setzt.

Auf der Basis dieser Verfahren und Daten erfüllt ein integriertes Evaluations- und Monitoringsystem den Autoren zufolge drei Funktionen: Akkreditierung und Zertifizierung (eine Funktion, die in den Niederlanden, wo es einen bedeutenden privaten Schulsektor gibt, eine wichtige Rolle spielt), Rechenschaftslegung (*Accountability*) sowie Diagnose und Organisationslernen.

Klieme (2005a) deutet das von Scheerens/Glas/Thomas (2003) entworfene Konzept einer systemischen und integrativen Sicht von Qualitätssicherung und Qualitätsentwicklung wie folgt:

„Die Autoren entwerfen das Bild eines empiriegestützten Schulentwicklungsprozesses, der durch eine Verbindung von externer und interner Evaluation vorangebracht wird. Evaluation sehen sie (Scheerens/Glas/Thomas 2003, S. 94) als das Herzstück einer lernenden Organisation, d.h. einer Schule, die ihre interne Effektivität und ihre Verantwortlichkeit gegenüber externen Ansprüchen verbessern will. Eine solche feedbackgestützte Strategie der Schulentwicklung sei vermutlich effektiver als Innovationsstrategien, die ‚vorwärts planend‘ arbeiten. Entscheidend für das Modell ist, dass Ergebnis- und Prozessqualität gleichermaßen zum Gegenstand des Feedbacks gemacht werden. Insbesondere sollten möglichst alle wichtigen Prozessfaktoren abgedeckt werden, die – gemäß den Ergebnissen der Schulwirkungsforschung (Scheerens/Bosker 1997) – für erfolgreiche Schulen typisch sind, also vor allem:

- Leistungsorientierung,
- Kooperation, Konsens und Kohärenz im Kollegium,
- pädagogische Führung,
- curriculare Qualität/Lerngelegenheiten,
- Schulklima und Schulkultur,
- Evaluation,
- Einbeziehung der Eltern“ (S. 46f).

Mit dieser Breite von Evaluationskriterien geht das von Scheerens und Mitautoren vorgestellte Konzept deutlich über die sogenannten *accountability systems* hinaus, wie sie in den USA entwickelt worden sind (vgl. Fuhrman/Elmore 2004). Dort dominiert immer noch eine rein leistungsbezogene Bewertung von Schulen, zumeist gestützt auf standardisierte Tests, die regelmäßig – im Extremfall jährlich für jede Schülerin und jeden Schüler – eingesetzt werden. Die Erfahrungen der USA belegen, dass ein solches engmaschiges, ausschließlich testbasiertes Monitoring-System zwar Effekte im Sinne von über die Jahre anwachsenden Leistungsniveaus zeigen kann – Effekte jedoch, deren Nachhaltigkeit nicht belegt ist, die möglicherweise auf elementare Leistungsbereiche begrenzt sind, welche schnell trainiert und leicht abgetestet werden können, und die vermutlich gerade jene Schulen, die besondere Probleme haben, gar nicht erreichen (vgl. Klieme 2004a). Gleichzeitig haben solche extensiven Testsysteme nachweislich negative Folgen, beispielsweise durch ein „teaching to the test“, das die Breite und die Tiefe des Curriculums beeinträchtigt. Experten der Schulentwicklung und Evaluation in den USA empfehlen daher auch dort den „Rückbau“ der Accountability-Systeme in Richtung einer vorsichtigeren, unterrichtsnahen, multiple Kriterien einbeziehenden Schulevaluation.

Das entscheidende Argument für die Berücksichtigung von Ergebnis- und Prozess-Aspekten ergibt sich aus der Frage, wie denn Qualitätsentwicklung bzw. – in der Terminologie von Scheerens et al. – die Funktion des „Organisationslernens“ realisiert werden soll. Eine bloße Feststellung von Lernergebnissen mag für die Rechenschaftslegung nach außen oder (in einem noch stärker deregulierten System) für eine Akkreditierung von Institutionen genügen. Eine Schule als „Lernende Organisation“ muss aber auch wissen, an welchen Prozessfaktoren sie ansetzen kann, um ihre Ergebnisse zu sichern bzw. zu verbessern. Evaluation ohne Prozessanalyse bedient bestenfalls legitimatorische Zwecke und Kontrollwünsche, ermöglicht jedoch keine Weiterentwicklung (zu dieser Unterscheidung von Evaluationszielen vgl. Abs/Klieme 2005, S. 46f).

Die vorliegende Expertise schließt sich diesem Modell der Qualitätssicherung und Qualitätsentwicklung an und betont,

- dass es sowohl *objektiver Indikatoren* (z.B. anhand schulübergreifender Standards gewonnene Testwerte) als auch der *professionellen Verpflichtung zur Selbstevaluation* entstammender subjektiver Urteile und Wahrnehmungen aller Beteiligten bedarf und

- dass es sowohl summative (die Erfüllung definierter Sollwerte überprüfende) als auch formative (die weitere Entwicklung stimulierende) Daten

braucht, wenn Evaluation mehr sein soll als reine bürokratische Kontrolle oder aber bloße „pädagogische Goodwillpolitik“ (Fend 2002). Beide haben sich entsprechend dem, was die Forschung darüber weiß, als wenig erfolgreich erwiesen.

Weiter bedeutet die hier vertretene Position, dass im Rahmen einer „neuen Bildungssteuerung“ nicht allein auf eine Evaluation des Outputs gesetzt werden darf, sondern zentrale Input- und Prozessmerkmale, von denen bekannt ist, dass erfolgreiche Schulen sie aufweisen, mitevaluiert werden müssen.

Während die Überprüfung von Bildungsstandards im engen Sinne durch testbezogene Outputmessung zwar möglich und auch sinnvoll ist, entscheidet letztlich die Qualität des in Prozesse umgesetzten schulischen Inputs (Bildungsstandards im weiten Sinne) darüber, wie weit diese auch erreicht werden. Dazu kommt, dass nur Schulen, die über ein breites Verständnis von Standards verfügen, das sowohl Ergebnis- als auch Prozessstandards umfasst, und sich in ebendieser Breite für deren Sicherung und Weiterentwicklung einsetzen, überhaupt in der Lage sein werden, Leistungsrückmeldungen zu verarbeiten und in verbesserten Input zu übersetzen, d.h. *zu lernen*. Verfahren der Evaluation von Bildungsstandards müssen sich deshalb sowohl auf Bildungsstandards im *engen* als auch im *weiten* Sinne richten: auf durch Tests messbare, *fachbezogene Outputstandards* und auf *schulpädagogische und (fach)didaktische Input- und Prozessstandards*, welche breiteren Evaluations- und Qualitätskriterien (auch im Sinne von Scheerens/Glas/Thomas 2003) entsprechen und über ein enges Verständnis von „accountability“ hinausgehen.

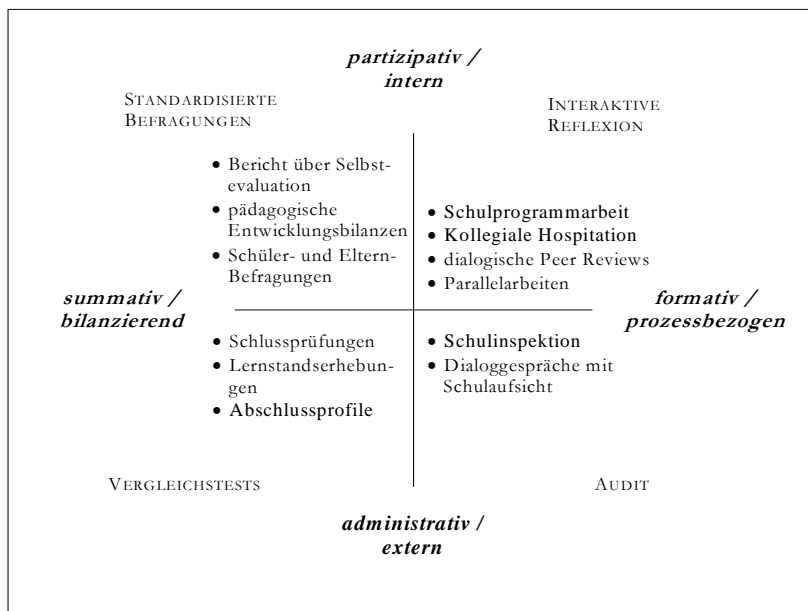
Was bedeutet dies in Bezug auf konkrete Verfahren der Fremd- und der Selbstevaluation?

Beide Arten sind notwendig und in Ergänzung zueinander zu verstehen. Auch wenn Bildungsstandards Orientierung und Zielklarheit liefern, genügt Selbstreferentialität allein nicht bei der Qualitätssicherung. Wirkungsbewusstsein kann nicht lediglich als professionelle Selbstverpflichtung verstanden werden. Es braucht kluge Mischungsverhältnisse von Selbst- und Fremdevaluation und dazugehörige bewährte, auch objektive, Instrumente. Aus der Schulentwicklungsforschung der letzten Jahre liegt diesbezüglich einiges an Know-how vor. Dieses Wissen muss

an die von den Bildungsstandards geforderte Qualitätsdiskussion angeschlossen werden. Evaluation wird sodann umso eher akzeptiert, je erkennbarer sie für die Evaluierten mit positiven Anreizen gekoppelt ist, nicht bloß nach Mängeln sucht, sondern auch Stärken betont und Erfolge sichtbar macht, das Gefühl eigener Wirksamkeit vermittelt, praktisch umsetzbare Ergebnisse liefert und mit Förderoptionen verbunden ist.

Unter Bezugnahme auf eine von Nisbet (1990) stammende zweidimensionale „Landkarte der Interessen von Evaluation“ sowie im Anschluss an Klieme (2005a) lassen sich die im Schulbereich zurzeit diskutierten Evaluationsverfahren und -instrumente in einer Matrix mit folgenden Dimensionen darstellen (Abb. 9): Bilanzierend-summative Kontrolle versus entwicklungsorientiert-prozessbegleitende Diagnose; administrativ gesteuerte, externe Rechenschaftslegung versus Einlösen professioneller Selbstverpflichtung mittels partizipativer Verfahren der Selbstevaluation.

Abbildung 9: Dimensionen und Typen von Schulevaluation



Quelle: In Anlehnung an Klieme (2005a) sowie Nisbet (1990).

Dass Schulen nicht nur extern evaluiert werden, sondern im Sinne professioneller Selbstverpflichtung Selbstevaluation betreiben sollen, findet seinen Niederschlag zunehmend auch in einer veränderten Gesetzgebung. Bisher waren Verfahren der externen und internen Evaluation gesetzlich kaum geregelt. Soweit der Begriff der Evaluation in Schulgesetzen überhaupt vorkam, beschränkte man sich auf allgemeine Hinweise, etwa darauf, dass es sich um eine wichtige, von Schulen und Schulaufsicht gemeinsam wahrzunehmende Aufgabe handle. Offensichtlich bewegt man sich in diesem Bereich vielerorts immer noch auf Neuland. Dennoch kommen die Dinge zunehmend in Fluss. So sind etwa die öffentlichen Schulen des Kantons Zürich im Rahmen des neuen Volksschulgesetzes zur Qualitätsentwicklung und -sicherung verpflichtet (Kanton Zürich 2005). Damit werden Rahmenbedingungen geschaffen, welche Weiterentwicklung verlangen und dadurch die Ermächtigung zur Selbstevaluation erfordern. Selbstevaluationsinstrumente gewinnen auf diese Weise nicht nur in Bezug auf die Steuerung von Einzelschulen, sondern auch hinsichtlich der Steuerung des Gesamtsystems Bedeutung. Entsprechend sind Instrumente notwendig, welche den Schulen und deren Lehrpersonen Rückmeldungen ermöglichen, welche die Grundlage für praktikable Maßnahmen und Ziele bilden.

SEIS. Ein Beispiel dafür lässt sich wiederum dem Programm der Bertelsmann-Stiftung entnehmen. Das Instrument bezieht sich auf die Gestaltung von „Selbstevaluationen in Schulen“ (*SEIS*) und ist als Unterstützung der internen Schulsteuerung zu verstehen. Zudem können damit externe Evaluationen vorbereitet werden. Konkret handelt es sich um einen Fragebogen, der im Blick auf fünf Dimensionen der Schulqualität eingesetzt werden kann. Die Fragebögen können auch online ausgefüllt werden. Die Datenauswertung erfolgt elektronisch, sie lässt auch Vergleiche mit anderen Schulen zu.¹⁹⁵ *SEIS* geht von einem Qualitätsverständnis aus, welches Lehr- und Lernprozesse als prioritäres Anliegen von Schule sieht und der Schulführung, dem Schulklima und der Schulkultur bei der Realisierung von lernförderlichen Rahmenbedingungen entscheidende Bedeutung beimisst. Vor der Erhebung soll über deren Durchführung unter den Betroffenen eine breit angelegte Diskussion stattfinden. Aufgrund der ausgefüllten und ausgewerteten Fragebö-

¹⁹⁵ *SEIS* macht Schule: <http://www.das-macht-schule.de/service/presse/PM4> (Stand 24.03.06).

gen werden in einem weiteren Schritt die Daten mit den Beteiligten sorgfältig interpretiert, Aufgaben definiert, Handlungsoptionen betrachtet und deren Umsetzung und Effektivität geprüft. Ebenso wird davon ausgegangen, dass Veränderungsprozesse langfristig gesehen werden müssen und der Einbindung sämtlicher Angehöriger eines Kollegiums bedürfen.

EVIT. Ähnliches gilt zum seit 2004 in Schleswig-Holstein eingeführten Verfahren *EVIT* (Evaluation im Team). Ausgehend vom Grundsatz, dass eine Schule nur für jenen Gestaltungsspielraum, den sie verantwortet, Rückmeldungen bekommen soll, werden Aspekte von Lern- und Arbeitsbedingungen erhoben (Ministerium für Bildung und Frauen des Landes Schleswig-Holstein 2005). Die teilnehmenden Schulen sollen über Stärken und Schwächen in ihrer Entwicklung und deren Wirkungen informiert werden sowie Impulse für ihre Arbeit erhalten. Das Konzept umfasst sechs Qualitätsbereiche: Bildungsprozesse, Schulische Effekte, Lern- und Arbeitsbedingungen, Schulleitung, Qualitätsmanagement und Kooperation. Neu ist hier die Anbindung an die KMK-Bildungsstandards, indem unter „Schulische Effekte“ die von den Schülerinnen und Schülern der jeweiligen Institution erbrachten Leistungen in den Hauptfächern am Landesdurchschnitt gemessen werden. Als Messinstrumente werden Fragebögen für Lehrer, Schülerinnen und Schüler und Eltern, die erwähnten Ergebnisse der Schülerinnen und Schüler, das Schulprogramm, aktuelle Daten aus der Schulstatistik sowie Fachkonferenzprotokolle eingesetzt. Ergänzend sind systematische Unterrichtsbeobachtungen, Gebäuderundgänge und Evaluationsgespräche mit sämtlichen bereits genannten Akteuren vorgesehen. Erhoben werden die Daten durch Expertenteams des Instituts zur Qualitätsentwicklung an Schulen Schleswig-Holsteins. Vorgesehen ist, dass die Schulen in einem Vier-Jahres-Turnus fortlaufend evaluiert werden.

Die beiden Beispiele *SEIS* und *EVIT* verweisen auf ein grundsätzliches Dilemma bei der Evaluation an der Einzelschule: Zum einen erfordern knappe Zeitbudgets und höhere Belastungen der Lehrpersonen leicht handhabbare Instrumente, welche mit vergleichsweise geringem Zeitaufwand bearbeitet werden können. Andererseits bedingt die Komplexität des Systems Einzelschule Zugänge auf verschiedenen Ebenen und zeitintensive Beobachtungen aus verschiedenen Perspektiven.

Grundsätzlich ist es nicht leicht, Fremdevaluationen und die daran anschließenden Rückmeldungen so zu gestalten, dass Schulen daraus lernen können. Lehrpersonen müssen die (aufwändigen) Evaluationen

mit einem Mehrwert verbinden bzw. als nutzbringend erfahren. Akzeptanz und erfahrene Nutzeffekte sind dabei umso größer,

- je besser die wahrgenommene methodische Qualität der eingesetzten Verfahren ist,
- je konkreter und handlungsbezogener das Feedback ist,
- je mehr sich die Rückmeldung neben Partikulärem auf gemeinsam akzeptierte Ziele und Werte der Schule bezieht,
- je diagnostischer die Rückmeldung in Bezug auf die Benennung von möglichen Ursachen von Mängeln bzw. Indikatoren des Erfolgs gestaltet ist sowie
- je präziser und praktisch verwertbarer mögliche Handlungsfolgen thematisiert werden.

Fremdevaluationen genießen in der Regel ein größeres öffentliches Vertrauen als Selbstevaluationen, auch wenn deren inhaltliche Validität umstritten ist. England praktiziert Inspektionen, während Schweden eher wieder davon abgekommen ist. Beide Länder würden aber der Aussage zustimmen, dass es sich lohnt, vermehrt in Meta-Evaluationen zu investieren. Hier wird nicht primär inhaltlich inspiziert, sondern es werden die Standards und die Güte der eigenen Selbstevaluation bewertet, indem deren Verfahren (dialogisch angelegte Peer-Reviews, Qualität der Schüler- und Elternbefragungen, aus einer Selbstevaluation gezogene Konsequenzen) unter die Lupe genommen werden. Dies geschieht auch vor dem Hintergrund der Einsicht, dass die Etablierung einer Evaluationskultur an den Schulen sich nicht von allein einstellt, sondern strategischer und methodischer Kompetenzen bedarf, die über die Ausbildung von Schulleitungen und das Angebot von bei Bedarf abrufbaren Instrumenten hinausgehen.¹⁹⁶

Gelingensbedingungen von Selbstevaluation

Selbstevaluationen werden in der Regel als wirksamer betrachtet in Bezug auf daraus gezogene Lerngewinne, aber auch nur dann, wenn ge-

¹⁹⁶ Vgl. dazu die Erfahrungen im NRW-Projekt „Selbstständige Schule“: <http://www.selbststaendige-schule.nrw.de> (Stand 12.07.06).

wisse Bedingungen erfüllt sind. Strittmatter (2000) formuliert die folgenden zehn Gelingensbedingungen nachhaltiger Selbstevaluation:

„Insgesamt lassen sich aus den Forschungen und Beobachtungen (z.B. von BeraterInnen) einige Regeln für das Gelingen nachhaltiger Selbstevaluation formulieren. Systematische Selbstevaluation wird dann zur festen Installation mit Gewinn für die Arbeit der Lehrpersonen und die Schulqualität,

(1) *wenn die Haltungen der Beteiligten stimmen*

Redlich wissen/lernen/mitsteuern wollen ist absolut entscheidend! Die FeedbackpartnerInnen nehmen die Haltung der Feedbacknehmenden Person in der Regel sehr aufmerksam wahr und verhalten sich entsprechend konstruktiv oder aber defensiv oder lustlos-fahrlässig. Und die Haltung bestimmt auch weitgehend, ob dann auch Folgerungen gezogen und umgesetzt werden;

(2) *wenn die Feedbackschleife durchgezogen wird*

- *vollständig* (inkl. Schlussfolgerungen, Umsetzung und Rückmeldung an Feedback Gebende),
- *trennscharf* (vor allem Datensichtung, Deutung und Wertung auseinander halten),
- häufig *partizipativ* (wo es wichtig und machbar ist, die Auswertung, Deutung und Schlussfolgerung mit den FeedbackpartnerInnen zusammen vornehmen);

(3) *wenn Konsequenzen machbar, sichtbar und freundlich erscheinen*

Blicke nur in Spiegel, wenn du weißt, dass du das, was du dort wahrscheinlich sehen wirst, anschauen und annehmen willst und kannst und es sich wahrscheinlich in etwas Produktives umsetzen lässt;

(4) *wenn eine SOFT-Balance gesucht wird*

Die ‚Mea-Culpa-Fehlersuche‘ mag eine edle christlich-pädagogische Attitude sein; mindestens ebenso wichtig ist das Entdecken und Bestätigen von Stärken, das Finden von Ausbaumöglichkeiten und das frühzeitige Erkennen von Gefährdungen.

(5) *wenn die Interessen/Zwecke/Rollen geklärt und voll akzeptiert sind*

Menschen in Feedbacksituationen fragen sich unwillkürlich, wer da wozu und mit welchem Hut auf dem Kopf etwas wissen will und wer dann sonst noch mit dem entstehenden Wissen bedient wird und was anstellt. Sie verhalten sich entsprechend offen, sorgfältig, konstruktiv, zugeknöpft oder sorglos. Ergiebige und faire Feedbacksituationen entstehen bei klar definierten und deklarierten Zwecken und Rollen.

(6) *wenn die Beziehungen in der Organisation vertrauenswürdig sind*

In einem verlässlichen Team- und Schulleitungsklima wird mehr Offenheit möglich, sinkt die Verletzungsgefahr, vermeidet man Selbst- und Fremdüberforderungen. Für heikle Situationen gibt es Orientierungswerte und Spielregeln.

(7) *wenn Wissen/Erfahrung nutzvoll weitergeleitet wird*

Selbstevaluation dient in erster Linie der Beschaffung von Steuerungs- und Rechtfertigungswissen. ‚Gutes Wissen‘ (im sokratischen Sinn) kann und soll anderen zur Verfügung gestellt werden: horizontal für die KollegInnen (voneinander lernen) und vertikal für die Schulleitungsorgane (für gute Arbeitsbedingungen und Vertrauen sorgen).

(8) *wenn die Energie dafür vorhanden/reserviert ist*

Seriöse Selbstevaluation braucht zeitliche, intellektuelle und emotionale Energie. Die Räume für den Aufwand müssen reserviert sein und geschützt werden. Und es sollte wenig energetische Konkurrenz durch andere Projekte/Ärgerthemen etc. bestehen. Nötigenfalls sollte der FQS-Aufwand bewusst angepasst werden.

(9) *wenn individuelle Umstände/Abweichungen einen erlaubten Platz haben*

Es ist möglich, dass individuell bestimmte FQS-Praktiken überfordernd erscheinen oder momentan wenig wichtig. Es muss ermöglicht werden, eine ständige oder temporäre Anpassung der FQS-Praktiken an die persönlichen Umstände vorzunehmen – vorausgesetzt, die subjektiv guten Gründe werden offen deklariert und einer Erlaubnis durch das Kollegium, das Team oder die Leitungsorgane unterzogen.

(10) wenn über ein funktionales „Handwerkszeug“ verfügt wird

Zu guter Letzt (!) trägt viel zum Gelingen bei, wenn die Evaluationswerkzeuge kompetent, zweckmäßig und variabel eingesetzt werden können. Die Aussagekraft der gewonnenen Daten steigt dadurch, Selbstevaluation erhält mehr Vertrauen und die Routinen verringern den Aufwand. (Es erträgt hier aber mehr Fehler, wenn die Grundhaltungen [Punkt 1.] stimmen!)

Es müssen nicht von Beginn weg alle diese Faktoren schon vorhanden bzw. günstig ausgeprägt sein. Und es kommt vor, dass mal eine Zeit lang eine oder zwei dieser Bedingungen fehlen oder gar negativ ausgeprägt sind, ohne dass das ganze System abstürzt. Andererseits können in schwachen Grundkonstellationen ein oder zwei dauerhaft negativ ausgeprägte Faktoren das System schwer gefährden. Die genannten Regeln sind also weniger Erfolgsrezepte, sondern ein Dauerauftrag für die Optimierung der Selbstevaluationsarbeit.“

Wichtig ist, dass auch Selbstevaluationen, welche gegenüber externen Verfahren den Vorteil einer höheren Akzeptanz durch die Evaluierten besitzen, in Bezug auf ihre Formen und Meta-Standards nicht einfach den Schulen überlassen werden. Eine stärkere Standardisierung ist notwendig, wenn Selbstevaluation nicht überall etwas Verschiedenes heißen soll. Angebote können von den Ländern, Kantonen oder inzwischen auch von kommerziellen Anbietern bereitgestellt werden. Auch vom IQB in Deutschland sind Aufgaben und Materialien zu erwarten, die für interne Evaluationszwecke von Schulen in Frage kommen. Eine Frage kann den Punkt betreffen, welche Evaluationen regelmäßig und verpflichtend durchzuführen seien (vgl. Schweden, Niederlande, England) und welche freiwillig bleiben sollen.

Weil systematische und verpflichtende, sowohl externe als auch interne Schulevaluationen in den deutschsprachigen Ländern bisher kaum eine Rolle gespielt haben, wird es notwendig sein, Evaluationsberater auszubilden und einzusetzen.

Begrüßenswert sind in diesem Zusammenhang Initiativen wie der BLK-Modellversuch zur Evaluation und Selbstevaluation von Schulen. Wie gesagt sollte Beliebigkeit vermieden werden. Deshalb stellt sich die Frage nach einer zentral koordinierten Entwicklung von Evaluationssystemen, um die Vergleichbarkeit der erhobenen Daten zu sichern. Ohne die Koordinierungsfunktion zentraler Einrichtungen drohen die

vielfältigen Evaluationsansätze in ein unüberschaubares Nebeneinander einzelner Initiativen zu zerfallen.

5.2.4 Kooperation von Lehrkräften in Fachgruppen und Netzwerken

Traditionelle Weiterbildung allein genügt nicht, um die Unterrichtspraxis zu verändern (Elmore 1996). Es braucht dazu Kooperation von Lehrpersonen in Fachschaften und (schulübergreifenden) Netzwerken („networking“). Beides stellt ein zentrales und mittlerweile bewährtes Werkzeug im Bereich der Unterrichts- und Schulentwicklung dar. Sie tragen zur Entwicklung der Schulkultur bei und fördern die Unterrichtsexpertise (Leithwood 2000; Rolff 2001). Nachfolgend werden Möglichkeiten der kooperativen Lehrerarbeit im Lichte von Erfahrungen und Studien im Hinblick auf Bildungsstandards vorgestellt – im Sinne von Voraussetzungen, Gelingensbedingungen, Handlungsregeln, praktischen Folgerungen und Herausforderungen.

Schulen repräsentieren individualistische Berufskulturen. Obgleich Lehrpersonen den Arbeitsort teilen und – je nach Schulform und Stufe – dieselben Klassen unterrichten, ist Unterrichtsarbeit als das eigentliche „Kerngeschäft“ in der Regel nach wie vor (anonyme) Einzelarbeit. Immer noch funktionieren viele Kollegien nach einem Autonomie-Paritäts-Muster (Lortie 1975; Altrichter/Eder 2004), wonach ihre Mitglieder sich in ihrer Berufsrolle als Einzelkämpfer mit einem ausgeprägten Gleichheitsgefühl verstehen und kein Kollege bzw. keine Kollegin sich deshalb traut, den anderen Vorschläge zu machen oder Kritik auszutauschen – kurzum: man lässt sich in Ruhe. Von außen betrachtet weisen viele Einzelschulen denn auch eine gefügeartige Gesamtstruktur auf (Rolff 1993), sie erscheinen eher als lose gekoppelte Systeme (Fend 1986), denn als nach Kohärenz strebende pädagogische Handlungseinheiten. Erst seit in den neunziger Jahren im deutschen Sprachraum ein Trend zu teilautonomen Schulen eingesetzt hat, waren Lehrerkollegien aufgefordert, in Zusammenarbeit mit ihren (neu eingerichteten) Schulleitungen Leitbilder für ihre Schulen und ein Qualitätsmanagement zu entwickeln. Wo dies geschah, bezog sich Kooperation bald nicht mehr nur auf organisatorische Belange, sondern erstreckte sich zunehmend auf Kernziele und Inhalte einer pädagogischen Schulentwicklung und wurde – neben Merkmalen wie Innovationsbereitschaft und pädagogi-

schem Engagement – als ein konstitutives Merkmal guter Schulen identifiziert (Steffens/Bargel 1993).

Seit sich Befunde erhärtet haben, dass die Qualität kollegialer Zusammenarbeit zwischen Lehrkräften ein zwischen guten und verbesserungsbedürftigen Schulen differenzierendes Merkmal darstellt (Scheerens/Bosker 1997; Bensen 2005) und dass es erfolgreichen Systemen gelingt, durch die Orientierung der Schulentwicklung an gemeinsamen Zielen der Isolierung der Lehrkräfte bzw. dem „Einzelkämpfertum“ entgegenzuwirken, ist Kooperation unter Lehrkräften auch in den deutschsprachigen Ländern zu einem Forschungsgegenstand geworden. International werden Kooperation von Lehrkräften und die damit verbundene Bildung dichter, kommunizierender Netzwerke schon länger unter variablen Etiketten („collegiality“, „collaboration“, „professional community“, „schools that learn“, „networks“ oder „consensus and cohesion“) erforscht (vgl. Bensen 2005). Unterscheiden lassen sich Formen der innerschulischen Kooperation von Formen der interschulischen Zusammenarbeit in schulübergreifenden Netzwerken.

Im deutschsprachigen Raum hat vor allem der Begriff der „professionellen Lerngemeinschaften von Lehrerinnen und Lehrern“ Eingang in die Diskussion gefunden (Bensen/Rolff 2006). Mit professionellen Lerngemeinschaften in Form von *Fachkonferenzen und anderen Formen der innerschulischen Kooperation* werden primär folgende Ziele verbunden:

- intensive Auseinandersetzung über nachhaltig bedeutsame Unterrichtsinhalte,
- erfolgreicher Erwerb neuer Unterrichtstechniken und eine Erweiterung des professionellen Wissens,
- Erweiterung der Lehrerrolle, insbesondere in Bezug auf Funktionen der Lernunterstützung und Förderung von Schülerinnen und Schülern,
- eine verbesserte Fähigkeit, sich auf die Bedürfnisse individueller Lernender einzustellen,
- eine höhere Berufszufriedenheit sowie
- höhere Bereitschaft und Motivation für Veränderung.

Beispielhaft können hier die Kollegialen Evaluationen (KoEv) genannt werden, in deren Rahmen Gruppen von sieben bis neun angehenden Lehrpersonen unter professioneller Anleitung Erfahrungen und Handlungen aus ihrer Tätigkeit reflektieren, subjektive Theorien rekonstruieren

ren bzw. modifizieren und neue Handlungsmöglichkeiten entwickeln. Angestrebt wird unter anderem ein Lernprinzip der doppelten Vermittlung – d.h. eines Lernens, bei dem die Teilnehmenden pädagogische Theorien mit Hilfe der übrigen Gruppenteilnehmer in der Selbstanwendung erfahren und auf diese Weise den Hiatus zwischen Theorie und Praxis zu überwinden beginnen (Neveling 2005).

In einer Analyse solcher Lerngemeinschaften in Deutschland zeigte sich, dass sich diese generell durch folgende Merkmale auszeichnen: durch ein höheres Ausmaß an kollegialer Reflexion, eine De-Privatisierung des unterrichtlichen Handelns, einen stärkeren Fokus auf Lernen statt auf Lehren und eine erhöhte Ausrichtung auf gemeinsame handlungsleitende Ziele (Bonsen/Rolff 2006).

Wodurch zeichnet sich Kooperation unter Lehrkräften aus bzw. welche Aspekte sind von Bedeutung? Als wichtig hat sich erwiesen, dass Kooperation für die Beteiligten erfahrbar macht, dass es sich lohnt und einen Mehrwert im Sinne von Arbeitserleichterungen und einer verbesserten Unterrichtsqualität mit sich bringt, wenn man in Fachgruppen zusammenarbeitet und eine Rückmeldekultur im Kollegium aufgebaut wird.

Exemplarisch sei hier auf eine schweizerisch-deutsche Studie verwiesen. Steinert/Klieme (2003) haben auf der Basis einer vom Deutschen Institut für Internationale Pädagogische Forschung (DIPF) durchgeführten Befragung von mehr als 4000 Lehrpersonen an 93 Schulen mit einer Rasch-Skalierung 20 Items ermittelt, welche nicht nur die berufliche Kooperation von Lehrpersonen beschreiben, sondern auch eine Zuteilung der Schulen zu einem höheren oder tieferen Level der Kooperation ermöglichen. Dabei werden sowohl unterrichtsbezogene Aspekte von Kooperation (z.B. „Wir haben eine gute fachspezifische Zusammenarbeit“) als auch organisationale Merkmale (z.B. „Wir werden rechtzeitig und ausreichend über wichtige Vorgänge informiert“) oder Aspekte des Personalmanagements (z.B. „Wir erarbeiten gemeinsam Strategien zur Bewältigung beruflicher Probleme“) berücksichtigt. Die am leichtesten zu realisierenden Formen von Kooperation bezogen sich auf fachspezifische Zusammenarbeit, Erarbeitung eines Schulprofils, kollegiale Mitbestimmung, Koordination der Unterrichtsarbeit innerhalb der Jahrgangsstufe, Informationsfluss und Beteiligung an der Festlegung der Stundenpläne (Steinert/Klieme/Maag Merki/Döbrich/Halbheer/Kunz 2006). Im Kanton Zürich wurden mit demselben Instrument die Lehrpersonen von neun Gymnasien befragt. Aufgrund ih-

rer Mittelwerte wurden die Schulen einem bestimmten Kooperationslevel zugeordnet. Im Rahmen eines Extremvergleichs konnte gezeigt werden, dass Schulen mit hohem Kooperationsgrad auch höhere Werte hinsichtlich verschiedener Skalen der Schulorganisation (z.B. Schulleitungshandeln, Innovationsbereitschaft) und der individuellen Ebene (z.B. geringfügigeres Erleben von Belastung), nicht aber hinsichtlich von Aspekten der Unterrichtsqualität erreichen (Kunz/Halbheer/Maag Merki 2005). Steiner/Klieme/Maag Merki/Döbrich/Halbheer/Kunz (2006) berichten Unterschiede zwischen verschiedenen Schulformen bezüglich der auf diese Weise erhobenen Kooperation: So erzielten in Hessen Gymnasien gesamthaft tiefere Grade der Kooperation als andere Formen der Sekundarstufe (ebd.), was angesichts der Binnenstruktur der Ersteren nicht erstaunt.

Während die innerdeutschen Unterschiede gut erklärbar sind, trifft dies auf die ebenfalls aufgetretenen Differenzen zwischen deutschen (hessischen) und Schweizer (Zürcher) Schulen nicht ohne Weiteres zu. In der Untersuchung zeigten sich durchgängig signifikante Unterschiede bezüglich sämtlicher erwähnten Aspekte von Kooperation. Und zwar erzielten die Schweizer Schulen die höheren Kooperationswerte als die hessischen (ebd.). Warum dies so ist, ist eine interessante Frage. Zum einen besteht die Möglichkeit, dass der im Zürcher Mittelschulgesetz von 1999 im Rahmen der Teilautonomie der Mittelschulen verankerte Auftrag zum Qualitätsmanagement die Ursache für eine veränderte Zusammenarbeitskultur bildet, während in Hessen die Schulaufsicht und die Regelung personaler und finanzieller Belange zentral geregelt sind und dem Bundesland obliegen. Andererseits ist es denkbar, dass die Ergebnisse Ausdruck eines je unterschiedlichen kulturellen Verständnisses von Kooperation sind, in dem Sinne, als Schweizer und deutsche Lehrpersonen von einem je anderen Berufsverständnis ausgehen. Kelchtermans (2006) erwähnt den Einfluss kontextueller Faktoren für die Entstehung von Kooperation:

„In the same vein, professional learning communities – in which collaboration and collegiality are supposed to play a key role – ought to be conceived of not so much as structural arrangements, but rather as cultural and political environments in which those forms of collaboration and collegiality can take place that really contribute to pupils' learning, teacher development and quality of school improvement“ (ebd., S. 234).

Inwiefern es sich im vorliegenden Fall um kulturelle Differenzen handelt, muss an dieser Stelle offen bleiben.

Weiter laufen unter den Begriffen „Bildungsnetzwerke“ oder „regionale Bildungslandschaften“ seit einigen Jahren verschiedene regionale und auch internationale Versuche wie etwa QiS (Quality in School) oder INIS (Internationales Netzwerk innovativer Schulsysteme), welche beide Qualitätsverbesserungen anstreben und zu diesem Zweck Kriterien formulieren, welche extern evaluiert werden (vgl. Solzbacher/Minderop 2007). Differenzierte und umfassende Rückmeldungen zu Auswirkungen auf verschiedenen Ebenen fehlen bislang.

Während bisher vor allem Prozesse, Formen und Wirkungen *innerschulischer Kooperation* angesprochen wurden, ist auf einen zweiten Typus zu verweisen: die *schulübergreifende (über)fachliche Netzwerkarbeit*. Als prototypisches Erfolgsmodell schulübergreifender Kooperation kann das bereits erwähnte (vgl. Kap. 5.1.7) BLK-Modellversuchsprogramm zur „Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts“ (SINUS, SINUS-Transfer) gelten (vgl. Prenzel 2000). Dem Programm liegt ein Implementationsansatz zugrunde, bei dem Lehrkräfte problemorientiert an Modulen ihren Unterricht weiterentwickeln, und in dem Kooperation eines der zentralen Prinzipien ist. Der Ansatz zeichnet sich durch folgende Merkmale aus (Ostermeier/Carstensen/Prenzel/Geiser 2004):

- Die Unterrichtsentwicklung erfolgt schulnah an Problembereichen des mathematisch-naturwissenschaftlichen Unterrichts. Ausgangspunkt ist ein System von elf Modulen, die so zugeschnitten sind, dass sie von Lehrergruppen an unterschiedlichen Schulen (für Lehrpersonen und Schülerinnen und Schüler) erfolgreich bearbeitet werden können.
- Kooperation ist ein zentrales Leitprinzip im Programm. Die beteiligten Lehrkräfte werden über die modulbezogene Arbeit an ihren Schulen mit Verfahren der kooperativen Unterrichtsentwicklung vertraut gemacht. Ziel ist eine langfristige Etablierung dieser Prozesse.
- Einrichtung von regionalen und überregionalen Schulnetzen, die ihre Modulbearbeitung abstimmen, Entwicklungen erproben, Erfahrungen auswerten und gemeinsam Fortbildungen auf Schul- und Netzwerkebene organisieren.

- Die Zusammenarbeit wird an der Schule selbst sowie auf der Ebene der Schulnetzwerke durch Koordinatorinnen und Koordinatoren unterstützt. Die Unterstützung beinhaltet Anregungen und Hilfen für die Unterrichtsentwicklung, Organisation des Austauschs und Dokumentation der Ergebnisse und Prozesse.
- Ein zentraler Programmträger hält Handreichungen, Materialien, Fortbildungs- und Beratungsangebote zu den Modulen bereit und sorgt durch Kooperation mit Einrichtungen, z.B. der Lehrer(fort)bildung, für eine unterstützende und innovationsfreundliche Infrastruktur.
- Einrichtung eines zentralen Servers zum Informationsmanagement und zur Unterstützung des Austauschs.
- Begleitung der Schul- und Netzwerkarbeit durch abgestimmte Evaluationsmaßnahmen, die den Schulen und dem Programmträger zuverlässige Rückmeldungen über Erfolg und Probleme des Programms liefern.

Unter den angebotenen Modulen können die Lehrpersonen diejenigen auswählen, welche ihnen bedeutsam erscheinen und auch über die Wahl der Teilaspekte bestimmen (ebd.).

Im Rahmen einer summativen Evaluation, in welcher 102 Schulen des Modellversuchs mit Schulen aus der PISA-Erweiterungstichprobe hinsichtlich Einstellungen gegenüber den Fächern Mathematik und Naturwissenschaften sowie bezüglich der erzielten Fachleistungen verglichen wurden (Prenzel/Carstensen/Senkbeil/Ostermeier/Seidel 2005), wurde auch nach der Häufigkeit von Kooperation gefragt. Die Resultate ergaben, dass

- die Häufigkeit der Kooperation vor allem in den Hauptschulen, in den Schulen mit mehreren Bildungsgängen und in den Realschulen der SINUS-Stichprobe höher liegt als bei den Schulen der Kontrollgruppe, dass dieser Effekt aber bei den Gymnasien und Integrierten Gesamtschulen nicht eintritt;
- die Leistungen in den SINUS-Schulen sowohl in Mathematik und in Naturwissenschaften in Hauptschulen, MBG und Integrierten Gesamtschulen höher liegen, nicht aber in Gymnasien und Realschulen;
- die Unterschiede insgesamt in den genannten Bereichen auf den tieferen Schulstufen stärker zu sein scheinen als in den Gymnasien (ebd.).

Ostermeier (2004) hat sodann die Entwicklung der Akzeptanz des Programms vertieft unter die Lupe genommen. Auf der Basis einer Befragung von über 500 Lehrpersonen aus 161 Schulen, ihren Schulleitungen sowie den Schüler- und Elternvertretungen ergaben sich folgende Befunde:

- Die Kooperationsaktivitäten der Lehrpersonen nahmen von Messzeitpunkt 1 zu Messzeitpunkt 2 leicht ab.
- Die Lehrpersonen erlebten die Kooperation schon während des ersten Messzeitpunktes als positiv – ein Effekt, der sich im Rahmen der zweiten Erhebung sogar noch verstärkte.
- Ebenso berichteten die Befragten über eine hohe Zufriedenheit mit dem Programm, welche ebenfalls in dessen Verlauf noch anstieg (ebd.).

Bilanzierend lässt sich zum SINUS-Modell und weiteren Formen der regelmäßigen kollegial-fachlichen und pädagogischen Zusammenarbeit feststellen, dass Kooperation dort eindeutige Effekte zeigt, wo sie unterrichtsbezogen, entwicklungsorientiert und bezogen auf gemeinsame Ziele erfolgt. Hier ist auch der wichtige Vorteil von Bildungsstandards zu sehen: Fachgruppen und Fachkonferenzen können mit Hilfe der Bildungsstandards besser kooperieren, weil für Planung und Auswertung eine gemeinsame Grundlage gegeben ist. Dabei ist zu beachten, dass Kooperation nicht nach einem festgelegten Muster abläuft, welches – über die Steuerung des Gesamtsystems eingeführt – für alle bestimmend werden könnte, sondern als kontextabhängig und situationspezifisch wahrgenommen wird. Die Angebots-Nutzungs-Struktur, wie sie die SINUS-Module darstellen, kann als Positivbeispiel dienen. Das SINUS-Programm zeigt überdies, dass eine fachbezogene, modularisierte Kooperation zwischen Lehrkräften in Netzwerken mindestens bei bestimmten Schularten¹⁹⁷ leistungsfördernde Effekte mit sich bringt. Eine Bedingung dafür ist der aktive Einbezug der Lehrkräfte in alle relevanten Entscheidungen und die klare Instrumentierung der Netzwerkarbeit. Die Lehrkräfte müssen bei ihrer Arbeit konkrete und empiri-

¹⁹⁷ Nach den vorliegenden Daten sind dies Hauptschulen, Schulen mit mehreren Bildungsgängen sowie Integrierte Gesamtschulen (Prenzel/Carstensen/Senkbeil/Ostermeier/Seidel 2005).

risch erprobte Anregungen erhalten, wie sie vorgehen können, wenn sie eine Repertoireerweiterung ihres Unterrichts vornehmen sollen.

Eine ähnliche Strategie verfolgt das durch 14 Bundesländer geförderte Projekt „Chemie im Kontext“ (CHiK; vgl. <http://www.chik.de/index2.htm>). Unter dem Stichwort „Symbiotische Implementationsstrategie“ (Gräsel/Parchmann 2004) werden Sets von 12-16 Lehrpersonen zusammengefasst, um neue Unterrichtseinheiten für das Fach Chemie zu entwickeln, welche sich dadurch charakterisieren lassen, dass sie von einem ko-konstruktivistischen Lernverständnis der Schülerinnen und Schüler ausgehen und die Bedeutung des Faches für das alltägliche Leben im Zentrum steht. Die Praxiserprobung durch die Lehrpersonen wird im entsprechenden Set diskutiert und reflektiert (Bünder/Parchmann 2004). Im Rahmen eines Längsschnittes wird das gesamte Projekt wissenschaftlich begleitet durch ein Team von Forscherinnen und Forschern des Leibniz-Institut für die Pädagogik der Naturwissenschaften in Kiel (IPN) sowie der Universitäten Wuppertal, Oldenburg und Dortmund. Ein Vergleich mit Kontrollklassen ergab u.a. günstigere Motivationswerte für die Probandinnen und Probanden der CHiK-Klassen. Ebenso stellen die befragten Schülerinnen und Schüler im Längsschnitt eine zunehmend effizientere Zeitnutzung der unterrichtenden Lehrpersonen fest (Chemie im Kontext 2007).

Von entscheidender Bedeutung ist auch, dass die im Zuge solcher Modellversuche gewonnenen positiven Erfahrungen publik werden und Schritte zur Verbreitung der besonders innovativen Elemente unternommen werden. Eine Befragung unterschiedlicher Zielgruppen (u.a. Ministerien, Schulaufsicht, Lehrerfortbildung) in verschiedenen deutschen Bundesländern ergab eine insgesamt gute Verankerung der SINUS-Transfer-Module, eine weite Verbreitung und eine positive Akzeptanz der Grundprinzipien, insbesondere der Kooperation, aber auch insgesamt geringe Kenntnisse über die Module und Grundprinzipien bei bestimmten Akteurgruppen (Jäger 2006). Neben der Bildungsadministration müssen insbesondere Lehrerbildner und die Lehrpersonen selber das innovative Potenzial von Netzwerken erkennen, nicht zuletzt um allfällige Berührungängste abzubauen zu können und sich möglichst vorurteilsfrei auf das Neue einlassen zu können. Diesbezüglich sind die im Zuge von SINUS-Transfer angesetzten Maßnahmen zur Dissemination des Programms als besonders wertvoll zu beurteilen. Ein eigentliches Transferkonzept soll erreichen, dass der Ansatz kooperativer unterrichtsbezogener Qualitätsentwicklung langfristig in die Breite getragen und nachhaltig in den teilnehmenden Schulen verankert wird (Zieg-

ler 2006). Im Zuge solcher Transfer- und Innovationsbemühungen gilt es immer auch zu beachten, dass die damit verbundenen Anstrengungen auf gewachsene Strukturen treffen, deren Veränderung eines ausgeprägtes Engagements bedürfen (Euler/Sloane 1998).

Netzwerkarbeit im überfachlichen Sinn hat sich auch im Projekt „Selbstständige Schule“ bewährt. Hier ist vor allem die Ebenenverschränkung in der Qualitätssicherung hervorzuheben. Aus Mitteln des Landes sowie der Kommunen sind regionale Entwicklungsfonds aufgebaut worden, die die regionalen Steuerungsgruppen wirksam unterstützen. In verschiedenen Modellregionen entstanden Bildungsbüros, die mit der Koordination und Steuerung der Qualitätssicherung beauftragt wurden, was einen effizienten Einsatz der Mittel sichert und Mehrspurigkeiten verhindert. Steht nur die einzelne Schule im Fokus, dann muss sie auch die ganze Entwicklungsarbeit leisten; Schulen lernen generell aber dann am besten, wenn sie ortsnahe und leicht zugängliche Beispiele zur Verfügung haben.

Die Implementation von Bildungsstandards verlangt von den Schulen den Aufbau neuer Verfahren, die Anpassung des Arbeitswissens der Organisation und nicht zuletzt die Überwindung hochgradiger Unsicherheiten. Die Erfolgskriterien ändern sich, die Standards steigern die Ansprüche an den Unterricht und die Ergebnisse werden transparent. Wenn diese durchaus riskanten Prozesse gelingen sollen, ist eine engmaschige Kooperation vor Ort und in der Region unausweichlich. Dabei haben sich schulische wie außerschulische Netzwerke insgesamt als von großer Bedeutung erwiesen; sie sorgen für wirksamen kollegialen Austausch im Fach und darauf bezogen für ein direktes Peer-Coaching. Nur durch ihre Etablierung können im Übrigen auch die beträchtlichen Investitionen in Implementationsstrategien wirksam eingesetzt werden. Anders wäre die einzelne Schule gehalten, jeweils den gesamten Entwicklungszyklus bei sich nachzuvollziehen.

Mit Blick auf die neuen Instrumente der Implementation, also Kerncurricula, Tests, Aufgabenkulturen und Rückmeldesysteme, sollten frühzeitig regionale Spurguppen eingerichtet werden, die für eine wirkungsvolle Vernetzung der Entwicklung sorgen. Ähnlich wie bei SINUS müssen Instrumente und Handreichungen zur Verfügung stehen, in deren Gebrauch die Lehrkräfte eingeführt werden. Die einzelnen Schulen können bestimmte Lehrkräfte damit beauftragen, die dann als innerschulische Multiplikatoren eingesetzt werden. So hat die Implementation von Bildungsstandards einen konkreten Adressaten und

zugleich eine regionale Verbreitung. Anders ist die Gefahr groß, dass eine diffuse oder chaotische Praxis entsteht, die schnell hohe Motivationsverluste mit sich bringt und, wenn überhaupt, nur eine ganz schwache Akzeptanz erzeugt.

5.3 Unterrichts- und Klassenebene

Steuerungsmaßnahmen in Bezug auf die Einführung von Bildungsstandards werden nur dann Erfolg haben, wenn es gelingt, sie für den Kern der pädagogischen Qualitätsentwicklung, für die Entwicklung des Unterrichts, fruchtbar zu machen. Die Frage lautet somit, wie Bildungsstandards in den Schulalltag Eingang finden können, oder mit welcher Art von Schulentwicklung nachhaltige Unterrichtsentwicklung mit Breitenwirkung ausgelöst und unterstützt werden kann. Eine Implementation von Bildungsstandards, die nicht bis auf die *Mikroebene des Unterrichts* durchdringt und die die Lehrpersonen und letztendlich die Schülerinnen und Schüler als eigenständig Lernende nicht erreicht, wird nichts bewirken. Für das Lehren *und* das Lernen gilt: keine Qualität der Produkte ohne entsprechende Prozessqualität; keine standardkonformen Lernergebnisse ohne vorangegangene Lerngelegenheiten für alle Schülerinnen und Schüler. Auf der Lehr-Lernebene entscheidet sich, ob die Reform wirksam ist. Schulentwicklung ist in ihrem Kern *Unterrichtsentwicklung*. Alle Maßnahmen der Einführung und Umsetzung einer standardbezogenen Qualitätssicherung – soweit sie die Schule und nicht deren weiteren Kontext betreffen – müssen deshalb mittelbar oder unmittelbar auf die Kernprozesse des Lehrens und Lernens und die Verbesserung ihrer Bedingungen gerichtet sein. Diese Erkenntnis steht im Einklang mit Forschungsergebnissen, die zunehmend die zentrale Bedeutung der Klassenebene – und damit von Unterrichtseffekten – für das Zustandekommen von Schulqualität und Schulleistung dokumentieren (vgl. Ditton, 2000). Stabile und substanzielle Verbesserungen von Schülerleistungen und weiteren Lernerträgen sind nur dann zu erwarten, wenn alle Lehrpersonen einer Schule einen qualitativ hochwertigen Unterricht anbieten und sich an dessen Ergebnissen als Bezugspunkt ihrer täglichen Arbeit orientieren.

Lehrerinnen und Lehrer sind somit auf Schulebene die wesentlichen Schlüsselfaktoren des Lernerfolgs ihrer Schülerinnen und Schüler und daher die Hauptträger und Hauptverantwortlichen der Reform. Sie gilt es zu aktivieren und mit Ressourcen und wirksamen Instrumenten, die

den Unterricht und das Lernen der Schülerinnen und Schüler im Blick haben, auszustatten. Mit anderen Worten: Es ist naiv, „standardbasierten Unterricht“ zu postulieren, ohne dementsprechende Entwicklungsprozesse, die auf eine Verbesserung der *unterrichtsbezogenen Lehrerprofessionalität* abzielen, in Gang zu setzen. Die Transformationsprozesse zwischen dem Erreichen eines geforderten Kompetenzniveaus und dem Unterricht sind dabei ungleich komplexer, als es eine naive Steuerungsstrategie ahnen lässt, welche fordert, „Bildungsstandards in didaktische Module zu übersetzen“. Auch ist es *nicht* das Ziel von Bildungsstandards, den Unterricht zu standardisieren. Standards definieren normative Erwartungen, auf die hin Schule bilden soll. Das verlangt über deren moralische Verinnerlichung hinaus nach einem tiefen fachdidaktischen Verständnis der Lehr-Lernprozesse. Lehrpersonen müssen nicht nur Bedingungen und Prozesse – die Lerngelegenheiten – gestalten können, die zu erwünschten Lernergebnissen führen, sie müssen auch wissen, an welchen Prozessfaktoren – zum Beispiel nach Rückmeldungen auf Tests und Vergleichsarbeiten – sie ansetzen können, um ihre Ergebnisse zu verbessern. Bildungsstandards im Unterricht umzusetzen bedeutet nicht nur zu wissen, welche (nicht-linearen) Prozesse mit einiger Wahrscheinlichkeit welchen „Output“ bewirken, sondern auch eine Rückkoppelung dieses „Outputs“ zum Prozess herzustellen und damit die Frage zu beantworten, wie sich extern erzeugte Leistungsinformationen in verbesserte Lernprozesse rückverwandeln lassen. Lehrpersonen müssen gute Intuitionen darüber entwickeln, wie Kompetenzen in Prozesse übersetzt werden bzw. wie sich ein kompetenzorientiertes Verständnis von Bildungsstandards in ein lehr-lerntheoretisches Aneignungsverständnis rückverflüssigen lässt.

Dass die Anforderungen an eine systemisch wirksame Unterrichtsentwicklung, die solches leistet, hoch sind, zeigen bereits Studien aus den neunziger Jahren (vgl. Kap. 3.1.4). Sie machen deutlich, dass die Implementation von standardbasierten Reformen, deren Erfolg an der tatsächlichen Wirkung auf Unterricht und Lernen gemessen wird, ein langwieriger Prozess ist. Diesen Befund bestätigt auch die Durchsicht von neueren Evaluationen. Eine Studie zu New Jersey zeigt z.B., dass viele Lehrpersonen auch nach intensiver, an neuen Standards orientierter Weiterbildung große Schwierigkeiten hatten, eine Lektion mit Aktivitäten zu planen, die auf einen selber ausgewählten Fachstandard hin führen (O’Shea 2003). O’Shea schließt daraus, dass die Abstimmung („alignment“) von Weiterbildung mit Standards und Curriculum nicht

ausreiche, wenn nicht die konkrete Unterrichtsplanung mit Fokus auf tief verstandene Lernprozesse und Lernergebnisse, die diese Standards einlösen, den Gegenstand bilde:

„(...) teachers need to learn how to plan a standards-based lesson that focuses on the achievement of outcomes related to the standards rather than simply conducting an activity aligned with standards. Furthermore, teachers need to learn how to evaluate student work samples in relation to the performance outcomes included in the standards-based lessons“ (ebd., S. 16).

Aufgrund der in dieser Expertise insgesamt analysierten Forschungsliteratur kann vermutet werden, dass

- die präzise Beschreibung von Kompetenzstandards und -zielen,
- einschließlich deren Umsetzung in Test- und Lernaufgaben,
- verbunden mit dem Verfügbarmachen von differenzierten Diagnosen des Lernstandes von Schülerinnen und Schülern an Schulen und Lehrpersonen,

nicht ausreicht, um dem Ziel einer wirksamen Optimierung und Weiterentwicklung des Unterrichts näher zu kommen. Oftmals mit Bildungsstandards verknüpfte Erwartungen auf der Unterrichtsebene wie die Förderung:

- des Umgangs mit einer anforderungsbezogenen Aufgabenkultur,
- der Unterrichtsplanung im Hinblick auf definierte Leistungserwartungen,
- der diagnostischen Kompetenz der Lehrerinnen und Lehrer,
- des Umgangs mit Heterogenität,
- der Evaluation von Unterricht durch interne und externe Verfahren,
- der Arbeit mit Lehrplänen

dürften nicht leicht umzusetzen sein. Das heißt, dass es auch auf der Ebene des Unterrichts und des Schülerlernens ergänzender Massnahmen, Verfahren und Werkzeuge der Unterrichtsentwicklung und des Lehrerlernens bedarf.

Während eine Reihe von Werkzeugen bereits auf der Schulebene dargestellt worden ist, werden im Folgenden vier weitere angesprochen. Sie beziehen sich auf Maßnahmen zur *Entwicklung guten Unterrichts* und betreffen weiter die Bedeutung von *Lehrmitteln* und die damit verbundene Etablierung einer anforderungsbezogenen, auf ein kompetenzori-

entiertes Verständnis von Bildungsstandards gerichteten *Aufgabenkultur*, auf den Nutzen und Gebrauch unterrichtsbezogener diagnostischer und evaluativer *Fachtests* und auf zwei neuere Werkzeuge der Personalentwicklung, die erst seit kurzem mit Erfolg auch in der Lehrer(fort)bildung eingesetzt werden: das *fachpädagogisch-didaktische Coaching* und die *videobasierte Unterrichtsreflexion*. Beide Instrumente besitzen ein breites Potenzial hinsichtlich der Möglichkeiten der Ausrichtung auf Qualitätsziele, wie sie von den Bildungsstandards angesprochen werden.

Abschließend wird auch noch die Schülerinnen- und Schüler Ebene thematisiert. Standards definieren verbindliche Erwartungen nicht nur für Lehrpersonen, sondern auch für die Lernenden. Den Schülerinnen und Schülern bleibt es – gemäß dem in dieser Expertise vertretenen Angebots-Nutzungsmodell – letztlich ein Stück weit selber überlassen, inwiefern sie die von der Schule bereitgestellten Lerngelegenheiten nutzen wollen. Die Schülerinnen und Schüler – in ihrem idealen Selbstverständnis als eigenständig Lernende – für die Ausrichtung der Lernaktivität an Bildungsstandards zu gewinnen, bleibt letztlich die kardinale Herausforderung, der sich das Unternehmen Bildungsstandards stellen muss.

5.3.1 Unterrichtsentwicklung als Kernaufgabe

Über etliche Jahre hinweg diente der Begriff der Schulentwicklung als Sammelkategorie für verschiedenste Innovationsprozesse auf der Meso- wie auch auf der Mikroebene des Schulsystems. Im Zuge einer eher dürftigen Befundlage zur Wirksamkeit von Maßnahmen, welche vor allem auf die organisatorische Entwicklung von Schulen zielen, ist indes eine Akzentverschiebung zu beobachten, indem der Fokus vermehrt auf unterrichtliche Innovation gelegt wird. Davon zeugen auch die zahlreicher werdenden Publikationen zum Thema der Unterrichtsentwicklung (z.B. Altrichter/Posch 1998; Klippert 2004; Bastian 2007; Gschwend/Claude 2004).

Es besteht mittlerweile Konsens darüber, dass es sich bei Maßnahmen, die der Weiterentwicklung des Unterrichts dienen, um den Kern der Schulentwicklung handelt (vgl. Klippert 2004). Gemeinhin werden unter dem Begriff systematische und gemeinsame Anstrengungen der am Unterricht beteiligten Akteure verstanden, welche zur Verbesserung des Lehrens und Lernens und seiner schulinternen Bedingungen beitragen können (Bastian 2007). Allerdings scheint auch hier die Gefahr

groß zu sein, dass dieser wenig trennscharfe Begriff an den damit verbundenen Übererwartungen scheitert, wenn außer Acht gelassen wird, dass jegliche Reform im Schulwesen auf ein komplexes und ausbalanciertes System trifft, das sich nur bedingt zielgerichtet und nachhaltig beeinflussen lässt. Vor dem Glauben an lineare Umsetzungsprozesse ist zu warnen und anstelle von Top-Down-Ansätzen sollen symbiotische Implementationsstrategien postuliert werden (vgl. Gräsel/Parchmann 2004; Gärtner 2007).

Reformmaßnahmen müssen die Lehrpersonen als Hauptakteure nicht nur rhetorisch, sondern tatsächlich erreichen, d.h. Vorgaben müssen in die Unterrichtspraxis integrierbar und alltagstauglich erscheinen und deren Wirkungen im Unterrichtsalltag erfahrbar sein (vgl. Bastian 2007; Klippert, 2004). Schulen und Lehrpersonen müssen daran glauben können, dass ihre Bemühungen zu Verbesserungen im eigenen Unterricht führen (Niggli 2000). Auch reicht es nicht aus, lediglich Ergebniserwartungen an Schulen zu kommunizieren oder bloß äusserlich die Bedingungen für eine Weiterentwicklung des Unterrichts zu schaffen – z.B. durch die Bereitstellung von Zeitgefäßen. Von zentraler Bedeutung ist, dass konkrete Verfahren und Werkzeuge in der Fortbildungsarbeit vermittelt werden, durch die Lehrpersonen und lokale Lehrerteams darin unterstützt werden, ihren Unterricht mit Blick auf Wissen und Können sowie eigenverantwortliches Lernen und Arbeiten *aller* Schülerinnen und Schüler weiterzuentwickeln. Besondere Herausforderungen stellt dabei zweifellos der didaktisch überlegte Umgang mit Heterogenität dar, insbesondere mit Blick auf die Förderung von (bildungsfernen) Jugendlichen mit Migrationshintergrund sowie die Schulformen mit Grundansprüchen, in denen sich viele leistungsschwache Schülerinnen und Schüler befinden.

Inwieweit die Implementation von Bildungsstandards durch Unterrichtsentwicklungsmaßnahmen wirksam gefördert werden kann, hängt u.a. davon ab, ob

- es gelingt, die Lehrerschaft von der Notwendigkeit der Reform zu überzeugen und die Standards als Zielvorgaben für einen (minimalen) Bildungskern internalisiert werden („shift in ‚reform ownership‘ from ‚external‘ to ‚internal‘“; Coburn 2003),
- ein Professionalitätsverständnis den Konsens der Lehrerschaft und die Unterstützung durch das „System“ findet, in dem sich Lehrerinnen und Lehrer – stärker als bisher – als selbstwirksame, verant-

- wortlich handelnde und reflektierende professionelle Praktiker verstehen,
- es gelingt, Unterrichtsentwicklung als verbindliche Aufgabe teilautonomer Schulen und Lehrerteams pragmatisch und bottom-up in lokalen Schulkontexten zu verankern,
 - die den Schulen und Lehrpersonen angebotenen Unterrichtsentwicklungsinstrumente es erlauben, sich (fach)didaktisch konkret und orientiert an komplexen, lokal gültigen Praxissituationen mit der Lehr-Lernprozesseite von Bildungsstandards zu beschäftigen,
 - ob die Entwicklungsarbeit sich an bedeutsamen, durch empirische Forschung und Evidenz erhärteten Dimensionen guten und wirksamen Unterrichts orientiert, d.h. in der inhaltlichen Ausrichtung nicht beliebig und lediglich selbstreferenziell ist,
 - genügend professionelle Ressourcen (Coaching-Angebote, Austausch- und Feedbackmöglichkeiten, Handreichungen, Best-Practice-Materialien und Dokumentationswerkzeuge) auf regionalen und überregionalen Ebenen zur Verfügung stehen, um einzelne Lehrpersonen, Lehrerteams und ganze Kollegien zu beraten und zu unterstützen,
 - die Fortbildungsarbeit auf der Unterrichtsebene nicht nur verpflichtend eingefordert und durch abgestimmte Evaluationsmaßnahmen überprüft, sondern auch berufsfeldbezogene, strukturelle Anreize sowie eine unterstützende Infrastruktur dafür geschaffen werden.

Eine wichtige Erkenntnis der Schulentwicklungsforschung besteht zudem darin, dass nachhaltige produktive Entwicklungsarbeit schulnah, d.h. – nicht ausschließlich, jedoch zu einem wichtigen Teil – „vor Ort“ und in Fachteams stattfinden soll (vgl. u.a. Abschnitt 5.2.4) Nach wie vor besteht die Standardsituation unterrichtlicher Professionalität mehr oder weniger darin, dass Lehrerinnen oder Lehrer in der Rolle von Einzelkämpfern und damit isoliert mit einer mehr oder weniger großen Schülergruppe arbeiten, ohne dass weitere Fachleute Anteil am Geschehen nehmen. Auch wenn Lehrpersonen stets auch eine individuelle Lehrverantwortung inne haben, ist eine nachhaltige Unterrichtsentwicklung auf Lehrerteams, Praxisgemeinschaften – auf Kooperation als Leitprinzip – angewiesen, in deren Rahmen professionelle Selbstreflexion, soziales Feedback und darauf bezogenes Um- und Neulernen ermöglicht und aktiv unterstützt werden. Darauf zu achten ist sodann, dass Unterrichtsentwicklungsmaßnahmen hinsichtlich intendierter Wir-

kungen periodisch evaluiert werden. Vielerorts besteht bereits heute im Zuge teilautonomer Strukturen die Verpflichtung zu Selbst- und Fremdevaluation (siehe dazu Abschnitt 5.2.3). Kriteriengeleitete und auf das Handeln der Lehrpersonen abgestimmte Evaluationsinstrumente können eine wichtige Funktion im Zuge des „Feintunings“ von Entwicklungsmaßnahmen einnehmen. Hierzu ist wiederum auch die Bereitschaft der Lehrpersonen zur Kooperation von entscheidender Bedeutung. Insgesamt wird deutlich, dass eine ernsthafte und auf Nachhaltigkeit hin angelegte Unterrichtsentwicklung etlicher „Baustellen“ bedarf (vgl. Messner 2006), deren innerer Bezug zueinander für die Betroffenen jederzeit ersichtlich bleiben muss.

Ergebnisse aus vielen, oftmals lokalen und kleineren Studien, in denen Unterrichtsentwicklungsmaßnahmen evaluiert wurden, zeigen, dass Unterrichtsentwicklung eine insgesamt anspruchsvolle und vor allem langfristige Angelegenheit darstellt. Das heißt, von kurzzeitigen, auch sehr sinnvollen Maßnahmen sind kaum durchgreifende Wirkungen auf die Leistungen und das Lernverhalten von Schülerinnen und Schülern zu erwarten. In einer vergleichsweise typischen Interventionsstudie im Zuge der Einführung neuer Rahmenlehrpläne im Fach Mathematik arbeiteten Schulhausteams aus Grundschulen und Gymnasien während rund zehn Sitzungen mit videobasierten Qualitätszirkeln. Im Vergleich zu einer vergleichbaren Kontrollgruppe veränderte sich der Unterricht der Interventionsgruppe in konstruktivistischer Weise, doch blieben Effekte auf der fachlichen Leistungsebene aus. Gleichwohl wurden die Lehrpersonen, die das Programm durchlaufen hatten und in videobasierten Qualitätszirkeln gelernt hatten, ihren Unterricht zu reflektieren und kollegiale unterrichtsbezogene Feedbacks zu erteilen, von ihren Schülerinnen und Schülern als kompetenter wahrgenommen als ihr Kolleginnen und Kollegen, welche in dieser Zeit ohne entsprechende Maßnahmen blieben (Gärtner 2007). Wie etliche Beispiele in den folgenden Unterkapiteln zeigen, ist die Entwicklung guten Unterrichts nicht auf die Schnelle zu haben, sondern benötigt neben der Bereitschaft an innovativen Versuchen teilzunehmen immer auch die Ausdauer aller Akteure. Wie insbesondere die SINUS-Studie (vgl. 5.2.4) gezeigt hat, werden in der Regel gegen zwei Jahre benötigt, um stabile Arbeitsgruppen an Schulen zu etablieren und mit einer zielbezogenen kooperativen Unterrichtsentwicklung zu beginnen und erste Fortschritte zu erzielen. Grosse Qualitätsgewinne benötigen nochmals weit größere Zeiträume der kontinuierlichen Zusammenarbeit.

Bezogen auf die Einführung von Bildungsstandards werden zwei Dinge deutlich: Erstens können Implementierungsansätze nicht ausschließlich top-down erfolgen, sondern müssen von den Akteuren her gedacht und in kooperativer Weise angestoßen werden. Die Entwicklung des Unterrichtssystems beginnt dabei mit der Reflexion des praktizierten eigenen Unterrichts. An dieser Reflexion müssen sich alle mit der Bildung von Lehrpersonen befassten Institutionen beteiligen. Zweitens muss bei einer mit Standards verbundenen Outputorientierung der Prozessebene des Unterrichts die notwendige Aufmerksamkeit geschenkt werden. Die Frage muss von zentralem Interesse sein, durch welche Prozesse und unter welchen Rahmenbedingungen die erstrebten Leistungen erbracht werden können. Bildungsstandards sind immer auch als *Prozessstandards* (Opportunity-to-learn-Standards; Ravitch 1995) und damit als *Lehr-Lernstandards* zu begreifen. Auch bei einer verstärkten Output-Orientierung darf die Qualität von Lehr-Lernprozessen als Teil des pädagogischen Auftrags nicht ausgeklammert werden.

Die nachfolgend aufgelisteten Merkmale stehen für ein Kondensat dessen, was in der empirischen Unterrichtsforschung, aber auch in einer zeitgemäßen Lehrerbildung gegenwärtig unter kognitionspädagogischen Gesichtspunkten als normativ wünschbare und empirisch erhärtete Gütemerkmale eines allgemein und fachdidaktisch gehaltvollen und ertragreichen Lehrerhandelns gelten kann (Aebli 1983; Helmke 2003; Meyer 2003; Resnick/Hall 2001; Reusser 2001):

- Klare Lernzeitstrukturierung und störungspräventive Klassenführung;
- lernförderliches, als warm und ermutigend perzipiertes Unterrichtsklima;
- vielfältige Motivierung im Rahmen einer produktiven Lernaufgabekultur im Geist des Problemlösens;
- Klarheit und Kohärenz des verstehensbezogenen fachdidaktischen Strukturbaus mit Blick auf Standards des Wissens und Könnens;
- Konsolidierung, intelligentes Üben und Durcharbeiten des Lernstoffs;
- sinnstiftende Unterrichtskommunikation und Gesprächsführung;
- adaptive Instruktion und Lernbegleitung (Unterrichts als „cognitive apprenticeship“);

- Förderung von Lernstrategien, Heranführen an selbständiges und reflexives Lernen;
- kognitiv aktivierende Unterrichtsgestaltung in angemessener Methoden- und Inszenierungsvielfalt;
- Individualisierung und Differenzierung: Umgang mit heterogenen Lernvoraussetzungen;
- Verfügbarkeit von Informations- und Wissensmedien;
- klare Leistungserwartungen und Kontrollen, wechselseitiges Feedback.

Postuliert wird, dass solche Merkmale in der Verknüpfung mit Fachstandards zu Ausgangspunkten einer unterrichtsbezogenen Fortbildungsarbeit gemacht werden.

In den „Principles and Standards for School Mathematics“ des National Council of Teachers of Mathematics (NCTM) (Ferrini-Mundy/Martin 2000) formulieren die Autoren sechs Prinzipien, welche für den Mathematikunterricht auf sämtlichen Stufen Gültigkeit beanspruchen, und die sich im Prinzip auch auf andere Fächer übertragen lassen:

- Gleichheitsprinzip: Exzellenz in Mathematik bedingt Erwartungen an und intensive Unterstützung für alle Schülerinnen und Schüler.
- Curriculumprinzip: Das Curriculum muss kohärent sein und die wichtigsten mathematischen Zusammenhänge abbilden.
- Lehrprinzip: Effektiver Mathematikunterricht muss danach fragen, was Schülerinnen und Schüler wissen und lernen müssen und auf dieser Basis herausfordernde Lernprozesse anregen.
- Lernprinzip: Schülerinnen und Schüler müssen ein tieferes Mathematikverständnis entwickeln können, das an Vorwissen und Erfahrung anknüpft.
- Beurteilungsprinzip: Beurteilung soll das Lernen unterstützen und Informationen für Lernende und Lehrende erzeugen.
- Verfahrensprinzip: Neue Technologien bilden wesentliche Elemente für mathematisches Lernen.

Bildungsstandards können nicht ausschließlich von ihrem Ende – outputbezogenen Tests – her gedacht werden. Sie müssen in engen Zusammenhang gestellt werden mit Maßnahmen, welche das Prozessgeschehen des Unterrichts nachhaltig beeinflussen. Entsprechende Maßnahmen benötigen nicht nur die Bereitschaft und das Einverständnis

der betroffenen Akteure, sondern auch deren verbindliche Kooperation im Sinne unterrichtsbezogener Schulentwicklung.

5.3.2 Lehrmittel und Aufgabenkulturen

Lehrpersonen, die sich an Bildungsstandards orientieren, werden im Unterricht nicht direkt mit diesen Standards arbeiten. Lernaufgaben und Lehrmittel stellen in der Didaktik seit jeher die häufigsten Materialisierungen von Lernzielen, Ergebniserwartungen – und neu: Bildungsstandards – im Unterricht dar. Es gibt eine ganze Reihe von Fachkulturen, darunter Mathematik und Naturwissenschaften, aber auch Geschichte und Geografie, in denen Lehrmittel und insbesondere Lernaufgaben als Bausteine, Werkzeuge und als didaktisches Mittel der (problemorientierten) Unterrichtsführung eine lange Tradition haben und eine zentrale Rolle spielen. Gute Aufgaben und Lernaufträge fordern das eigenständige Lernen der Schülerinnen und Schüler heraus und lenken es in produktive fachliche Bahnen. *Gute fachliche Lernaufgaben materialisieren jene Wissens- und Könnenskomponenten, lösen jene Denk- und Arbeitsprozesse aus und aktivieren jene analytischen und synthetischen Figuren des Problemlösens, Argumentierens, Betrachtens und Deutens, um die es in einem bestimmten Fach im Kern geht und die dessen intellektuelle Kultur ausmachen.* Offiziellen und inoffiziellen Lehrmitteln entnommene oder mündlich präsentierte Aufgabenstellungen bestimmen nicht selten (namentlich in der Mathematik, in der gegen 80% des Unterrichts aus Aufgabenlösen bestehen; vgl. Reusser/Pauli 2003) den Unterricht in allen seinen Phasen und damit das potenzielle Niveau der *kognitiven Aktivierung* und *Verknüpfungstätigkeit* und folglich des *möglichen Kompetenzerwerbs* bei den Lernenden.

Stärker als Lehrpläne bilden *Lehrmittel* das Rückgrat des Unterrichts und üben einen nachhaltigen Einfluss auf die Modellierung einer fachspezifischen Lern- und Arbeitskultur aus. Sie stehen für den „tatsächlichen“ Lehrplan, vor allem, wenn sie wie etwa in großen Teilen der Schweiz für bestimmte Fächer verbindlich vorgeschrieben sind. Gerade auf den Stufen der Volksschule, wo Lehrpersonen mehrere Fächer zu unterrichten haben, repräsentieren Lehrmittel den Kanon des zu vermittelnden domänenspezifische Basiswissen und sichern den Anschluss an fachdidaktische Stoffaufbaukulturen. Lehrmittel entsprechen dem Bedürfnis der Lehrpersonen nach fundiertem und adressatengerechtem Wissen, vor allem aber leisten sie eine für die Lehrpersonen notwendige didaktisierte Stoffreduktion (Bertschy 2004). Da Lehrbücher tendenziell

von einer Position des Wissens und nicht von einer solchen des Lernens her verfasst sind, fehlt ihnen oft der „Geist des Problemlösens“. Den Lehrpersonen obliegt es, den „fertigen Stoff“ zu verflüssigen und in genetisch-konstruktivistisches Lehr-Lernhandeln umzusetzen (Aebli 1961; Wagenschein 1968). Bildungsstandards können dabei helfen, die Erwartungshorizonte an dadurch zu erreichende Kompetenzen des Wissens und Könnens präziser und verbindlicher zu kommunizieren.

Obwohl Lehrmittel im pädagogischen Alltag eine zentrale Rolle spielen, existiert wenig Forschung zu ihrem tatsächlichen Gebrauch und zu ihrer Wirksamkeit. Dabei ist auch von Interesse, welche Merkmale von Unterrichtsmaterialien von Lehrpersonen geschätzt werden. Adamina (2004) etwa nennt die Wichtigkeit einer hohen Strukturierung und Aufbereitung, weiter das Vorhandensein möglichst konkreter Aufgabenstellungen und Einsatzmöglichkeiten für selbständiges Arbeiten. Zudem sollen Lehrbuchtexte optimal verständlich sein und durch Bilder und grafische Darstellungen ergänzt sein. Die Unterlagen sollten sowohl variantenreich im Unterricht einsetzbar und möglichst offen konzipiert, aber auch strukturiert sein und den Zielen des Lehrplans entsprechen. Bei der Betrachtung des Angebots verschiedener staatlicher und privater Anbieter fällt zudem auf, dass neuere Lernmaterialien sich in zunehmendem Maße an den zur Verfügung stehenden elektronischen Möglichkeiten orientieren und Kindern und Jugendlichen insofern entgegenkommen, indem sie virtuelle Spiel- und Erlebniswelten adaptieren, um auf diese Weise Fachwissen zu vermitteln. Kritisch gefragt werden muss hier, inwieweit ein explodierender Lehrmittelmarkt überhaupt wünschenswert ist und ob nicht zumindest für selektionsrelevante Fächer verbindliche Lehrmittel die Regel sein sollten.

Lehrmittel werden also eine bedeutsame Rolle spielen, wenn Unterricht reformiert wird. Sie eignen sich für die Vermittlung fachdidaktischer Zugänge und Ergebniserwartungen in einer Domäne. Allerdings ist zu beachten, dass Lehrpersonen bei der Benutzung von Lehrmitteln mit deren fachdidaktischen Kernideen vertraut sein müssen (vgl. Wittmann 2004) – es wäre wiederum fahrlässig, einen automatischen Effekt von der Verwendung eines bestimmten Lehrwerks auf die Lernleistungen der Schülerinnen und Schüler zu postulieren.

Einen Ansatz in diese Richtung postulieren Blum et al. (2006), indem sie nach einer anschaulichen Beschreibung zentraler mathematischer Kompetenzen Aspekte kompetenzorientierten Mathematikunterrichtes benennen und im dritten Teil ihres Buches Beispiele kompetenzorientierter Mathematikaufgaben präsentieren. Dabei spielen die

Variationen verschiedener Aufgaben, multiple Lösungswege und Realitätsbezüge wichtige Rollen. Konkret werden sechs mathematische Kompetenzen genannt und in der Folge detailliert beschrieben (mathematisch Argumentieren, Problemlösen, mathematisches Modellieren, mathematische Darstellungen verwenden, mit symbolischen und formalen Elementen der Mathematik umgehen sowie mathematisch Kommunizieren). Lehrpersonen erhalten damit eine handlungsnah Zielorientierung und ein Werkzeug, das ihnen helfen kann, mathematische Grunderfahrungen zu vermitteln und zu reflektieren. Die Rückverflüssigung von Standards in sinnvolle und nachhaltige Lernaktivitäten kann mit der Einführung solcher innovativer Lehrmittel unterstützt werden; Lehrpersonen lassen sich für fachdidaktische Anliegen gewinnen, wenn ihnen auf pragmatische Weise Unterstützung in ihrer Kern-tätigkeit angeboten werden kann.

TIMSS und danach PISA haben deutlich gemacht, dass z.B. deutsche Schülerinnen und Schüler mit dem Lösen von Routineaufgaben recht gut zurechtkommen, jedoch erhebliche Schwierigkeiten zeigen mit Aufgaben, bei denen es um Anwendung, sinnvolle Übertragung und Transformation des Gelernten geht. Weiter wurde u.a. in der *TIMSS 1999 Videostudie* deutlich, dass elementarste Routineaufgaben *international* den überwiegenden Teil der Lernzeit im Mathematikunterricht beanspruchen und dass dort, wo Lernaufgaben das Potenzial zu „höherwertigem“, über prozedurales Routinehandeln hinausgehendem Denken hätten, dieses Potenzial häufig unerkannt und ungenutzt bleibt (Hiebert/Gallimore/Garnier/Givvin/Hollingsworth/Jacobs 2003; Pauli/Reusser 2006). Über die Aufgabenbearbeitungskultur in einem Fach nachzudenken und diese zu verbessern – durch die Orientierung von Aufgaben an gestuften fachlichen Kompetenzvorstellungen –, muss deshalb als produktiver Impuls für eine an Kriterien des anspruchsvollen fachlichen Lernens bzw. an Bildungsstandards ausgerichtete Unterrichtsentwicklung gesehen werden.

Im deutschsprachigen Raum ist vor allen seit dem Modellversuchsprogramm SINUS die Bedeutung von Aufgabenkulturen erkannt und ihre Weiterentwicklung, eingebettet in ein Gesamtkonzept der kooperativen Unterrichtsentwicklung, in Fachgruppen und Netzwerken postuliert worden. In einem eigens zu diesem Thema konzipierten Modul konnten die Lehrpersonen lernen, wie sie abwechslungsreiche Anwendungs- und Übungsaufgaben in variierenden Kontexten verwenden und auf diese Weise zur Vernetzung und Flexibilisierung von Wissen beitra-

gen können. Ziel ist es, vermehrt Aufgaben zu entwickeln, zu erproben und einzusetzen, die „Verstehen fördern und fordern, realistische Anwendungen betreffen, mehrere Lösungswege zulassen oder systematisch früheren Stoff aufgreifen und wiederholen. Zugleich sollen die Schülerinnen und Schüler bei der Lösung von Aufgaben begleitet, unterstützt und angeleitet werden“ (Prenzel 2005, S. 261). Im Anschluss an SINUS und an die weiterführenden Entwicklungsarbeiten bei PISA werden mittlerweile kompetenzorientierte Aufgaben in den verschiedensten, mit Bildungsstandards kompatiblen oder verknüpften Funktionen eingesetzt: als Lern- und Prüfungsaufgaben für den Unterricht und darauf bezogene Klassenarbeiten, als Diagnoseaufgaben für Lernstandserhebungen und Vergleichsarbeiten und als Testaufgaben für Schulleistungsprüfungen und Evaluationen im Anschluss an TIMSS und PISA.

In Bezug auf die Etablierung sinnvoller Aufgabenkulturen sind wiederum Lehrmittel von nicht zu unterschätzender Bedeutung. Neben der Darstellung fachdidaktischer Bezüge können Lehrpersonen konkrete Hilfestellungen geboten werden, indem standardkompatible (Muster-) Beispiele vorgestellt werden. In der erwähnten Arbeit von Blum et al. (2006) wird dabei u.a. auf die Notwendigkeit intelligenten Übens, auf Projektorientierung und Realitätsbezüge sowie auf die Notwendigkeit von Variation und multiplen Lösungswegen hingewiesen. Letztlich geht es dabei um einen langfristigen Kompetenzaufbau, nicht nur innerhalb eines Schuljahres, sondern über mehrere Klassenstufen hinweg. Dies kann aber nur gelingen, wenn den anwendenden Lehrpersonen Typologie und Kompetenzstrukturen der Aufgaben erschlossen sind. Das heißt, dass Lehrpersonen nicht nur die Tiefenstrukturen kompetenzorientierter Aufbauprozesse insgesamt, sondern auch die inhärenten Kompetenzfacetten jeder einzelnen Aufgabe verstehen – mithin zu kognitiven Aufgabenanalysen fähig sein – müssen. Nicht unterschätzt werden darf dabei der Aufwand, der für den Prozess der Aufgabenentwicklung zu leisten ist. Standardkompatible Aufgabenkollektionen müssen in Zusammenarbeit mit Experten aus Fachdidaktik und Fachwissenschaft entwickelt, bewertet, klassifiziert und gegebenenfalls überarbeitet werden (vgl. Ehmke et al. 2006).

Auf der Unterrichtsebene können fachliche Standards durch Aufgabenbeispiele und Wege ihrer Bearbeitung veranschaulicht werden. Gute Beispiele sollen Lehrpersonen im Detail verdeutlichen, wie Kompetenzen unterschiedlicher Stufung mit spezifischen, aufgabenbezogenen Bearbeitungs- und Problemlöseprozessen zusammenhängen. Sodann sollen sie es Lehrpersonen erlauben, über verschiedene Kompetenzstu-

fen und Bearbeitungsniveaus und über unterschiedliche Lösungswege von Aufgaben nachzudenken sowie Lernprozesse über Aufgaben unterschiedlicher Komplexität – und damit abgestimmt auf unterschiedliche Fähigkeitsniveaus – zu organisieren. Wenn Aufgaben einem nachhaltigen Kompetenzaufbau dienen sollen, müssen sie ferner Teil einer Lernumgebung sein, welche hilft, neue Themen zu erschließen, langfristige Arbeitsplanungen vorzunehmen, und individuelles und produktives Üben sowie komplexes Anwenden zu ermöglichen (Blum et al. 2006).

Sich zu vergegenwärtigen, welche Anforderungen mit einer Aufgabe verbunden sind, ist eine wichtige Voraussetzung für eine fachdidaktisch produktive und pädagogisch-lerndiagnostisch sensible Begleitung und Unterstützung von Aufgabenbearbeitungsprozessen, insbesondere bei schwächeren Schülerinnen und Schülern.

5.3.3 Abrufbare Fach- und Leistungstests

Ob ein Bildungsmonitoring oder eine Lernstandserhebung günstige Auswirkungen auf die konkrete Unterrichtsarbeit hat oder nicht, hängt nicht zuletzt davon ab, ob Rückmeldungen für die Lehrpersonen produktive Nutzungsmöglichkeiten bieten. Unterrichtende sind auf Daten angewiesen, die ihnen bei der Diagnose von Lernleistungen – bezogen auf die Klasse und auf individuelle Schülerinnen und Schüler – dienlich sind. Dabei ist vor allem an die Erweiterung einer klassenintern notgedrungen meist sozialen Bezugsnormorientierung zu denken: Für Lehrkräfte bestanden bis vor wenigen Jahren kaum valide Vergleichsmöglichkeiten der Klassenarbeiten ihrer Schülerinnen und Schüler, weil kriteriale Tests fehlten, die sich einerseits an fachlichen Leistungskriterien orientieren, andererseits die Chance bieten, die erreichten Ergebnisse mit denjenigen einer schulübergreifenden Stichprobe zu vergleichen. Entsprechende Möglichkeiten ergeben sich heute mit Vergleichs- respektive Parallel- oder Orientierungsarbeiten, wie sie in etlichen deutschen Bundesländern bereits zur Anwendung kommen. Unter den unterschiedlichen Begriffen wird dabei Ähnliches verstanden. Die Kernidee besteht darin, den Leistungsstand von Lernenden anhand von Fachtests zu prüfen und in Beziehung zum Leistungsniveau einer in Bezug auf Eingangsvoraussetzungen und Rahmenbedingungen vergleichbaren Altersgruppe zu setzen.

In Hessen etwa absolvieren sämtliche Schülerinnen und Schüler in der zweiten Hälfte der zweiten und der dritten Klasse eine Orientie-

rungsarbeit in den Fächern Deutsch und Mathematik. Die Aufgabenstellungen entsprechen den Bildungsstandards der beiden Fächer. Neben der absoluten Leistungsbestimmung erlauben die Arbeiten die Analyse von Stärken und Schwächen in einzelnen fachlichen Teilgebieten. Die als Diagnoseinstrumente konzipierten Tests sollen dabei nicht für die Benotungen eingesetzt werden. Die Gesamtergebnisse der teilnehmenden Klassen werden auf der Website des Kultusministeriums veröffentlicht.¹⁹⁸

In ähnlicher Weise werden in Baden-Württemberg verpflichtende Diagnose- und Vergleichsarbeiten in der zweiten, sechsten und achten Klasse in verschiedenen Fächern durchgeführt. Auch hier sollen landesweite Vergleichswerte erzeugt werden, an denen die Lehrpersonen den durchschnittlichen Leistungsstand ihrer Klasse messen können. Angestrebt wird damit eine zielgerichtete und systematische Qualitätsentwicklung vor Ort.¹⁹⁹

Während in Deutschland und anderen Ländern eine gegenseitige Abstimmung von Orientierungsarbeiten mit den bereits eingeführten Bildungsstandards festzustellen ist, orientieren sich die in der Schweiz verwendeten Instrumente an Aufgabensammlungen, welche auf der Grundlage der bestehenden Lehrpläne entwickelt worden sind. Im Unterschied zu den zitierten deutschen Beispielen können die Schweizer Instrumente zum Teil von den Lehrpersonen selbst abgerufen werden.

Ganz allgemein ist zu vermuten, dass mit der Etablierung einer verstärkt wirkungsorientierten Evaluationskultur standardisierte Schulleistungs- und Fachtests, die von Lehrpersonen auch selber eingesetzt werden können, insgesamt zunehmend an Bedeutung gewinnen werden.

Ein Instrument ist das im Kanton St. Gallen erprobte und in den Kantonen Graubünden und Zürich eingesetzte *Klassencockpit*. Es handelt sich um ein modular aufgebautes Internet-Angebot für die Bestimmung des Leistungsstandes einer Klasse in Deutsch und Mathematik. Die mittels einer repräsentativen Basisstichprobe standardisierten Leistungstests werden von den Lehrkräften korrigiert und über eine Online-Datenbank ausgewertet. Die Datenbank enthält die Daten anderer Klassen, die einen kantonalen Durchschnitt bilden, so dass ein Vergleich der Leistungen möglich wird. Die Vergleichsdaten können klas-

¹⁹⁸ Vgl. dazu: http://www.kultusministerium.hessen.de/irj/HKM_Internet?cid=55bce4ece72a32c66452d270d824f900 (Stand 15.09.06).

¹⁹⁹ Vgl. dazu: <http://www.dva-bw.de/> (Stand 15.09.06).

senspezifisch wie schülerspezifisch abgerufen werden, die Eingabe ist anonym. Bearbeitet werden Testaufgaben aus den beiden Promotionsfächern Mathematik und Deutsch. Eine externe Evaluation ergab eine hohe Akzeptanz bei den Lehrpersonen, nicht zuletzt deshalb, weil auf diese Weise ein valides und nicht für die schulische Selektion vorgesehenes Vergleichsinstrument bereitsteht. Kritisiert wird vor allem der noch restriktive Einsatz des Instruments. In Befragungen zeigt sich, dass die Lehrerinnen und Lehrer den Test und seine Auswertung nicht nur zur Standortbestimmung nutzen wollen, sondern auch zur Anpassung der Schülerbeurteilung, zur Planung der Übertritte von der Primar- in die Sekundarschule und nicht zuletzt zur Optimierung des eigenen Unterrichts (Moser 2003b). Allerdings ist auch zu bemängeln, dass Interpretationshilfen für die Lehrpersonen fehlen (Schneewind 2006).

Im Kanton Aargau wird mit dem Projekt *Check 5* ebenfalls ein Instrument zur unabhängigen Standortbestimmung der Schülerinnen und Schüler eingesetzt. Neben der Erfassung fachlicher Kompetenzen in Deutsch und Mathematik wird der Fokus auch auf überfachliche Fertigkeiten (kooperatives Problemlösen, selbstreguliertes Lernen) sowie auf die Wahrnehmung von Unterrichtsaspekten durch die Lernenden gerichtet. Als übergeordnete Ziele werden die Verbesserung der Unterrichtsqualität und die Förderung der Schülerinnen und Schüler genannt. Rückmeldungen an die Lehrpersonen werden in anonymisierter Form vorgenommen. Auf deren Grundlage sind die teilnehmenden Lehrpersonen verpflichtet, anhand eines vorgegebenen Analyserasters Maßnahmen zur Unterrichtsentwicklung zu skizzieren und sie den übrigen Lehrerinnen und Lehrern auf einem Internetportal zur Verfügung zu stellen. Der Einsatz dieses für fünfte Klassen vorgesehenen Instruments ist freiwillig. Im Jahr 2005 nahmen 306 Lehrpersonen mit mehr als 5300 Schülerinnen und Schülern an *Check 5* teil. Somit sind 80 Prozent aller Aargauer Klassen an dieser Leistungsmessung beteiligt (Kanton Aargau, Departement Bildung, Kultur 2006; Moser/Keller 2005). Der Rahmen eines standardisierten Fachtests wird damit bedeutsam erweitert, indem auch die Diagnostik und Weiterentwicklung von Unterricht thematisiert werden. Unklar erscheint allerdings, in welche Richtung eine solche Weiterentwicklung erfolgen soll. Die Lehrpersonen evaluieren ihre getroffenen Maßnahmen selbst, eine externe Evaluation, welche Veränderungen feststellen könnte, existiert bislang nicht (Schneewind 2006).

Um der schleichenden Entwertung des Volksschulabschlusses durch firmeneigene Assessments entgegenzutreten, bietet das Projekt *Stellwerke*

einen Online-Test an, der es Schulabgängern erlaubt, sich bezüglich der an verschiedenen Lehrstellen verlangten Kompetenzen einzuschätzen (Moser 2004). Mit Hilfe einer Analyse des Leistungsstandes jedes einzelnen Schülers soll ermittelt werden, wie und in welchen Bereichen die Schülerinnen und Schüler bis zum Ende des 9. Schuljahres gezielt gefördert werden können, um das für eine Lehrstelle geforderte Niveau erreichen zu können. Dieses Förderprogramm ersetzt in Teilen den Lehrplan, der zugunsten von gezielten Nachbesserungen in bestimmten Kompetenzbereichen gelockert oder suspendiert wird. Die Probandinnen und Probanden bearbeiten online Aufgaben aus den Themenbereichen Deutsch, Französisch, Englisch und Vorstellungsvermögen. Je nach Fähigkeiten können leichte, mittlere und schwierige Fachtests gelöst werden. Dabei werden unterschiedliche Aufgabentypen eingesetzt (Eingabefelder, Multiple-Choice, Lückentext, Aufzählung, Hotspot, Zuordnung, Audioaufgaben, aber auch Flash-basierte Aufgaben) (Corray/Graf 2004). Die erbrachten Leistungen werden in einem Profil dargestellt.

Auch wenn sich die geschilderten Schweizer Instrumente infolge fehlender Entwicklungsarbeiten (noch) nicht auf Bildungsstandards im neueren und präzisen Wortsinne beziehen, vermitteln sie einen Eindruck von den Möglichkeiten abrufbarer Fachtests und den Bedürfnissen und möglichen Zielen, welche die Nutzer schon jetzt damit verbinden. Und sie pilotieren Wege, die auch für den adaptiven Gebrauch von an Bildungsstandards orientierten Fachtests im Klassenzimmer in Zukunft begangen werden können.

Die Evaluation der einzelnen Instrumente verweist auf eine insgesamt gute Akzeptanz durch die Lehrpersonen, welche das Informationspotenzial von standardisierten Fachtests als wertvolle Unterstützung für die Leistungsdiagnostik einschätzen. Standardisierte Fach- und Leistungstests können, wenn sie fachdidaktisch gut konstruiert sind, auch ein hilfreiches Werkzeug beim Wechsel von einer sozialen, auf das Leistungsniveau der jeweiligen Klasse bezogenen, hin zu einer sachlich-kriterialen Bezugsnormorientierung sein, indem sie das Erreichen fachlicher Kompetenzen durch Lernende in den Vordergrund rücken. Allerdings wird der Nutzen trotz der breiten Akzeptanz von Ergebnismeldungen bisher als eher gering eingeschätzt (vgl. Schneewind 2006). Offensichtlich fällt es den Lehrpersonen sehr schwer, die vielfältigen Rückmeldungen wieder in produktivere Lehr-Lernprozesse in ihren eigenen Klassen zu übersetzen. Lehrpersonen müssten die Ergebnisse der Tests aber nutzen können, um sie als hilfreiche Diagnose ihrer

eigenen und der Arbeit ihrer Schülerinnen und Schüler einsetzen zu können. Dazu wären eine Ausrichtung der Aufgaben an klareren Kriterien, die den Lehrpersonen bekannt sind, sowie Interpretationshilfen für die Lehrkräfte notwendig. Eine detaillierte Analyse der Nutzung der Ergebnisse von Check 5 durch die Lehrpersonen (Tresch 2007) ergab denn auch als häufigsten Beweggrund zur Teilnahme die Möglichkeit, die Leistungen der eigenen mit denjenigen anderer Klassen vergleichen zu können. In bedeutend geringerem Maße wird die Erreichung der Lehrplanziele genannt. Die Tests werden vor allem deshalb geschätzt, weil sie eine Reflexion des Unterrichts und die Leistungsbeurteilung der Schülerinnen und Schüler ermöglichen. Die Grenzen des Potenzials von Leistungstests werden in der fehlenden Abstimmung auf Unterrichtsinhalte und in der Verschiebung des Augenmerks ausschließlich auf messbare Ziele gesehen. Als Konsequenz für zukünftige Schulleistungsmessungen werden die Abstimmung auf curriculare Ziele, aussagekräftige Ergebnisrückmeldungen, die Ermöglichung externer Unterstützungsangebote für die Lehrpersonen, eine sorgfältige Balance zwischen verpflichtender Teilnahme und genügenden Freiräumen im Umgang mit den Testresultaten sowie eine transparente Kommunikation bezüglich des Instrumenteneinsatzes genannt.

Indem sie auf ganz bestimmte Aspekte von Lernergebnissen fokussieren, vermögen Leistungstests zudem nur ein eingeschränktes Segment des vielschichtigen Lehr-Lerngeschehens abzubilden. Um andere Aspekte und insbesondere „weichere“ Prozesserträge in den Blick zu nehmen, bedarf es der komplementären Ergänzung durch formative und förderorientierte Diagnosen. In je größerer Distanz vom Lerngeschehen in der Klasse Leistungstests ausgearbeitet werden, desto weniger sind sie für die Diagnose von individuellen Lernständen nutzbar. Um individuelle Förderdiagnostik zu betreiben, braucht es feiner auf das konkrete Lerngeschehen abgestufte Verfahren als die der Statusdiagnostik dienenden Tests. Abrufbare, auf Bildungsstandards abgestützte Fachtests könnten hier vermutlich einen Beitrag leisten, wenn Lehrpersonen sie produktiv einzusetzen in der Lage sind. Unter dem Begriff „Standards von unten“ schlägt Winter (2004, 2005) darüber hinaus Formen der individuellen Rechenschaftslegung vor, welche neben Prüfungen und Tests Portfolios, Lerntagebücher und gemeinsame Präsentationen in die Beurteilung einbeziehen, um das Spektrum an Informationen für alle am Lehr- und Lerngeschehen Beteiligten zu erweitern. Dies erscheint insofern notwendig, als im Rahmen eines erweiterten

Lehr-Lern-Begriffs selbsttätiges Lernen und damit verbunden der Gebrauch von Lernstrategien und Arbeitstechniken an Bedeutung zunehmen. Selbsttätig Lernende bleiben jedoch in Bezug auf ihr Lernen auf die Anleitung, Unterstützung und Beratung durch ihre Lehrkräfte angewiesen. Diese Beratung wird umso besser ausfallen, je differenzierter und aspektreicher die Feedback-Grundlagen zum Lerngeschehen sind, auf die zurückgegriffen werden kann.

Standardisierte, abrufbare Fach- und Leistungstests müssen auf fachdidaktische valide Weise erfassen, was Schülerinnen und Schüler können sollen. Sie sind entlang von eindeutigen und allen Betroffenen bekannten Leistungskriterien zu formulieren. Im Rahmen der Ausrichtung des Lernens auf Bildungsstandards stellen sie dann einen Mehrwert dar, wenn sie differenzierte und klar interpretierbare Leistungsrückmeldungen ermöglichen und förderdiagnostisch einsetzbar sind. Sie stellen jedoch keinen vollständigen Ersatz einer am Klassenlerngeschehen und an individuellen Lernprozessen einzelner Schülerinnen und Schüler ausgerichteten Diagnostik dar, sondern sind vielmehr eine notwendige Ergänzung.

5.3.4 Fachspezifisch-pädagogisches Coaching

Wenn bisher Aufgaben- und Testkulturen beschrieben worden sind, die auf eine nachhaltige Implementation von Standards auf der Ebene des Schülerlernens zielen, geht es in den folgenden beiden Abschnitten um Instrumente, die sich direkt an das Handeln und die unterrichtsbezogene Lehrerprofessionalität richten. Lehrpersonen sind neben den Schülerinnen und Schülern die Hauptakteure, deren professionelle Kompetenzen durch *innovative Verfahren und Instrumente der Fortbildung* weiterentwickelt werden müssen. Traditionelle Verfahren der Fort- und Weiterbildung, die meist als reine Wissensvermittlung und Belehrung ohne Nachbetreuung gestaltet sind, wurden in den letzten Jahren dahingehend kritisiert, dass sie häufig praxisfern erfolgen, wenig individualisiert und situiert angelegt sind und deshalb ziemlich wirkungslos bleiben (Hawley/Valli 1999; Messner/Reusser 2000; Staub 2004). Unter den neueren Verfahren der Personalentwicklung (u.a. Hospitationen, Lehrer-Tandems, Formen der Supervision und der kollegialen Beratung, Peer-Coaching, Arbeitsgruppen, Lernzirkeln), welche direkt auf die Unterrichtsebene zielen, gehört das *Coaching des pädagogisch-fachdidaktischen Handelns* von Lehrpersonen zu den vielversprechenden Ansätzen (Staub 2001; West/Staub 2003). Dabei geht es darum, Lehrpersonen zeitlich

begrenzt, fachbezogen und „on the job“ in der Planung, Durchführung und Auswertung (Reflexion) ihres Unterrichtshandelns im Hinblick auf erfolgreiche Lernprozesse ihrer Schülerinnen und Schüler zu unterstützen.

Die Erwartung an Lehrpersonen, als reflektierende Praktiker ihr professionelles Handeln kontinuierlich zu überdenken und zu reflektieren, ist nicht neu (vgl. Schön 1987). Lehrpersonen verfügen in Bezug auf die Grundlagen ihres professionellen Handelns dazu auch über verschiedene Arten von Wissen. Darunter erweist sich seit längerem vor allem fachspezifisch-pädagogisches Wissen als bedeutsam (Shulman 1986, 1987; Bromme 1995; Staub 2001). Es handelt sich dabei um eine Verschmelzung von Wissen verschiedener Herkunft mit der eigenen Erfahrung und setzt sich aus disziplinärem Fachwissen, allgemein pädagogisch-didaktischem und psychologisch-didaktischem Wissen und Erfahrungen von Anwendungen in konkreten Handlungssituationen zusammen (Staub 2004). Da es in weiten Teilen habituiert ist und eng mit im Laufe der eigenen pädagogischen Laufbahn erworbenen und gefestigten Überzeugungen (beliefs) zusammenhängt (vgl. Dann 1983, 1994), ist es nur schwer zugänglich und veränderbar. Es handelt sich im Sinne von Polanyi (1985) um „implizites“ (tacit) Wissen, das durch herkömmliche, „akademisch“ organisierte Fortbildungen in den Formen traditioneller Wissensvermittlung kaum verändert wird. Da eine hohe Konsistenz zwischen den impliziten, subjektiven Theorien (den wissensmäßigen Grundlagen des Lehrens und Lernens) und dem Handeln von Lehrpersonen (vgl. Dann 1994; Wahl 2001) besteht und ebenfalls Zusammenhänge zum Schülerlernen nachweisbar sind (Staub/Stern 2002), ist dies der Punkt, an dem innovative Fortbildungen vor allem ansetzen müssen. Soll die Implementation von Bildungsstandards tatsächlich die Unterrichtsebene erreichen, so muss schwerpunktmäßig in innovative Formen der Fortbildung von Lehrpersonen investiert werden.

In dem von Staub im Kontext des amerikanischen „Institute for Learning“²⁰⁰ entwickelten innovativen Coaching-Modell (Staub 2001, 2004; West/Staub 2003) beteiligt sich ein im Unterrichtsfach kompetenter Coach – in der Regel handelt es sich um erfolgreiche (ehemalige) Lehrkräfte und/oder erfahrene Weiterbilderinnen und Weiterbilder, die

²⁰⁰ Vgl. <http://www.instituteforlearning.org/about.html> (Stand 16.09.06).

für ihre Aufgabe ausgebildet werden – gemeinsam mit einer (Gruppe von) Lehrperson(en) an der Planung, Durchführung und Auswertung von Unterricht. Der Coach ist dabei nicht bloß ein beobachtender Prozessberater mit (allgemeinen) Reflexionskompetenzen, sondern er übernimmt in der Rolle einer in allen Belangen der Arbeit aktiv mitwirkenden Fachperson *Mitverantwortung* für die Gestaltung der gecoachten Unterrichtseinheiten und für das durch den Unterricht erzeugte Lernen der Schülerinnen und Schüler. Der Ansatz hebt sich somit ab von allgemeinen Coaching-Modellen, in denen der Coach „primär als *Prozessberater* und nicht als direkter Problemlöser“ agiert (Rauen 1999, S. 30; zit. aus Staub 2004, S. 130). Zudem wickelt sich das Coaching innerhalb des authentischen beruflichen Handelns ab, was beim herkömmlichen Prozess-Coaching in der Regel nicht der Fall ist.

Staub umschreibt das Verfahren des „Content-Focused Coaching“SM wie folgt (2004):

„Das Coaching orientiert sich an allgemeinen Lehr-Lernprinzipien sowie Kernperspektiven für die Gestaltung und Reflexion von Lektionsdesigns. Das unmittelbare Ziel besteht in der gemeinsamen Gestaltung eines für das Lernen der Schüler möglichst optimalen Unterrichts. Das langfristige Ziel Fachspezifisch-Pädagogischen Coachings ist die nachhaltige Entwicklung von Unterrichtskompetenz. Dabei sollen insbesondere allgemein-didaktische *Planungs- und Reflexionsstrategien* eingeübt und *fachspezifisch-pädagogisches Wissen* erworben werden“ (S. 124).

Staub entwickelt ein Konzept schulinterner Weiterbildung und macht Vorschläge auf verschiedenen Ebenen (ebd.):

„Mit Bezug auf die Schulen geht es um die *organisatorisch-institutionelle* Ermöglichung und Gestaltung von *Settings* für die Zusammenarbeit von fachkompetenten Coachs und Lehrpersonen. Die in solchen Settings entstehenden Lerngelegenheiten werden wesentlich dadurch bestimmt, worüber Coach und Lehrperson kommunizieren. Bezüglich der *Inhalte* von Coaching-Gesprächen macht Fachspezifisch-Pädagogisches Coaching Vorschläge dazu, welche pädagogisch-psychologischen und psychologisch-didaktischen Perspektiven in den Besprechungen von Unterricht besonders zu berücksichtigen sind. Ein wesentliches Merkmal dieses Coaching-Ansatzes besteht weiter darin, dass Coaching-Gespräche über Fragen der Unterrichtsgestaltung in einem hohen Maße als *inhaltspezifische Thematisierungen von Lehr-Lernprozessen* erfolgen. Zur Her-

ausbildung einer entsprechend fokussierten Gesprächskultur wird die Arbeit mit konzeptuellen Werkzeugen in der Form von flexibel zu handhabenden Leitfäden vorgeschlagen. Das durch die Gespräche zwischen Coach und Lehrperson(en) geförderte Lernen wird weiter von der Art und Weise beeinflusst, wie Inhalte aufgegriffen und verhandelt werden. Dies hängt sowohl mit dem Rollenverständnis der Akteure wie auch mit der Qualität der Gesprächsführung zusammen. Fachspezifisch-Pädagogisches Coaching strebt eine möglichst *dialogische Gestaltung* von Coaching-Interaktionen an. Das Coaching findet in der Regel verteilt über mindestens ein ganzes Schuljahr alle ein bis zwei Wochen statt. Es wird jedoch auch so gearbeitet, dass Coachs in Blöcken jeweils mehrere Tage sehr intensiv mit bestimmten Lehrpersonen zusammenarbeiten und zwischen den Blöcken längere Phasen eigenständiger Arbeit liegen“ (ebd. S. 124).

Auf der konkreten Handlungsebene besteht der Coaching-Prozess aus drei konsekutiven Phasen oder Elementen. In der ersten Phase, bei der es um die *Unterrichtsplanung* geht, stellen Coach und Lehrperson beispielsweise Überlegungen dazu an,

- welcher Art das beabsichtigte Lernen ist,
- welches die zu erreichenden Lernziele und Standards sind,
- an welches Vorwissen der Lernenden angeknüpft werden kann,
- welche Schwierigkeiten zu erwarten sind und
- welche Gelegenheiten den Schülerinnen und Schülern geboten werden könnten, ihr Denken offen zu legen (Staub 2006).

In der sich an die Vorbesprechung anschließenden *Unterrichtsphase* übernimmt der Coach sodann Mitverantwortung am Lernen der Schülerinnen und Schüler, indem er sich bedarfsweise und nach Absprache mit dem Coachee (der Klassen- oder Fachlehrperson) aktiv an Unterrichtsgesprächen und einzelnen Elementen der Lektion beteiligt. Die darauf folgende *Unterrichtsnachbesprechung* dient dem „Coaching der didaktischen Reflexion“ (Eggenberger/Staub 2001), in deren Rahmen auf Unterrichtsepisoden Bezug genommen wird, wiederum primär mit Fokus auf die Lernprozesse der Schülerinnen und Schüler.

Das Konzept wurde bisher vor allem im Bereich der Grundschulmathematik in einem großen Schulkreis von New York City (West/ Staub 2003) und in der amerikanischen Lehrerfortbildung in Schuldistrikten in

Providence (RI), Los Angeles (CA), Denver (CO) und Austin (TX) im Englischunterricht über mehrere Jahre eingesetzt. In jüngerer Zeit findet es auch in der schweizerischen Lehrer(fort)bildung zunehmend Beachtung.

Es muss kaum besonders betont werden, wie sehr sich ein solches unterrichtsbezogenes Coaching-Verfahren eignen kann, um Unterricht unter Qualitätsgesichtspunkten, wie sie Bildungsstandards im engeren (als leistungsbezogene Ergebnisstandards) und im weiteren Sinne (als Prozess- und Input-Standards) darstellen, weiterzuentwickeln.

Wichtig ist, dass Coaching als ein Verfahren der lehrpersonen- und lehr-lernprozessbezogenen Unterrichtsentwicklung nicht primär als remediales Werkzeug verstanden wird. Auch wenn es seine Funktion dort, wo Standards von den Lernenden nur unzureichend oder nicht erfüllt werden, sicherlich in besonderer Weise entfalten kann, sollten alle Lehrpersonen davon im Sinne eines Instrumentes der Entwicklung eines an den Lernprozessen der Schülerinnen und Schüler ausgerichteten Unterrichts profitieren können.

Unterrichtsbezogenes Coaching sollte demnach in der Funktion eines Privilegs und nicht als unerwünschte und aufgezwungene Nachhilfemaßnahme für unbeholfene Lehrpersonen in die Lehrer(fort)bildungssysteme eingeführt und dort vermehrt eingesetzt werden. Allerdings kann nicht verschwiegen werden, dass innovative Coachingformen durchaus aufwändige und deshalb keine billigen Fortbildungsinstrumente darstellen. Sie können aber auch als langfristige Investition betrachtet werden, zum Beispiel wenn dadurch ein Beitrag zu einer nachhaltige(re)n Implementation von Bildungsstandards geleistet wird.

5.3.5 Videobasierte Analyse und Reflexion von Unterricht

Als weiteres innovatives Instrument des lehrpersonen- und unterrichtsbezogenen professionellen Lernens mit Potenzial für die didaktische Arbeit mit Bildungsstandards gewinnen Unterrichtsvideos als Medium der Lehrer(fort)bildung zunehmend an Bedeutung (Brophy 2004; Welzel/Stadler 2005). Unterrichtsvideos rücken, wie die Unterrichtsvorbereitung, -gestaltung und -reflexion mit einem Coach, die didaktische Arbeit der Lehrpersonen in den Fokus von Qualitätsbemühungen. Lehrende müssen wissen, was guter Unterricht ist und auf welche Weise sich die Lerngelegenheiten für ihre Schülerinnen und Schüler verbessern lassen (vgl. Darling-Hammond/Bransford 2006). Ein grundsätzlicher Tatbestand und gleichzeitig ein Problem besteht darin, dass Lehr-

personen ihre Arbeit in der Regel ausschließlich in Gegenwart ihrer Klasse und unter Ausschluss von Berufskolleginnen und -kollegen verrichten und ihrerseits selbst wenig Gelegenheit haben, Kolleginnen und Kollegen beim Unterrichten zu beobachten. Eine von außen gespiegelte Reflexion des eigenen Tuns und der dahinterstehenden – häufig unhinterfragten – Grundannahmen über Lehr-Lernprozesse findet daher eher selten statt und muss über entsprechende Hilfsmittel ermöglicht werden. Die in den vergangenen Jahren in mannigfaltigen Ausbildungskontexten (z.B. Einübung und Analyse von Bewegungsabläufen im Sport, Kommunikations-, Beratungs-, Verkaufsschulung) zum Einsatz kommende Videografie wird im Gefolge der TIMSS-Videostudien sowie sich daran anschließender Studien (vgl. Pauli/Reusser 2006) zunehmend auch für den Unterrichtsbereich bedeutsam. In den vergangenen Jahren wurden deshalb an verschiedenen Orten videobasierte Werkzeuge für die Aus- und Weiterbildung von Lehrpersonen entwickelt und erprobt. Welche neuen Möglichkeiten ergeben sich daraus?

Im Wesentlichen bieten videobasierte Fortbildungen folgende Vorteile (Reusser 2005b):

- Unterrichtsprozesse können in ihrer Komplexität und Variabilität sichtbar gemacht werden,
- Videos besitzen eine hohe Anschaulichkeit und Realitätsnähe,
- flüchtige Praxissituationen können strukturiert beobachtet werden,
- wiederholte Beobachtung und Archivierung sind möglich,
- unterrichtsbezogene Denk- und Handlungsmuster können objektiviert werden,
- Referenzierung und Reflexion von Standards und „best practice“ werden ebenso möglich wie eine problemorientierte Analyse,
- Theorien, Kernideen und Konzepte des Lehrens und Lernens können in Sichtstrukturen unterschiedlichen Handelns übersetzt werden,
- die (fach-)didaktische Verständigung über Lehr- und Lernprozesse wird erleichtert,
- verschiedene Analysemethoden werden miteinander verbunden und
- Videos können vielseitig eingesetzt werden.

Diese Optionen ermöglichen Lernprozesse im Bereich der Lehrer(fort)bildung, wie sie unter dem Sammelbegriff des *situierten und reflexiven Lernens* diskutiert werden. Fortbildungsansätze videobasierten Ler-

nens, in denen auch Videos der Teilnehmerinnen und Teilnehmer zum Einsatz kommen, lassen sich dadurch charakterisieren, dass

- reflexives Lernen in Lehrpraktika und in Fortbildungen situationsgebunden, datengestützt und im Selbstbezug erfolgt,
- didaktisches Handlungswissen in authentischen und sozialen Praxis-kontexten erworben wird,
- durch die Präsenz realer Anwendungssituationen die Entstehung „trägen Wissens“ vermieden wird,
- Lerngelegenheiten zur Verbindung von Theorie und Praxis geschaffen werden,
- Inhalte aus multiplen Perspektiven problemorientiert bearbeitet werden und
- selbstgesteuertes, auch kooperatives Lernen in größtmöglicher Berufsnähe erfolgt (ebd.).

Entsprechende Fortbildungsangebote, welche hinsichtlich ihrer Wirkungen wissenschaftlich ausgewertet wurden, sind noch relativ selten. Im Rahmen des DFG-Schwerpunktprogramms BIQUA („Bildungsqualität von Schule“) wurde am IPN in Kiel eine videobasierte Lehrerfortbildung mit einer virtuellen Lernumgebung für Physiklehrkräfte entwickelt. Auf dieser Grundlage können Physiklehrpersonen gruppenweise oder einzeln ihren Unterricht reflektieren und weiterentwickeln. Dabei interessiert unter anderem, wie lernförderlich die Probandinnen und Probanden das Medium Video erleben, welche Unterschiede in der Bearbeitung fremder und eigener Unterrichtsstunden bestehen und welche Rolle theoretische Fokussierungen und unterschiedlich stark strukturierte Aufgabenstellungen spielen (Seidel/Prenzel/Rimmele/Meyer/Dalehefte 2004; Kobarg/Schwindt/Seidel 2007).

Im Rahmen der Studie „Unterrichtsqualität und mathematisches Verständnis in verschiedenen Unterrichtskulturen (Klieme/Reusser/Pauli 2003) wurde 2004/2005 ebenfalls eine Fortbildung mit deutschen und Schweizer Lehrkräften durchgeführt. Die Teilnehmenden analysierten und reflektieren dabei ihr professionelles didaktisches Handeln unter Gesichtspunkten eines kognitiv aktivierenden Unterrichts und der Lernwirkungen auf Schülerinnen und Schüler. Kleingruppen von drei bis fünf Teilnehmenden unterstützen sich im Sinne professioneller Lerngemeinschaften bei der Planung und Reflexion von Unterricht (Ratzka/Lipowsky/Krammer/Pauli 2005).

Das Projekt „MuBiL“ (Mathematik unterrichten – Binationales und videobasiertes Lehrerinnen- und Lehrerfortbildungsprojekt), das von der Universität Augsburg und der Pädagogischen Hochschule Zentralschweiz Luzern gemeinsam durchgeführt wurde, bot einen Kontext, in dem sich Lehrpersonen gegenseitig im Unterricht besuchten, videografierten und gemeinsam Unterrichtsmaterialien entwickelten (Kuntze 2004).

In ähnlicher Weise, aber weitgehend ohne Benutzung des Mediums Video, werden in Japan bereits seit etlichen Jahren im Rahmen von sogenannten „Lesson studies“ Lehrpersonen unter Nutzung kooperativer Settings schulintern weitergebildet (Fernandez/Yoshida 2004; Morita 2005). Dabei werden in einem Kollegium gemeinsam – an bestimmten Stoffen, Lernzielen und Standards orientierte – Studienlektionen erarbeitet, präsentiert und unter Beteiligung externer Fachleute diskutiert, evaluiert und verschriftlicht. Die auf diese Weise verdichteten Resultate bilden dann die Basis für weitere Reflexion und Standortbestimmung. Auch in den USA haben „Lesson studies“ in den letzten Jahren Verbreitung gefunden und ergänzen zunehmend traditionelle Formen der Fortbildung von Lehrpersonen. Dabei werden auch Unterrichtsvideos zur Unterstützung des Lernprozesses eingesetzt (Lewis/Perry/Murata 2006; Wiburg/Brown 2007).

Forschungsergebnisse zu videobasiertem Lernen von Lehrpersonen sind noch relativ rar. Damit bleiben viele Fragen, die sich zu seiner differentiellen Wirksamkeit stellen, vorderhand unbeantwortet. Sherin/Han (2004) konnten in einer Studie nachweisen, dass sich im Laufe einer einjährigen Sequenz videobasierten Lernens mit monatlichen Treffen die Art der Gesprächsbeiträge der Teilnehmenden veränderte. Während sie zu Beginn vor allem das Verhalten der Lehrpersonen beurteilten, verlagerte sich ihre Aufmerksamkeit später vermehrt auf die Lernprozesse der Schülerinnen und Schüler. Ebenso wurde im Laufe der Zeit eher versucht, beobachtbares Verhalten zu verstehen, als dass alternative didaktische Konzepte vorgeschlagen wurden (ebd.). In einer Pilotstudie identifizierten Krammer/Hugener (2005) interessante Lernaufgaben, welche sich sinnvoll auf den Stoff beziehen, die angemessene Auswahl der Videos und eine tiefe technische Arbeitsbelastung als produktive Bedingungen für den Einsatz von Unterrichtsvideos. Und in der einjährigen schweizerisch-deutschen Lehrerfortbildung schätzten die beteiligten Lehrpersonen – bei hoher Akzeptanz des neuen Instrumentes – die Erträge positiv ein (Krammer/Ratzka et al. 2006). Eine

objektivierte Messung der Entwicklung des unterrichtsbezogenen Wissens anhand von Videobefragungen bestätigt diese positiven subjektiven Einschätzungen der Lehrpersonen und zeigt, dass die Teilnehmenden ihr Wissen über Möglichkeiten der Gestaltung eines anregenden, am Lernen der Schülerinnen und Schüler orientierten Unterrichts im Laufe der Fortbildung erweitert haben (Krammer/Schnetzler/Ratzka/Pauli/Reusser 2007). Kuntze (2006) wies u.a. nach, dass die Teilnehmenden des Projektes „MuBiL“ im Anschluss an die Fortbildung die Qualität von Unterricht anhand einheitlicherer Kriterien beurteilten als zu Beginn.

Allerdings gilt es auch, die Herausforderungen im Blick zu behalten, welche bei der Reflexion und Analyse von Unterrichtsvideos in der Lehre entstehen. So müssen sich die Teilnehmenden ihrer Subjektivität beim Betrachten der Sequenzen bewusst sein. Ebenso ist auf eine von Respekt geprägte Diskussionskultur zu achten. Weiter benötigt eine eingehende Reflexion gezielte Anleitung durch geschulte Dozierende. Nicht zuletzt muss der Auswahl und Aufbereitung geeigneter Sequenzen die nötige Aufmerksamkeit beigemessen werden (Krammer/Reusser 2004, 2005; Sherin et al. 2006). Staub (2005) verweist zudem nachdrücklich auf den Aufbau eines theoriebezogenen Verständnisses bzw. von Gesichtspunkten der Qualität von inhaltspezifischen Lehr- Lernprozessen beim Einsatz von Videos im fachspezifisch-pädagogischen Coaching. Videos allein bilden höchstens den „Rohstoff“ für ein vielversprechendes Fortbildungswerkzeug, welches sein Potenzial für die Unterrichtsentwicklung nur in sorgfältig gestalteten, auf die Bedürfnisse von Lehrpersonen und auf klare Ziele abgestimmten Aus- und Fortbildungssettings zu entfalten vermag (Krammer/Reusser 2004; Seago 2004). Nur auf diese Weise können sie zu einem Motor der Reflexion und Weiterentwicklung des professionellen Handelns von Lehrpersonen werden (ebd.).

Hinsichtlich der Einführung und Implementierung fachlicher Standards im Unterricht gilt Ähnliches wie für Ansätze des fachspezifisch-pädagogischen Coachings: eine nachhaltige Reflexion professionellen Handelns muss auf erfolgreiche Lernprozesse der Schülerinnen und Schüler bezogen werden. Wo Lernergebnisse von Schülerinnen und Schülern im Sinne der durch Bildungsstandards beschriebenen Kompetenzen angestrebt werden, müssen auch die dahin führenden Prozesse und deren Qualität in den Blick genommen werden (Stäudel/Blum 2005). Ein entsprechendes Qualitätsbewusstsein kann sich aber nur dort entwickeln, wo sich Lehrpersonen sowohl mit theoretischen Ansätzen

und praktischen inhaltlichen Aufgaben als auch mit der (selbst-)kritischen Beurteilung des Erreichten auseinander setzen (ebd.). Videobasierte Aus- und Weiterbildung von Lehrpersonen oder fachspezifisch-pädagogisches Coaching bieten hier ideale Voraussetzungen, weil sie auf eine gründliche und kritische Auseinandersetzung mit dem eigenen Unterricht abzielen. Denkbar sind auch Weiterbildungsszenarien, welche beide vorgestellten Mittel miteinander kombinieren: Coaching kann dabei den Einführungsschritt markieren, welcher Lehrpersonen für das Geschehen im eigenen Unterricht sensibilisiert, während die Arbeit mit Videos auf gründliche Reflexion abzielt. Insgesamt zeigen diese beiden Instrumente Möglichkeiten der Personalentwicklung, welche die Lehrpersonen selber als Ort und Ausgangspunkt von Entwicklungsprozessen sehen.

Dass dieser Weg – sowohl in zeitlicher als auch in finanzieller Hinsicht – aufwändig ist, muss dabei im Auge behalten werden.

5.3.6 Lehren und Lernen – Perspektive und Verantwortung der Schülerinnen und Schüler

Lehrerinnen und Lehrer sind – wie in dieser Expertise mehrfach betont – zwar die primären Adressaten der Bemühungen um eine Implementation von Bildungsstandards. Letztlich ist aber von entscheidender Bedeutung, ob und wie Schülerinnen und Schüler auf deren Basis lernen. Schülerinnen und Schüler und ihre Lernleistungen sind jedoch nicht einfach eine abhängige Variable im Bildungssystem. Nimmt man einen Begriff des *aktiven Lernens* zur Grundlage, welcher die Lernenden nicht als passive Rezipienten, sondern als verstehens- und problemorientierte, zunehmend selbstgesteuerte und für ihr Lernen verantwortliche Akteure begreift, erhalten diese ihre Rolle ebenfalls als unabhängige – durch Unterricht nicht beliebig steuerbare – Koproduzenten ihrer Lernergebnisse. Dass Lehrpersonen sich in der Folge vermehrt auf eine Rolle als unterstützende Lerngerüste (*scaffolds*) und Lernbegleiter einstellen und ein gegenüber traditionellen Formen der direkten Instruktion erweitertes didaktisches Profil entwickeln müssen, stellt das Komplement und die Kehrseite des in den letzten Jahrzehnten vollzogenen „Paradigmenwechsels“ von einer behavioristischen zu einer kognitiv-konstruktivistischen Lehr-Lernphilosophie dar. Mit anderen Worten: Durch die Qualitätsinitiative „Bildungsstandards“ werden nicht nur die Lehrkräfte, sondern auch die Schülerinnen und Schüler herausgefordert.

Nicht nur von den Lehrpersonen wird erwartet, dass sie sich in neuer Weise auf die Qualitätsfragen ihres unterrichtlichen Handelns einlassen, auch von den Schülerinnen und Schülern wird erwartet, dass sie nicht in der Rolle als passive Unterrichtsteilnehmer verharren, sondern sich als verantwortliche Subjekte auf Wissen und Können und damit auf eine Leistungspartnerschaft einlassen und sich aktiv am Lerngeschehen beteiligen.

Im Rahmen der erwähnten Neuzentrierung des Lehr-Lernbegriffs und den dadurch ausgelösten Diskussionen um eine „neue Lernkultur“ und um einen multikriterialen Begriff von Bildungswirksamkeit, stellt der Unterricht eine Angebotsstruktur dar, welche von den Lernenden genutzt werden will. Angebots-Nutzungsmodelle sehen die Schülerinnen und Schüler denn auch als aktive (von ihrer Herkunft, ihren Lernstrategien und von ihrer Motivation, d.h. von der Mitgift des sozialen und kulturellen Kapitals des Herkunftsmilieus geprägte), mehr oder weniger fähige *Nutzer* eines bestehenden Angebots (Reusser/Pauli 1999; Fend 2000; Helmke 2003). Von entscheidender Bedeutung in Bezug auf die „Outcomes“ bzw. den Ertrag von Bildungsprozessen ist daher immer auch die *Qualität der Nutzung von Lerngelegenheiten*. Im Zuge der „kognitiven Wende“ in der Lern- und Kognitionspsychologie wird schon seit längerem von eigenständigen (Beck/Guldemann/Zutavern 1991) oder selbstregulierten Lernerinnen und Lernern (u.a. Boekaerts 1996; Pintrich 2000) gesprochen. Auch PISA geht davon aus, dass Lernende in der Lage sind, „sich selbstständig Lernziele zu setzen, dem Inhalt und Ziel angemessene Techniken und Strategien auszuwählen und sie auch einzusetzen. Ferner halten sie ihre Motivation aufrecht, bewerten die Zielerreichung während und nach Abschluss des Lernprozesses und korrigieren – wenn nötig – die Lernstrategie“ (Deutsches PISA-Konsortium 2001, S. 271).

Ein adaptiver und flexibler Umgang mit Lerntechniken und -strategien setzt jedoch voraus, dass Schülerinnen und Schüler solche auch kennen. Lehrpersonen sind deshalb auf vielfältige Weise herausgefordert, die Vermittlung von fachlichem Wissen an den Erwerb sinnvoller und möglichst generalisierbarer Strategien zu koppeln. Wie an verschiedener Stelle in dieser Expertise bereits erwähnt, stoßen entsprechende innovative didaktische Konzepte bei den Lehrpersonen auf gute Resonanz und zeitigen erfreuliche Wirkungen auf motivationaler und leistungsbezogener Seite. So ergab etwa eine Evaluation des Forschungsprojekts WELL (wechselseitiges Lehren und Lernen), in dem Lehrpersonen mit Techniken kooperativen Lernens vertraut gemacht wurden,

dass Schülerinnen und Schüler im Biologieunterricht auf der Sekundarstufe hinsichtlich Motivation und fachlicher Leistung günstigere Werte erreichten, wenn sie mit Gruppenpuzzle, strukturierter Kontroverse oder Partner- und Gruppeninterview selbständig arbeiteten; dies gegenüber Schülerinnen und Schülern einer Vergleichsgruppe, deren Lehrpersonen keine derartigen kognitiv aktivierenden Lern- und Arbeitsformen vermittelten (vgl. Huber 2005).

Schülerinnen und Schüler erscheinen in einer solchen Sichtweise – neben den Lehrpersonen und den weiteren pädagogischen Akteuren – ebenfalls ein Stück weit *selbst gegenüber ihrem Lernen verpflichtet* und für dessen Ergebnisse *haftbar*. Damit sie jedoch überhaupt selbständig lernen und dafür Verantwortung übernehmen können, müssen sie ebenfalls dessen Ziele kennen und über Jahre Gelegenheit haben, kognitive, kommunikative, soziale, motivationale und metakognitive Kompetenzen zu ihrer Erreichung auszubilden und zu entwickeln (Beck/Guldimann/Zutavern 1991). Die entscheidende Kernaufgabe eines unter multiplen Qualitätsmerkmalen attraktiven Unterrichts besteht dabei darin, ihnen die erforderlichen Lerngelegenheiten zu verschaffen – und dies ebenfalls im Hinblick auf eine nachhaltige, die Lernenden als Subjekte erreichende Implementation von Bildungsstandards. Standards müssen demnach so gefasst werden, dass sie nicht nur von Lehrpersonen, sondern letztlich auch von den betroffenen Schülerinnen und Schülern und ihren Eltern als Orientierungsmarken und Ziele ihres Lernens verstanden werden. Für die Lehrpersonen bedeutet dies mindestens dreierlei: dass sie die Erwartungen an das Lernen der Schülerinnen und Schüler transparent kommunizieren, dass sie qualitativ unterrichten und dass sie Schülerinnen und Schüler und Eltern periodisch darüber informieren, welche Ziele erreicht wurden und welche nicht.

Ein zentraler Aspekt in Bezug auf die Forderung, dass Bildungsstandards nicht nur ein zentrales Kriterium für das Lehrerhandeln, sondern auch für das Lernen und die Lernergebnisse jeder einzelnen Schülerin und jedes einzelnen Schülers werden, ist die *Kultur der Leistungsbewertung und der Leistungsdiagnostik* in unseren Schulen. Produktives Lehren und Lernen bedarf in wechselseitiger Abstimmung differenzierter, bilanzierender und diagnostisch-formativer Rückmeldungen zu den Ergebnissen und Prozessen des Lernens. Die traditionelle Kultur der Leistungsbewertung (die ebenfalls eine analoge Lehr-Lernkultur widerspiegelt) sieht jedoch summative Tests als oftmals einzigen Abschluss eines Prozesses der Wissensaneignung vor (Winter 2002). Mit der Note erhält der Schü-

ler bzw. die Schülerin eine Rückmeldung zu der in einer Klausur erbrachten Leistung. Diese an sich schon „schwierige und fehleranfällige Aufgabe für den Lehrer“ (Schrader/Helmke 2002) bringt den Lernenden und den Lernprozessen nur begrenzten Nutzen.

Entsprechend fordert Winter (ebd., S. 52) eine Leistungsbewertung,

„die vor allem qualitativ gute Rückmeldung liefert, die genutzt werden kann, während gelernt und gearbeitet wird, die hilft, zu guten Leistungen zu gelangen. Dazu gehören Maßnahmen der gemeinsamen Vorschau, Lernkontrakte, Reflexionen im Prozess sowie die Anleitung zur Selbstbewertung und die Vermittlung von Bewertungskriterien an die Schülerinnen und Schüler. Dazu passen Lernpartnerschaften und Lernarrangements, in denen Lehrerinnen und Lehrer Zeit haben, zu beobachten und zu beraten.“

Bildungsstandards können für die Leistungsbewertung insofern durchaus positive Konsequenzen haben, als sie den Wechsel von einer vorwiegend sozialen zu einer kriterialen Bezugsnorm fördern. Schülerinnen und Schüler erhalten idealerweise über die reinen Ergebnisse hinausgehende, diagnostisch-informative Rückmeldungen, wie sich ihre Leistung in Bezug auf bestimmte fachliche Kriterien verhält, und können diese zur Entwicklung und Korrektur ihres Lernverhaltens nutzen. Jedoch ist dies ein äußerst anspruchsvolles Unterfangen, das – wenn es gelingen soll – über bilanzierende Schulleistungstests hinaus eines zusätzlichen und erheblichen lerndiagnostischen Aufwandes bedarf. Ergebnisse aus Vergleichsarbeiten können hier ebenfalls primär Aussagen über (summativ) erreichte Lernstände vermitteln, ohne aber die entsprechenden Prozessmerkmale zu berücksichtigen. Schneewind (2006) verweist nachdrücklich auf die Berücksichtigung der Bedürfnisse des Empfängers, damit dieser die Rückmeldungen in Lernzuwächse umsetzen kann. Auch müssen die Feedbacks auf beeinflussbare und veränderbare Handlungsaspekte zielen, damit die Lernmotivation erhalten bleibt.

Hier zeigt sich die Notwendigkeit einer zusätzlichen individuellen Bezugsnorm, welche Leistungen im Längsschnitt betrachtet und erzielte Ergebnisse daran misst, was von den Lernenden in diesem Bereich bisher erbracht wurde und worin die Fortschritte liegen. Schwankungen im Lernverlauf können auf diese Weise ebenso sichtbar gemacht werden wie Leistungszuwächse. Damit rückt eine an fachlichen Kriterien ausgerichtete, über systemische Trendanalysen (wie beim Bildungsmonitoring) hinausgehende, auch auf individuelles Lernen gerichtete *längsschnitt-*

liche Leistungs- und pädagogische Förderdiagnostik ins Zentrum. Bis dahin ist es sicherlich ein sehr weiter Weg, der (auch wissenschaftlich) nur schrittweise instrumentiert und begangen werden kann. Zudem berühren die anvisierten Veränderungen die Grundlagen unseres Bildungs- und Unterrichtssystems. Immerhin sind Schritte in die vorgezeichnete Richtung nicht nur denkbar, sondern werden an einzelnen Orten auch gegangen. Denkbar sind etwa ergänzende Rechenschaftslegungen der Lernprozesse in Form von Portfoliomethoden oder Lern- und Reisetagebüchern (Winter 2004; Ruf/Gallin 1991), wobei die Lernenden in die Auswahl der zu berücksichtigenden Lernprodukte und Prozessdokumente einbezogen werden können. Schülerinnen und Schüler müssen (zusammen mit ihren Eltern) damit im Rahmen der Einführung von Bildungsstandards genauso zu einer (allerdings begrenzten) *Accountability* verpflichtet werden wie Lehrpersonen und Schulen. Standards können und müssen im Lichte des skizzierten Verständnisses eines zunehmend selbstverantwortlichen Lernens auch von den Schülerinnen und Schülern als sinnvolle Ziele und Erwartungen an zu erreichende Leistungs- und Kompetenzniveaus verinnerlicht werden.

Eine erfolgreiche Implementation von schulischen Standards durchdringt somit sämtliche Ebenen des Schulsystems:

- Schulen sind herausgefordert, Lehrenden und Lernenden möglichst optimale Arbeits- und Lernbedingungen zu schaffen;
- Lehrpersonen entwickeln ihren Unterricht im Sinne produktiver und auf Verstehen ausgerichteter Lernprozesse;
- Schülerinnen und Schüler sehen sich – idealerweise unterstützt durch das häusliche Umfeld – als eigenständige Lernerinnen und Lerner für die Erreichung zumindest minimaler Kompetenzziele mitverantwortlich.

Nur wenn es gelingt, Standards auf allen Ebenen des Bildungssystems auf nachhaltiges Lernen hin zu beziehen und wenn sich auch die Schülerinnen und Schüler selbst im Verlaufe ihrer Schulkarrieren in wachsendem Maße zur Realisierung von Bildungserwartungen und -zielen verpflichtet fühlen, kann das Unternehmen „Bildungsstandards“ Erfolg haben.

6 Stand der Diskussion und Entwicklung in den Auftragsländern

Die bildungspolitischen Ziele und Erwartungen, die sich gegenwärtig mit Bildungsstandards verknüpfen, insbesondere die Forderung nach der Vergleichbarkeit der Lerninhalte und der Bewertungsmaßstäbe, gewannen längst vor den Diskussionen im Anschluss an TIMSS und PISA an Bedeutung. Entsprechend werden heute im Zusammenhang mit Bildungsstandards auch scheinbar „alte“ Fragen wie Strategien der Lernzieloperationalisierung, als zentrale Problemstellung der Curriculumreformen der siebziger Jahre, in neuer begrifflicher Einkleidung wieder verhandelt (Klieme 2004a). Allerdings ist seinerzeit weder über regelmäßige Leistungstests noch über neue Formen der Evaluation diskutiert worden. Auch das Bildungsmonitoring war kein Thema. Ziele und Resultate standen in keinem empirisch kontrollierbaren Zusammenhang, auch weil die Forschung in den deutschsprachigen Ländern zu wenig entwickelt war.

Das hat sich inzwischen deutlich geändert. Die besondere Bedeutung, die Bildungsstandards als staatlichem Steuerungsmittel zukommt, steht in einem klaren Zusammenhang mit den Ergebnissen der internationalen Leistungsstudien. In den politischen und konzeptionellen Papieren fehlt kaum der Hinweis auf die Tatsache, dass erfolgreiche PISA-Länder in Skandinavien oder im angelsächsischen Raum ihre Anstrengungen zur Systemsteuerung seit längerem auf eine Ergebnisorientierung umgestellt hätten. Daraus ziehen politische Gremien den Schluss, dass eine „vorrangige Inputsteuerung allein nicht zu den erwünschten Ergebnissen im Bildungssystem führt. Die Festlegung und Überprüfung der erwarteten Leistungen müssen hinzukommen.“ (Kultusministerkonferenz (KMK) 2005, S. 5; für Österreich vgl. ähnlich: Bundesministerium für Bildung, Wissenschaft und Kultur (BMBWK) 2004, S. 1). Im Vordergrund steht auch in den vier Auftragsländern eine neue Politik, die Qualitätssicherung auf verschiedenen Ebenen an messbaren Ergebnissen orientiert, mit dem Ziel, über Rückmeldungen den Unterricht und das Lernen zu optimieren.

Weitere pädagogische und bildungspolitische Reformthemen, teilweise mit längerer Kontinuität, schließen sich in sämtlichen vier Ländern, wenn auch mit unterschiedlicher Gewichtung, an:

- Konkretisierung von Bildungszielen für Fächer oder Lernbereiche mit Hilfe von Kerncurricula;
- Erhöhung der Autonomie und zugleich der Verantwortung der Schulen einschließlich der Formulierung eigener Curricula;

- Betonung der autonomen Professionalität der Lehrkräfte, die Gestaltung von Lehr-/Lernprozessen soll mit der Einführung von Bildungsstandards entwickelt und nicht eingeschränkt werden;
- Bildungsstandards als fairer Referenzrahmen für die Notengebung und die Vergleichbarkeit von Abschlüssen;
- Förderung individueller Leistungen sowie der Leistung des Systems vor dem Hintergrund einer internationalen Wettbewerbssituation sowie in einer rasch sich wandelnden „Wissensgesellschaft“;
- Beförderung der Chancengleichheit angesichts breiter Leistungsvarianzen in Abhängigkeit von sozialer und kultureller Herkunft.

Ausgehend von diesen Postulaten soll im Folgenden ein Überblick über den Stand der Entwicklung von Bildungsstandards in Deutschland, Österreich, der Schweiz und Luxemburg gegeben werden. Wie bereits angesprochen, sind vor dem Hintergrund globalerer steuerungspolitischer Trends zum Teil weitgehende Überschneidungen bezüglich Erwartungen und Zielsetzungen zwischen den Ländern festzustellen. Unterschiedlich ist hingegen der Stand der Umsetzung sowie in Teilen das Vorgehen und die Reichweite der Reformen; dies hat wiederum mit den unterschiedlichen Ausgangslagen, u.a. den politisch-strukturellen Voraussetzungen, zu tun. Je nach Stand der Entwicklung und vorhandenen Quellen, besonders aber auch je nach Spezifika potentieller Problemlagen, finden sich zum Teil unterschiedliche Gewichtungen in der Darstellung der Länder. So ist für Österreich auf dem Papier ein Implementationskonzept vorhanden, das in Pilotphasen erprobt und allenfalls weiterentwickelt werden kann, während in Deutschland die Bildungsstandards zwar beschlossen sind, die föderale Struktur jedoch besondere Implementationsfragen aufwirft; anders ist aber die Ausgangslage im relativ kleinräumigen Großherzogtum Luxemburg oder in der Schweizer Konföderation, wo der politische Beschluss der Bildungsstandards und damit auch Implementationsfragen unmittelbar bevorstehen. Sind ein stimmiges Konzept der Implementation von Bildungsstandards und eine Abschätzung der Folgen insgesamt erst rudimentär vorhanden, dürfte die Entwicklung selbst unumkehrbar sein, ungeachtet dessen, dass auch in den Auftragsländern Kritik gegeben ist.

6.1 Deutschland

6.1.1 Diskussionen und Problemwahrnehmungen

Die Notwendigkeit der Sicherung und Weiterentwicklung der Qualität des Bildungssystems stand in Deutschland bereits vor der Publikation der Ergebnisse von PISA 2000 auf der politischen Agenda, wie etwa am „Konstanzer Beschluss“ (Beschluss der 280. Kultusministerkonferenz vom 23./24.10.1997) zur Qualitätssicherung in Schulen mittels Evaluation und Vergleichsuntersuchungen abzulesen ist. Die Bundesrepublik Deutschland hatte bereits in den sechziger Jahren an Vergleichsstudien teilgenommen und sich dann auch in allen großen Projekten engagiert (Study of Reading Literacy, TIMSS, International Adult Literacy Survey). Die Ergebnisse der ersten PISA-Studie kamen für die Fachwelt keineswegs unerwartet und der sogenannte „PISA-Schock“ stellte insofern nur den öffentlichkeitswirksamen Höhepunkt bereits anhaltender Reformdiskussionen dar.

Die Befunde der ersten PISA-Studie zeigten in sämtlichen untersuchten Kompetenzbereichen mittlere Ergebnisse der deutschen Schülerinnen und Schüler, die unter dem OECD-Durchschnitt lagen. Dies trifft insgesamt auch für die Erhebung von 2003 zu, wenngleich in Teilbereichen der Mathematik und der naturwissenschaftlichen Kompetenz gegenüber PISA 2000 signifikante Zuwächse zu beobachten sind. Vor allem im unteren Leistungssegment waren die Ergebnisse vergleichsweise schlecht. Prägnant war zudem der Zusammenhang zwischen sozialer Herkunft und Kompetenzerwerb, dies besonders im Bereich der Lesekompetenz.

Daraus wurde gefolgert, dass in Deutschland Niveauanforderungen der schulischen Bildung nicht ausreichend eingehalten würden. Der Vergleich der Bundesländer untereinander in der Ergänzungsstudie PISA-E machte zudem deutlich, dass Leistungsniveaus und Benotungsmaßstäbe zwischen den Ländern, aber auch innerhalb der Länder zwischen Schulformen und einzelnen Schulen, zum Teil stark variieren. Aus dem Vergleich mit anderen Teilnehmerländern, die eine outputorientierte Qualitätssicherung kennen, wurde der Schluss gezogen, dass in Deutschland inhaltliche Standards für den Unterricht fehlten, die verbindlich gesetzt seien und deshalb erreicht werden müssten. So entstand in der bildungspolitischen Diskussion der Ruf nach national einheitlichen Standards, die sowohl die Art als auch das Niveau von schulischen Anforderungen regeln und prüfbar machen sollten.

Vorhaben, die auf eine Intensivierung der Qualitätssicherung abzielen, sind in Deutschland seit Beginn der neunziger Jahre verstärkt wahrzunehmen. Im

Mittelpunkt stand und steht die Erhöhung der Eigenständigkeit der Schulen bei gleichzeitiger Zurücknahme der Schulverwaltung und Kontrollfunktion der Schulaufsicht, verbunden mit einem Ausbau von Unterstützungssystemen und Beratungsangeboten sowie einer neuen Rolle der Schulleiter (Döbert 2002, S. 99f.). Dennoch hat sich in Deutschland das „traditionelle Modell“ der Qualitätssicherung (Rolff 1995), das weitgehend auf Rechts- und Verwaltungsvorschriften basiert und kaum Evaluationen kennt, als immer noch sehr stabil behauptet. Änderungen der Steuerungsphilosophie und erste praktische Schritte sind erst in den letzten Jahren festzustellen.

Der Grund ergibt sich aus den Ergebnissen der internationalen Schulleistungsvergleiche und deren öffentlicher Diskussion, die in Deutschland besonders heftig geführt wurde. Die Weiterentwicklung des deutschen Bildungswesens in Richtung mehr Effektivität und Transparenz wurde zu einem zentralen und dauerhaften Thema, das politisches Handeln verlangte. In den Argumentationen wurden auch politisch bestimmte Eigenheiten des deutschen Bildungssystems erneut in den Mittelpunkt gerückt, die bereits frühere Debatten bestimmt hatten, wie die Gliederung des Schulsystems und die föderale Zuständigkeitsstruktur. Es war sogar von einer perennierenden „Bildungskatastrophe“ die Rede, weil frühere Reformen offenbar nicht zum gewünschten Resultat geführt hatten.

Mit dem Ziel, über die Qualität und Zukunftsfähigkeit des deutschen Bildungssystems zu beraten und Empfehlungen auszuarbeiten, lancierten Bund und Länder die Diskussionsplattform „Forum Bildung“ (1999-2002). Dazu eingeladen wurden als ständige Mitglieder Vertreter aus Wirtschaft, Wissenschaft und Gesellschaft. Fünf Themenschwerpunkte standen im Zentrum:

- Bildungs- und Qualifikationsziele von morgen,
- Förderung von Chancengleichheit,
- Qualitätssicherung im internationalen Wettbewerb,
- lebenslanges Lernen und
- neue Lern- und Lehrkulturen.

Aus den Arbeiten der Arbeitsgruppen gingen verschiedene Publikationen sowie zwölf Empfehlungen hervor (Forum Bildung 2001). Im letzten Punkt dieser Empfehlungen wurde die Forderung nach einer stärkeren Eigenverantwortung für Bildungseinrichtungen vertreten, verbunden mit der Bereitschaft zur Rechenschaftslegung und so zur internen wie externen Evaluationen. Die Schlussfolgerung an der Abschlusstagung im Januar 2002 lautete, dass eine erfolgreiche Bildungsreform von einem „von oben“ bestimmten Konzept ausgehen müsse,

das an der Basis auszugestalten sei (Döbert/Klieme/Sroka 2004, S. 334). Das Forum hob zudem die Notwendigkeit hervor, regelmäßig einen Bericht über das deutsche Bildungssystem zu erstellen, der den Stand der nationalen Entwicklung dokumentiert. Die Einführung eines solchen Berichtes wurde 2002 von der deutschen Kultusministerkonferenz (KMK) beschlossen.

Eine Reihe der vom Forum vorgeschlagenen Maßnahmen findet sich in den sieben Handlungsfeldern der KMK wieder, die im Anschluss an die Analyse der PISA-Ergebnisse im Dezember 2001 formuliert wurden. Zwei dieser Felder beziehen sich auf

- Maßnahmen zur konsequenten Weiterentwicklung und Sicherung der Qualität von Unterricht und Schulen auf der Grundlage von verbindlichen Standards sowie einer ergebnisorientierten Evaluation (Handlungsfeld 5);
- Maßnahmen zur Verbesserung der Professionalität der Lehrertätigkeit, insbesondere im Hinblick auf diagnostische und methodische Kompetenz als Bestandteil systematischer Schulentwicklung (Handlungsfeld 6).

Anlässlich der 298. Plenarsitzung vom 23./24. Mai 2002 fassten die Kultusminister den Grundsatzbeschluss, für ausgewählte Schnittstellen der allgemeinbildenden Schularten *Bildungsstandards* zu erarbeiten, in den Ländern verbindlich einzuführen und überprüfen zu lassen. Der Begriff „Bildungsstandards“ erlebte danach eine rasante Karriere und steht heute im Mittelpunkt einer intensiven Entwicklungsarbeit, an der alle deutschen Bundesländer beteiligt sind.

6.1.2 Konzeption und Funktion von Bildungsstandards

Fünf Monate nach dem Mai-Beschluss erfolgte eine Präzisierung dahingehend, dass die Länder in landesweiten bzw. länderübergreifenden Orientierungs- und Vergleichsarbeiten überprüfen sollten, in welchem Umfang die Standards tatsächlich erreicht würden. Die Vergleichs- und Orientierungsarbeiten sollten der Qualitätssicherung dienen und den Lernprozess begleiten (Beschluss der 299. Kultusministerkonferenz vom 17./18.10. 2002). Die Überprüfung sollte nicht auf das Ende der schulischen Laufbahn konzentriert sein, um so durch Maßnahmen der individuellen Förderung die Zielerreichung zu ermöglichen.

Gleichzeitig hatte das Bundesministerium für Bildung und Forschung die Expertise *Zur Entwicklung nationaler Bildungsstandards* in Auftrag gegeben. Wie oben erwähnt, lautete die zentrale Aussage der 2003 publizierten Expertise, dass die Festlegung der fachlichen Standards von fachdidaktisch und pädagogisch-psychologisch begründeten Kompetenzmodellen auszugehen habe, die darstellen, in welchen Teildimensionen und Graduierungen (Kompetenzstufen)

Erwartungen an das fachliche Lernen beschrieben und überprüft werden können. Diese Idee wurde nicht verworfen, aber auch nicht unmittelbar realisiert. Die KMK ging davon aus, dass die Grundlagenarbeit für die Entwicklung von Kompetenzmodellen noch nicht geleistet sei und daher unerschwinglich entschieden werden müsse.

Die Entwicklung von Kompetenzmodellen ist fachdidaktisch und psychologisch höchst anspruchsvoll. Kompetenzmodelle sind als Hypothesen anzusehen, die empirisch geprüft werden müssen (Klieme 2004b). Aufgaben, die den Kompetenzstufen zugeordnet sind, illustrieren die erreichbaren Niveaus, aber sie bilden noch keinen Beleg für deren Gültigkeit. Im Anschluss an die theoretischen Arbeiten, die Abgrenzung und Formulierung von Kompetenz-Teildimensionen sowie die Konstruktion und Zuordnung von Aufgabenstellungen, müssen die Überprüfung des Modells und die Normierung der Test-items anhand einer Stichprobe erfolgen.²⁰¹ Erst nach der Beschreibung von Kompetenzdimensionen und -niveaus sowie von Leistungsverteilungen (Normierung) können normative Erwartungen, d.h. Standards im eigentlichen Sinn (z.B. Mindeststandards oder schulformspezifische Regelstandards) bestimmt werden (Klieme 2004b). Neben offenen technischen Fragen etwa nach Verfahren zur Einschätzung von Anforderungsmerkmalen, der Abgrenzung von Niveaus, der Berücksichtigung fächerübergreifender Kompetenzen und schwer messbarer Bereiche stellt sich die grundsätzliche Frage, inwiefern Kompetenzmodelle imstande sind, die Lernentwicklungen abzubilden, also den Prozess zu erfassen. Während die Expertise nachdrücklich die Festlegung von Mindeststandards empfahl, entschied sich die KMK mit der Begründung des Fehlens der notwendigen validierten Kompetenzmodelle für Regelstandards. Die Argumentation für dieses „pragmatische Vorgehen“ (Kultusministerkonferenz (KMK) 2004, S. 14) lautete weiter, dass die Definition von Mindeststandards in einem ersten Schritt die Gefahr berge, Teile der Schülerschaft zu unter- bzw. zu überfordern. Die „Gewinnung von Kompetenzstufen“ stehe im Mittelpunkt der weiteren Entwicklung und bilde die Voraussetzung für die Definition von Mindeststandards (ebd., S. 15). Es wird in diesem Zusammenhang gesagt, dass auch die Expertise von Klieme/Avenarius/Blum et al. „Zur Entwicklung nationaler Bildungsstandards“ (2003) auf diese Notwendigkeit einer längerfristigen Ent-

²⁰¹ Diese Validierung und Normierung geschieht für den Bereich Mathematik nach dem Feldtest mit dem PISA-Durchgang 2006. Bis Ende 2008 sollen auch die Standards in den Fächern Deutsch und den Fremdsprachen Englisch und Französisch erarbeitet sein. Die entsprechenden Arbeiten in den Fächern Biologie, Physik, Chemie (Mittlerer Schulabschluss) werden bis nach 2009 dauern.

wicklungs- und Erprobungsperspektive verweise. Schließlich wurde bereits darin festgehalten, dass die Festlegung realistischer Mindeststandards Zeit brauche und solche erst durch eine Erprobung in den Schulen und nach ersten Erfahrungen mit Tests gesichert werden könnten.

Die deutsche Kultusministerkonferenz (KMK) hat am 16. Dezember 2004 dann zustimmend von dem Papier Kenntnis genommen, welches die Einführung von nationalen Bildungsstandards als Teil einer neuen Politikstrategie begründet. In diesem Papier ist von einem „Paradigmenwechsel in der Bildungspolitik“ die Rede, der abziele auf eine „Outcome-Orientierung“, damit verbunden auf Rechenschaftslegung und Systemmonitoring (Kultusministerkonferenz (KMK) 2005, S. 6). Die tatsächlich erreichten Ergebnisse sollten mehr Gewicht erhalten als bisher und zur Systemsteuerung eingesetzt werden. Dies deckt sich weitgehend mit den Argumenten der Expertise von 2003 sowie mit dem internationalen Diskussionsstand.

Bereits vorher, am 15. Oktober 2004, sind inhaltliche Bildungsstandards beschlossen worden, die nationale Vorgaben darstellen. Dabei kommt ein bestimmtes Anforderungsprofil zur Anwendung, das wie folgt beschrieben wird:

- Bildungsstandards greifen Grundprinzipien des jeweiligen Unterrichtsfaches auf,
- sie beschreiben die fachbezogenen Kompetenzen einschließlich zugrundeliegender Wissensbestände, die Schülerinnen und Schüler bis zu einem bestimmten Zeitpunkt ihres Bildungsganges erreicht haben sollen,
- zielen auf systematisches und vernetztes Lernen und folgen so dem Prinzip des kumulativen Kompetenzerwerbs,
- beschreiben erwartete Leistungen im Rahmen von Anforderungsbereichen,
- beziehen sich auf den Kernbereich des jeweiligen Faches und geben den Schulen Gestaltungsräume für ihre jeweilige pädagogische Arbeit,
- weisen ein mittleres Anforderungsniveau (Regelstandards) aus,
- werden durch Aufgabenbeispiele veranschaulicht (ebd.).

Dieser Bezug auf Leistungen und Leistungsmessung wird anschließend in den größeren Kontext der allgemeinen Bildungsziele gestellt, also nicht isoliert betrachtet. Schulqualität, so die KMK, ist „selbstverständlich mehr“ als „das Messen von Schülerleistungen anhand von Standards“ (ebd.). Beides, die „funktionale Aufgabe von Bildungsstandards“ und die „Ziele einer zeitgemäßen Allgemeinbildung“, stehen, so die KMK, nicht im Widerspruch zueinander. Bildungsstandards übernehmen eine bestimmte Funktion im Prozess der Qualitätssicherung; nur in dieser Hinsicht kann also von einem Paradigmenwechsel gesprochen werden.

Die Niveauanforderungen werden gesondert gefasst und nach „Mindest-, Regel- und Maximalstandards“ unterschieden. Diese Unterscheidung ist für die politischen Erwartungen der KMK maßgeblich geworden. Die entsprechenden Definitionen lauten wie folgt:

- **Mindest- oder Minimalstandards** beziehen sich auf ein definiertes Minimum an Kompetenzen, das alle Schülerinnen und Schüler zu einem vorher festgelegten Zeitpunkt erreicht haben müssen. Ein Unterschreiten des definierten Minimalniveaus am Ende des für die Überprüfung anvisierten Zeitpunktes würde mit erheblichen Schwierigkeiten dieser Schülerinnen und Schüler beim Übergang ins Berufsleben einhergehen. In einem solchen Fall ergeben sich dringend notwendige Maßnahmen der individuellen Förderung.
- **Regelstandards** beschreiben Kompetenzen, welche im „Durchschnitt“ oder „in der Regel“ von den Schülerinnen und Schülern einer Jahrgangsstufe erreicht werden sollen. Am Ende der Schullaufbahn würden Regelstandards entsprechend das Ausmaß an Kompetenz und Wissen kennzeichnen, über das z.B. ein durchschnittlicher Zehntklässler verfügen sollte.
- Als höchste Niveaustufe der Bildungsstandards können auch **Exzellenz- oder Maximalstandards** definiert werden. Die im oberen Leistungsniveau angesiedelten Kompetenzen beziehen sich darauf, was die besten Schülerinnen und Schüler der jeweiligen Jahrgangsstufen können sollten (ebd., S. 9).

Die von der KMK verabschiedeten Bildungsstandards greifen allgemeine Bildungsziele auf und legen fest, welche Kompetenzen die Schülerinnen und Schüler bis zu einer bestimmten Jahrgangsstufe im Hinblick auf „wesentliche Inhalte“ erworben haben sollen. „Die Bildungsstandards konzentrieren sich auf Kernbereiche eines Faches und beschreiben erwartete Lernergebnisse.“ Es handelt sich dabei um „eine Mischung aus Inhalts- und Output-Standards“ (ebd.), die als „Regelstandards“ zu verstehen sind (ebd.).

Mit diesem Entscheid verbinden sich bestimmte politische Erwartungen, die sich auf drei Ebenen beziehen. Die erste Ebene ist die der Systementwicklung, die vor dem Hintergrund des „Gleichheitsgebots“ beschrieben wird. Dieses Gebot wird definiert als „gleiche Bildungschancen für alle Jugendlichen, unabhängig vom Wohnort“ (ebd., S. 10). Gesellschaftliche Veränderungen wie die zunehmende Mobilität der Bevölkerung, das Verhältnis von Allgemeinbildung und Berufsbildung, gestiegene Anforderungen an die schulische Bildung, Forderungen seitens der Elternschaft und schließlich die Erwartungen der Abneh-

mer schulischer Bildung wie Hochschulen und Sozialpartner machen die Einführung von Bildungsstandards erforderlich.

Die zweite Ebene ist die der Schul- und Unterrichtsentwicklung. Sie wird unter Berücksichtigung eines Gutachtens beschrieben, das 2003 vorlag. In diesem Gutachten wird der deutschen Bildungspolitik empfohlen, dem Beispiel anderer Staaten, besonders demjenigen der erfolgreichen PISA-Länder zu folgen, und den Schulen „mehr Gestaltungsspielraum“ einzuräumen, „im Gegenzug“ dann aber auch die Steuerung durch Bildungsstandards zu intensivieren. Bildungsstandards, heißt es, können die inhaltliche Steuerung der schulischen Praxis „transparent, verbindlich und letztlich auch überprüfbar“ machen und so zur Qualitätsentwicklung „wesentlich“ beitragen (Avenarius/Ditton/Döbert et al. 2003, S. 109).

Dieser Empfehlung liegt ein Befund zur klassischen Lehrplansteuerung zugrunde, der wie folgt lautet: Die Stundentafeln an deutschen Schulen sind auch im historischen Vergleich konstant bis hoch konstant. Lehrpläne dagegen sind ständig geändert oder modifiziert worden, im internationalen Vergleich weisen sie „eine wohl einmalige Anzahl und Vielfalt aus“. Dieser Steuerungsansatz sieht auf der einen Seite eine konstante und konservative Verteilung der zur Verfügung stehenden Zeit vor. Auf der anderen Seite steht eine hohe und stete Erneuerung der inhaltlichen Vorgaben, die zu tun hat mit der „an Lehrplänen geknüpften direkten Steuerungserwartungen in Deutschland“ (ebd., S. 108). Diese Erwartung ist möglicherweise gerade die Ursache für die faktisch „eher geringe Steuerwirkung der Lehrpläne“ (ebd.). Daher wird auf dieser Ebene eine Umsteuerung, orientiert an wenigen, aber konkreten Vorgaben (ob als Kerncurricula oder Bildungsstandards) und zu entwickelnder schulbezogener Curricula, empfohlen.

Diese Konkretisierung des Auftrages durch Bildungsstandards wird bezogen auf Verfahren der Evaluation. Das heißt, dass sich die beiden allgemeinen Ziele, also „die Verbesserung der Ergebnisse und die Vergleichbarkeit der Chancen“, nicht automatisch einstellen werden. „Es bedarf hierzu einer systematischen Schulentwicklung mit interner und externer Evaluation. (...) Bildungsstandards liefern dafür die notwendigen Vergleichsmaßstäbe. Sie beschreiben erwartete Leistungen und sind somit ein Maßstab, an dem die tatsächlichen Leistungen der Schülerinnen und Schüler gemessen werden können“ (ebd., S. 10f.).

Die dritte Ebene bezieht sich auf den Unterricht. Grundsätzlich wird darauf verwiesen, dass Bildungsstandards „das Zulassen individueller Lernwege“ ermöglichen. Es geht also um die Erhöhung der Autonomie bei gleichzeitiger Veränderung der Kontrolle. Ausdrücklich heißt es: „Bildungsstandards standardisieren *nicht* die schulischen Lehr- und Lernprozesse“ (ebd., S. 11; Hervor-

hebung d. Verf.). Vielmehr definieren sie „eine normative Erwartung, auf die hin Schule erziehen und bilden soll“ (ebd.).

Die inhaltlichen Bildungsstandards der KMK sind auf Übergänge und Abschlüsse bezogen. Sie gelten für das Ende des Primarbereichs (Jahrgangsstufe 4) sowie für das Ende des Sekundarbereichs I (Jahrgangsstufe 9 und 10).²⁰² Es handelt sich um Standards, die für verschiedene Fächer und Stufen bezogen auf ein „mittleres Anforderungsniveau“ bereits vorliegen. Regelstandards orientieren sich an *Kompetenzen* und *Anforderungsbereichen*, nicht jedoch oder noch nicht an *Kompetenzstufen*. Empirisch abgesicherte Kompetenzstufenmodelle liegen, wie gesagt, derzeit kaum vor, sie sollten daher nicht voreilig ins Zentrum der Standardentwicklung gerückt werden. Die fachlichen Standards werden durch Aufgabenbeispiele veranschaulicht, „indem sie diese exemplarisch mit konkreten Inhalten verbinden“. Die Beispiele sollen deutlich machen, welche Leistung jeweils erbracht werden muss, um die Standards zu erfüllen. Es handelt sich nicht um Testformate für Abschlussprüfungen oder für Berechtigungen. Vielmehr dienen die Beispiele „zur Veranschaulichung der durch Standards definierten Zielvorgaben“ (ebd., S. 16f.).²⁰³

Mit den Aufgabenbeispielen sind Leistungsanforderungen verbunden. Für Aussagen über die „Angemessenheit, Qualität und Komplexität“ der Anforderungen stehen in den einzelnen Lehrplänen jeweils *Anforderungsbereiche* zur Verfügung, die aufgrund der Erfahrungswerte von Lehrkräften formuliert wurden. Sie stellen eine „Orientierung“ dar, die Aufschluss gibt, in welchen Bereichen „sich Leistungen der Schülerinnen und Schüler erfahrungsgemäss bewegen“ (ebd., S. 17).

Damit rücken Standards in die Nähe der Ziele in den früheren Lehrplänen. Bildungsstandards werden als eine Art Konkretisierung oder Präzisierung der Lehrpläne verstanden. Zum Verhältnis von Lehrplänen und Bildungsstandards heißt es im Begründungspapier der KMK:

²⁰² Abschluss Hauptschule bzw. Realschule.

²⁰³ „Die Aufgabenbeispiele illustrieren ausserdem eine für das jeweilige Fach charakteristische Spannweite von Aufgabentypen zur Überprüfung von Kompetenzen bzw. Standards. Die Aufgaben sind drei Anforderungsbereichen zugeordnet. Diese orientieren sich an der Beschreibung in den Einheitlichen Prüfungsanforderungen in der Abiturprüfung (EPA), an theoretischen Grundlagen der PISA-Studie und dem Gemeinsamen europäischen Referenzrahmen für Sprachen“ (Kultusministerkonferenz (KMK) 2005, S. 17).

„Beim schulischen Lernen geht es um Wissen, um Haltungen, Einstellungen, Interessen und grundlegende Fähigkeiten, die Schülerinnen und Schüler erwerben sollen. In *Lehrplänen* werden diese in Lernzielen und -inhalten aufgelistet und zeitlich angeordnet. *Bildungsstandards* hingegen arbeiten die zentralen Kompetenzbereiche heraus, die im Laufe der schulischen Ausbildung aufgebaut werden sollen. *Bildungsstandards* formulieren die zu erreichende Zielebene bis zu einer bestimmten Jahrgangsstufe bzw. bis zu einem bestimmten Abschluss, während *Lehrpläne* den Weg zur Zielerreichung beschreiben und strukturieren. *Bildungsstandards* stellen also eine auf das *Ergebnis* schulischen Lernens gerichtete Form curricularer Vorgaben dar“ (ebd., S. 17 f.; Hervorhebungen d. Verf.).

Aber auch Lehrpläne waren immer ergebnisorientiert, wenngleich nie sehr explizit und bislang kaum überprüft. In der oben zitierten Version der KMK leiten die Lehrpläne die Bildungsstandards an, die innerhalb des Rahmens der Lehrpläne zentrale Kompetenzbereiche bestimmen sollen. Sie wären damit von Kerncurricula kaum zu unterscheiden, so dass „neu“ im System der Output-Steuerung auf dieser Linie der Differenzierung eigentlich nur ist, „dass empirisch untersucht werden soll, inwieweit die Leistungserwartungen tatsächlich eingelöst worden sind“ (ebd., S. 18). „Output-Steuerung“ wäre so Überprüfung der Zielerreichung, eine Konzentration der curricularen Vorgaben auf „zentrale Kompetenzbereiche“ vorausgesetzt.

Daher kann es auf der Input-Seite weiterhin Lehr- und Rahmenpläne sowie Curricula geben, die die Lernziele und die Lerninhalte „systematisch und in ihrer zeitlichen Abfolge“ beschreiben und den Lehrkräften „Hinweise“ für das methodisch-didaktische Vorgehen und für Lernerfolgskontrollen geben (ebd.). Ausdrücklich heißt es: „Bildungsstandards sind keine Handreichungen für den Unterricht“. Auch die Aufgabenbeispiele sind nicht auf „unterrichtliche Voraussetzungen“ hin ausgewiesen worden (ebd., S. 17). Wie dann aber mehr Verbindlichkeit erreicht werden soll, ist eine noch offene Frage und ein zentrales Problem der Implementation. Die wesentliche Unterscheidung ist, folgt man den weiteren Ausführungen, diejenige zwischen einzelnen oder allen Lernzielen und Lerninhalten auf der einen, zentralen Zielen und Konzepten sowie grundlegenden Kompetenzen auf der anderen Seite.

Vier Vereinbarungen über Bildungsstandards sowie elf Vorschläge für fachliche Bildungsstandards liegen vor. Es sind dies Vereinbarungen der deutschen Bundesländer über Bildungsstandards für den Primarbereich (Jahrgangsstufe 4), für den Hauptschulbereich (Jahrgangsstufe 9) sowie zwei Vereinbarungen für den Mittleren Schulabschluss (Jahrgangsstufe 10). Fachliche Standards im Primarbereich liegen vor in den Fächern Deutsch und Mathematik. Für den Hauptschulabschluss sind dies Standards in den Fächern Deutsch, Mathematik

und für die erste Fremdsprache. Für den Mittleren Schulabschluss liegen Standards vor für Deutsch, Mathematik, die erste Fremdsprache sowie für die drei Naturwissenschaften Biologie, Chemie und Physik. Für diese Fächer wurde eine gesonderte Vereinbarung getroffen.²⁰⁴

Die Länder verpflichten sich in den Vereinbarungen, die Standards zu implementieren und anzuwenden. Im Blick auf den Primarbereich wurde beschlossen, die Bildungsstandards in den Fächern Deutsch und Mathematik „zu Beginn des Schuljahres 2005/2006 als Grundlagen der fachspezifischen Anforderungen für den Unterricht“ zu übernehmen. Die Implementation und Anwendung der Standards „betrifft insbesondere die Lehrplanarbeit, die Schulentwicklung und die Lehreraus- und -fortbildung“. Die Länder kommen überein, weitere Aufgabenbeispiele zusätzlich zu den in den Beschreibungen der Bildungsstandards vorliegenden zu entwickeln. In landesweiten bzw. länderübergreifenden Orientierungs- und Vergleichsarbeiten wird festgestellt, „in welchem Umfang die Standards erreicht werden“. Das kann am Ende der Jahrgangsstufe erfolgen, aber auch zu einem früheren Zeitpunkt, „um Interventionen zu ermöglichen“.

Zum Komplex Tests wurde zu diesem Zeitpunkt nur so viel gesagt:

„Die Standards und ihre Einhaltung werden unter Berücksichtigung der Entwicklung in den Fachwissenschaften, in der Fachdidaktik und in der Schulpraxis durch eine von den Ländern gemeinsam beauftragte wissenschaftliche Einrichtung überprüft und auf der Basis validierter Tests weiter entwickelt“ (Vereinbarung über Bildungsstandards für den Primarbereich 2004).

Dieser Passus ist in allen vier Vereinbarungen identisch. Bildungsstandards in den zwischen Dezember 2003 und 2004 verabschiedeten Fassungen sind Kerncurricula mit Aufgabensammlungen. Das Erreichen der Ziele wird mit Vergleichs- und Orientierungsarbeiten überprüft, also durch die Administration mit einem bewährten und nicht mit einem neuen Instrument. Wie weit, wo und mit welcher Zweckbindung neue Instrumente wie Tests oder ein System des Bildungsmonitorings zur Anwendung kommen, wird erst allmählich geklärt.

In der Kritik an den Richtungsentscheidungen der deutschen Bildungspolitik wird grundsätzlich in Frage gestellt, ob sich Bildungsstandards auf der Basis von Kompetenzmodellen entwickeln und definieren lassen, weil es bisher keine

²⁰⁴ <http://www.kmk.org/schul/Bildungsstandards/bildungsstandards-neu.htm>
(Stand 23. 07.06)

Entwicklungs- und Lernpsychologie des Kindes- und Jugendalters gebe, in der jene Komponenten und Stufen ermittelt worden wären, aus denen sich altersspezifische „Kompetenzmodelle“ konstruieren und entsprechende Testverfahren generieren ließen (Herrmann 2003, S. 634). Auch die Frage, wieweit sich Bildungsforschung auf politische Anforderungen einlassen dürfte oder sollte, ist kritisch gestellt worden (Bellmann 2006).

Spezifisch für Deutschland ist eine von verschiedenen Seiten vorgebrachte allgemeine Kritik am Konzept der Standards, die geprägt ist von humanistischen, kritisch-emanzipatorischen oder auch reformpädagogischen Konzepten. Darauf bauen pädagogische Erwartungen auf, die unverträglich zu sein scheinen mit „mechanistischen“ Bildungsreformen. Bildungsstandards werden in diesen Stellungnahmen in einen Gegensatz zu „Bildung“ gestellt oder in Konkurrenz zu diesem Ideal wahrgenommen. Den theoretischen Annahmen, die dem Konzept der Standards zugrunde liegen – etwa *literacy* –, wird abgesprochen, einen Beitrag zur Bildung im emphatischen Sinn zu leisten. Bildung sei Menschenbildung und die Standards würden den damit verbundenen Anspruch untergraben, weil ihnen ein lediglich basales, funktionales oder gar utilitaristisches Bildungsverständnis zugrunde liege. Der Objektivierungsanspruch von Standards stehe dem traditionellen Subjektivitäts- und Persönlichkeitsdenken von „Bildung“ entgegen. Da Bildung nicht auf messbare basale Fähigkeiten reduzierbar sei, dürfe eigentlich nicht von *Bildungsstandards* gesprochen werden (vgl. Gewerkschaft Erziehung und Wissenschaft (GEW) 2004a, S. 154).

Neben solchen Kritiken, die die gesamte Richtung der Politik betreffen, finden sich in verschiedenen Themen- und Beiheften erziehungswissenschaftlicher Zeitschriften sowie in Stellungnahmen von fachdidaktischen Gesellschaften und Lehrerorganisationen (etwa GEW²⁰⁵) auch Kontroversen, die professionspolitisch ausgerichtet sind und sich mit den neuen Strategien der Steuerung befassen. Es geht um die Konsequenzen einer Output-Orientierung für Schule und Unterricht, wobei zumeist vorausgesetzt wird, dass die technischen Probleme der Steuerung sich lösen lassen, aber pragmatische Fragen wie Ressourcen und eine Zunahme der Belastungen ohne Gegenwert zentrales Gewicht erhielten. Hier wird auch deutlich, dass Akzeptanz für die neue Politikstrategie vor dem Hintergrund verbreiteter Skepsis gewonnen werden muss.

Die Konzeption der KMK-Bildungsstandards²⁰⁶ wurde schließlich auch vor dem Hintergrund der Expertise (Klieme/Avenarius/Blum et al. 2003) kritisiert.

²⁰⁵ <http://www.gew.de/GEW-Positionen.html> (Stand 14.03.06)

²⁰⁶ Vgl. für die fächerspezifische Kritik an den KMK-Standards: Die Deutsche Schule, 96 (2004), 8. Beiheft.

So folgte man mit der Setzung von abschlussbezogenen Regelstandards gerade nicht der nachdrücklich Empfehlung der Autoren, mit den nationalen Bildungsstandards das verbindliche *Minimale*niveau schulformübergreifend festzuschreiben. Abgesehen von der politischen Dimension hängt dieser Entscheid technisch auch damit zusammen, dass zum Zeitpunkt dieser Festlegung theoretisch fundierte Kompetenzmodelle, wie sie die Expertise als wissenschaftliche Grundlage von Bildungsstandards unbedingt forderte, noch in keiner Reichweite waren. Messner (2004) stellte denn auch fest, dass es sich bei den vorliegenden KMK-Bildungsstandards lediglich um Lernzielkataloge handeln würde. Aus der Sicht der Gewerkschaft Erziehung und Wissenschaft (GEW) (2004a) fehlte schließlich ein echter Bezug zu Bildungszielen; insgesamt könne den KMK-Standards lediglich ein vorläufiger Charakter zugesprochen werden. Diese Kritik trifft insbesondere die Deutsch-Standards, während die Fremdsprachen-Standards, die sich am Gemeinsamen Europäischen Referenzrahmen (GER) orientieren, vergleichsweise gut wegkommen.

Bemängelt wird von Seiten der Gewerkschaft Erziehung und Wissenschaft (GEW) auch das Fehlen eines Gesamtkonzeptes der pädagogischen Qualitäts- und Schulentwicklung, in welches „gute“ Standards eingebettet sein müssten (Gewerkschaft Erziehung und Wissenschaft (GEW) 2004a). Kritisiert wird des Weiteren die Ausrichtung der Bildungsstandards auf die Übergänge und Abschlüsse des deutschen Schulsystems, was diese zu Selektionsinstrumenten mache und die Idee der Förderung fraglich erscheinen lasse. Der Entscheid zugunsten von *Regel*standards lasse bezweifeln, dass diese von der Mehrzahl der Hauptschülerinnen und -schüler erreichbar seien. Insgesamt fällt die Analyse der KMK-Standards und -Standardentwürfe auf der Folie der Merkmale guter Bildungsstandards nach Klieme/Avenarius/Blum et al. (2003) durch die GEW äußerst negativ aus. Dies betrifft etwa die Anzahl der Standards (zum Beispiel 112 Standards für den Hauptschulabschluss im Fach Deutsch), was der Forderung der Fokussierung zuwiderlaufe. Im Zentrum stehe die Entwicklung von Kenntnissen, statt Fertigkeiten und Fähigkeiten (Demmer/Schweitzer 2005; Helmke 2004; Lorenz 2005; Sill 2004). Auch die Überprüfbarkeit der Standards steht durch vage Formulierungen, die nicht operationalisierbar seien, in Frage.

Schließlich wird heftig kritisiert, dass in den Ländern sehr unterschiedliche Standards entwickelt werden, die ebenfalls nur Lehrpläne unter neuem Namen darstellen. Von einem nationalen Vorgehen könne so gar nicht die Rede sein (Saldern/Paulsen 2004; Gewerkschaft Erziehung und Wissenschaft (GEW) 2004b). Als irrational und nicht gerade kostenbewusst wird auch empfunden, dass verschiedene Bundesländer-Gruppierungen neben Standards auch Tests in eigener Zuständigkeit entwickeln, obwohl die KMK mit dem Institut zur Qua-

litätsentwicklung im Bildungswesen (IQB) eine zentrale Institution geschaffen hat (Saldern/Paulsen 2004).

Die Diagnose bezüglich der Implementationstätigkeiten in den Bundesländern lautete damals von Seiten der GEW (2004b, S. 201) wie folgt:

„In keinem Bundesland liegen bislang umfassende Implementationskonzepte vor. Durchgeführt werden lediglich Schulleitungsinstruktionen. Inhaltsreiche Fortbildung – vor allem auch für Lehrerinnen und Lehrer – findet nicht statt. Stattdessen werden Lehrkräfte mit Tests konfrontiert, die sie größtenteils nicht einordnen können. Die meisten Lehrkräfte wissen nicht, welche ‚Philosophie‘ und didaktische Konzeption mit Standards verbunden ist bzw. verbunden werden könnte. Auch sie verstehen Standards – gemäß deutscher Tradition – überwiegend als weiteres Selektions- und Kontrollinstrument. Diagnosefähigkeit wird dann missverstanden als die Fähigkeit, die Schülerinnen und Schüler in die vermeintlich ‚richtige‘ Schulform zu sortieren, nicht jedoch als die Fähigkeit, Lernprobleme zu erkennen und über Lernstrategien zu verfügen, die sie überwinden helfen.“

Dieser Situation soll mit dem *Institut zur Qualitätsentwicklung im Bildungswesen* (IQB) abgeholfen werden. Es wurde mit Beschluss von 4. Dezember 2003 von den Ländern als wissenschaftliche Einrichtung an der Humboldt-Universität in Berlin ins Leben gerufen, und zwar mit dem Zweck der Weiterentwicklung, Operationalisierung, Normierung und Überprüfung von Bildungsstandards; zudem soll es die Implementation der Standards wissenschaftlich begleiten. Konkret beschäftigt es sich heute mit den Teilbereichen:

- Generierung von großen Aufgabensammlungen zur Operationalisierung der Standards;
- Formulierung von Kompetenzmodellen;
- Formulierung von Vergleichsaufgaben;
- Erarbeitung von computergestützten Test-, Auswertungs- und Rückmeldesystemen;
- Bildungsmonitoring;
- Begleitforschung zur Implementation der Bildungsstandards.

Weiterentwickelt werden zum momentanen Zeitpunkt die Bildungsstandards Mathematik Sekundarstufe I (9. Jahrgangsstufe, 8. und 10. Jahrgangsstufe) und Primarbereich (3. und 4. Jahrgangsstufe), Deutsch Primarstufe sowie Englisch und Französisch Sekundarstufe I. Zu den Bildungsstandards Mathematik 9. Jahrgangsstufe existiert bereits eine Publikation (Blum/Drüke-Noe/Hartung

et al. 2006), deren explizites Ziel es ist, Visionen kompetenzorientierten Unterrichts im Sinne der Standards zu dokumentieren. Die enthaltenen Aufgabenbeispiele sollen die Standards veranschaulichen, indem sie diese exemplarisch mit konkreten Inhalten verbinden; erhofft wird dadurch die Entwicklung einer „anforderungsbezogenen Aufgabekultur verbunden mit einer Unterrichtskultur, die auf unterschiedliche Schülervoraussetzungen eingeht“ (IQB, Bildungsstandards Mathematik Sekundarstufe I)²⁰⁷.

Den fortgeschrittenen Arbeiten im Bereich der Mathematik Sekundarstufe I kommt Modellcharakter zu; demnach besteht das generelle Prozedere 1. in der Entwicklung von Testaufgaben, 2. der Pilotierung der Aufgaben, 3. einer Normierungsstudie, 4. der Definition von Skalen. Zusammen mit der Entwicklung von Kompetenzskalen werden außerdem verschiedene Rückmeldeformate für diverse Entscheidungsträger (u.a. Lehrkräfte, Schülerinnen und Schüler, Eltern) erprobt und bereitgestellt. Der Zeitplan sieht vor, dass die Skalen und Standards in den genannten Fächern bis 2008 bzw. 2009 vorliegen. Es ist vorgesehen, dass die Ländervergleiche (bislang PISA-E und IGLU-E) in Zukunft vom IQB durchgeführt werden und sich auf die Fächer Deutsch und Mathematik in der Grundschule sowie Deutsch, Mathematik, die erste Fremdsprache und die drei Naturwissenschaften in der Sekundarstufe I beziehen. Die Terminierungen orientieren sich dabei an den Durchführungszeitpunkten der internationalen Studien PISA, TIMSS und IGLU.

Neben der Weiterentwicklung der Bildungsstandards und der Bereitstellung von Testaufgaben nimmt sich das IQB inzwischen auch der Frage der Implementation an; sowohl ein Projekt zur qualitativ- als auch eines zur quantitativ-empirischen Begleitforschung sind konzipiert. Im Ersteren geht es um die Frage, was Lehrpersonen unternehmen, um die Orientierung an Bildungsstandards in die Gestaltung des Unterrichts zu integrieren; und weiter, welche Prozesse sich an Schulen ereignen, die die Arbeit mit Bildungsstandards in ihrem schulischen Alltagshandeln implementieren. Zur Beantwortung der Fragen werden in einem ersten Teilprojekt im Bereich der Forschung zur Lehrerprofessionalität implizite Orientierungen der Lehrkräfte bzw. implizite Strukturen in den Schulen rekonstruiert. Dazu sollen an mehreren Schulen in verschiedenen Bundesländern u.a. Gruppendiskussionen mit Lehrkräften durchgeführt und Konferenzen und Teamsitzungen teilnehmend beobachtet werden. Das zweite Teilprojekt fokussiert die Unterrichtsebene; Gegenstand ist der kompetenzorientierte Unterricht im Fach Französisch, der mittels Unterrichtsbeobachtung ana-

²⁰⁷ http://www.iqb.hu-berlin.de/arbbereiche/projekte?pg=p_8 (Stand 22.04.07).

lysiert wird. Die Studie ist im Grenzbereich qualitativ-rekonstruktiver Unterrichtsforschung und unterrichtsbezogener Aktionsforschung angesiedelt. Ein Projekt im Bereich quantitativ-empirischer Begleitforschung konzentriert sich auf das Innovationsmanagement von Lehrkräften, und zwar auf der theoretischen Basis des *Concern Based Adoption Model* (CBAM), einem Prozessmodell der Akzeptanz und der Umsetzung von Bildungsinnovationen durch Lehrkräfte (vgl. Loucks-Horsley 1996; Hall/Hord 1987; Hord/Rutherford/Huling-Austin et al. 1987; Loucks-Horsley/Stiegelbauer 1991).

6.1.3 Überprüfung der Leistungen: Tests und Bildungsmonitoring

Im *Bildungsbericht für Deutschland* (Avenarius/Ditton/Döbert et al. 2003) nannten sämtliche deutschen Bundesländer auf die Frage nach den bildungspolitischen Schwerpunkten Leistungserhebungen verschiedener Art (S. 280). Die Kultusministerkonferenz hatte sich 2002 auf den Beschluss beschränkt, dass bereits in der Grundschule landesweite *oder* länderübergreifende Orientierungs- und Vergleichsarbeiten zur Überprüfung der Standards durchgeführt werden sollen (KMK 2002). Inzwischen haben sämtliche Bundesländer Verfahren der Leistungsevaluation eingeführt. Die Tests finden überwiegend in Form von Vergleichsarbeiten, Orientierungsarbeiten, Lernstandserhebungen und Parallelarbeiten statt und sind dabei schwerpunktmäßig eher curriculum- oder kompetenzorientiert (Hovestadt/Kessler 2005). Problematisch ist, dass unter den verschiedenen Bezeichnungen oftmals ähnliche Initiativen verstanden werden; bringen jene unterschiedliche konzeptionelle Ansätze und Funktionen zum Ausdruck, sind sie „oft nur noch für Insider verständlich. Ihre Vielfalt erschwert der breiteren Öffentlichkeit das Verständnis“ (Avenarius/Ditton/Döbert et al. 2003, S. 268; Helmke/Hosenfeld 2003; Münder 2007). Insgesamt zeigt sich bezüglich externer Leistungsevaluationen, dass sämtliche Bundesländer Verfahren einsetzen, dies jedoch in unterschiedlicher Weise und mit Blick auf unterschiedliche Funktionen.

In Parallelarbeiten bearbeiten Parallelklassen einer oder mehrerer Schulen gleiche Testaufgaben: Die Arbeiten werden von den Lehrpersonen gemeinsam oder wechselseitig beurteilt. Damit ermöglichen sie einen Vergleich zwischen den Klassen und sollen darüber hinaus einen innerschulischen Diskurs über Lerninhalte, Unterrichtsmethoden und Benotungspraktiken auslösen. Parallelarbeiten werden von sieben Bundesländern durchgeführt, Vergleichsarbeiten finden in allen Bundesländern statt, wobei die Unterscheidung beider Evaluationsinstrumente nicht in allen Fällen eindeutig ist. Vergleichs-/Orientierungsarbeiten (auch Lernstandserhebungen) beruhen auf zentral entwickelten Aufgaben und werden landesweit oder länderübergreifend flächendeckend durchge-

führt. Sie ermöglichen den Vergleich zwischen Klassen und Schulen, die Überprüfung der Einhaltung (nationaler) Bildungsstandards und dienen der Rechenschaftslegung der Schulen. In der Regel sind die Lehrkräfte an den Schulen mit der Durchführung betraut. Eine länderübergreifende Initiative von sieben Bundesländern zu Vergleichsarbeiten in Deutsch und Mathematik existiert mit dem Projekt VERA.²⁰⁸ Neben diesen schullaufbahnbegleitenden Maßnahmen führen sämtliche Bundesländer mit Ausnahme von Rheinland-Pfalz zentrale Abschlussprüfungen durch (Eikenbusch 2007). Auch hier ist jedoch die konkrete Ausgestaltung, z. B. bezogen auf den Anteil zentral gestellter Aufgaben, einbezogener Schultypen oder Fächer sowie das Korrekturverfahren, sehr unterschiedlich (van Ackeren 2007).

In einer Analyse aus dem Jahr 2005 (Hovestadt/Kessler) konnten Rheinland-Pfalz und Bayern stellvertretend für zwei Typen in der Anwendung von Tests/Prüfungen angeführt werden. Rheinland-Pfalz wäre demnach ein Land mit zurückhaltender Anwendung von Tests. Es beteiligt sich am Projekt VERA, wobei die Ergebnisse der Vergleichsarbeiten verschiedene Funktionen haben und u.a. zur Schulempfehlung beitragen sollen. Ansonsten werden nur noch schulinterne Parallelarbeiten durchgeführt, und zwar in allen Fächern. Sie finden in der 5. und 7. Klasse statt und sollen mit der Wahl dieser Zeitpunkte vor allem Informationen für die Gestaltung des Unterrichts liefern. Sie geben Auskunft über die individuellen Lernausgangslagen bei Beginn bzw. in der Mitte einer Schullaufbahn. Vergleichsarbeiten am Ende der Laufbahn oder zentrale Abschlussprüfungen sind nicht vorgesehen. Bayern dagegen ist das Land mit der wohl häufigsten Anwendung von Evaluationsinstrumenten, wobei es keine Parallelarbeiten, sondern nur die zentral normierten Verfahren nutzt. Die Hauptfunktion liegt in der Bewertung und Normierung von Schülerleistungen und damit auch von Unterrichts- und Schulleistungen. Bayern führt gleichzeitig zu den zentralen Abschlussprüfungen auch noch individuelle Aufnahmeprüfungen für die Sekundarstufe II durch (Hovestadt/Kessler 2005, S. 10).

²⁰⁸ Vergleichsarbeiten Deutsch und Mathematik am Ende der 3. Klassenstufe. Inzwischen nehmen sämtliche 16 Bundesländer an der Erhebung teil. VERA ist ein standardisierter Leistungstest, der sich an den Gütekriterien von Schultests orientiert. Das Projekt ist an der Universität Landau angesiedelt (Koordination: Ingmar Hosenfeld und Andreas Helmke); es verfolgt u.a. folgende Ziele: Unterrichtsentwicklung; Erfassung und Verbesserung der Diagnosefähigkeit und -genauigkeit; Ermittlung von Lern- und Förderbedarf bei den Schülerinnen und Schülern; Standardsicherung und -entwicklung. Der Test orientiert sich an den KMK-Standards und Lehrplänen.

Unterschiedliche (schul-)politische Traditionen und Zielsetzungen, aber etwa auch verschiedene Ausgangslagen im Blick auf die PISA-Ergebnisse, lassen sich auch im Vergleich der von den Ländern Brandenburg und Rheinland-Pfalz getroffenen Maßnahmen aufzeigen (Nessel 2005). Beide Bundesländer stellen die Qualitätssicherung in den Mittelpunkt, die KMK-Beschlüsse erfahren dabei aber unterschiedliche Ausprägungen. Während in der Grundschule mit dem Einsatz von VERA noch recht ähnlich verfahren wird, zeigen sich deutliche Unterschiede auf der Sekundarstufe I. Rheinland-Pfalz betont die an Standards orientierte Qualitätsentwicklung in den Schulen – etwa durch die Einführung von „Erwartungshorizonten“ am Ende der 6. und 8. Klasse und schuleigener Arbeitspläne –, Brandenburg hingegen setzt eher auf die Kontrolle der Outcomes. Dies führte in Brandenburg konsequenterweise zur Einführung zentraler Abschlussprüfungen am Ende der 10. und 13. Klasse (Bieber 2007). Das Ziel, die Leistungen des Schulsystems insgesamt zu verbessern, wird mit der Verbindung von zentralen Leistungsüberprüfungen und einem Visitationssystem zu einer externen Schulevaluation angestrebt. Rheinland-Pfalz hingegen setzt auf die Vorstellung, dass seine auf Unterrichtsentwicklung ausgerichteten Konzepte der Qualitätssicherung zentrale Prüfungen überflüssig machen, und bettet die Umsetzung von Standards in ein bereits 1999 initiiertes Konzept der Schulprogrammarbeit ein (ebd., S. 24).

Thüringen, als ein weiteres Beispiel, führt seit dem Schuljahr 2003/04 landesweit eine vergleichende Leistungsmessung in Form von „Kompetenztests“ in den Klassenstufen 3 und 6 in Deutsch und Mathematik durch.²⁰⁹ Die Erarbeitung von Aufgabenbeispielen erfolgt zum Teil in Kooperation mit Bayern. Ziel der Kompetenztests ist es, diagnostische Daten zur individuellen Förderung und Weiterentwicklung des Unterrichts zu gewinnen, um damit eine systematische Integration der nationalen Bildungsstandards in das schulische Qualitätsentwicklungssystem zu gewährleisten (Mannsfeld 2004, S. 315). In Länderkooperation mit Hessen, Baden-Württemberg, Bayern, Hamburg, Nordrhein-Westfalen, Niedersachsen, Sachsen, Sachsen-Anhalt und Schleswig-Holstein werden zudem (freiwillige) Kompetenztests auf der 8. Jahrgangsstufe in Mathematik durchgeführt. Die Kompetenztests beanspruchen mit ihrem intermediären Charakter, über einen reinen Leistungsvergleich der Schülerinnen und Schüler hinauszugehen, Förderbedarf rechtzeitig zu erkennen und darauf zu reagieren.

²⁰⁹ Neu ist in der 6. Klasse die freiwillige Teilnahme an einem Kompetenztest in Englisch.

Die Umsetzung der länderübergreifenden KMK-Bildungsstandards, etwa durch die Schaffung neuer (Kern-)Lehrpläne, ebenso wie die Durchführung von Tests zur Standardsicherung wird zum gegenwärtigen Zeitpunkt länderspezifisch vorgenommen. Parallel zum Betrieb des IQB bauen diverse Länder eigene Qualitäts- und Evaluationsagenturen auf, um Instrumente für externe Evaluationen zu entwickeln, die Schulen bei der Qualitätsentwicklung zu beraten und das länderspezifische Bildungsmonitoring durchzuführen; andere Länder haben entsprechende Abteilungen innerhalb der pädagogischen Landesinstitute eingerichtet.

Am 2. Juni 2006 haben die Kultusminister eine Gesamtstrategie zum Bildungsmonitoring²¹⁰ beschlossen, die neue Akzente setzt. Dabei werden vier Instrumente für das *Bildungsmonitoring* in Aussicht gestellt:

- zentrale Überprüfung des Erreichens der Bildungsstandards in einem Ländervergleich,
- Vergleichsarbeiten in Anbindung oder Ankoppelung an die Bildungsstandards zur landesweiten Überprüfung der Leistungsfähigkeit einzelner Schulen,
- internationale Schulleistungsuntersuchungen und
- gemeinsame Bildungsberichterstattung von Bund und Ländern (Kultusministerkonferenz (KMK) 2006b, S. 1f.).

Die KMK-Bildungsstandards sollen voraussichtlich ab 2009, d.h. nach den notwendigen Validierungs- und Normierungsarbeiten, in regelmäßigen Abständen von ca. drei bis fünf Jahren länderübergreifend getestet werden. Die normierten Aufgaben können dann als gemeinsamer Referenzrahmen gelten für die Ebenen

- schulische Evaluation,
- Bildungsmonitoring in den Ländern,
- Ländervergleich und
- Bildungsberichterstattung (Kultusministerkonferenz (KMK) 2006, S. 4).

²¹⁰ Vgl. die Publikation: Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring (2006), online erhältlich unter:
http://www.kmk.org/schul/Bildungsmonitoring_Brosch%FCre_Endf.pdf
(Stand 22.04.07).

Die Überprüfung erfolgt zentral durch das IQB auf der Grundlage repräsentativer Stichproben aus den Jahrgangsstufen 3, 8 und 9, d.h. mit einem einjährigen Zeitabstand vor den Übergängen des Schulsystems.

Derzeit gilt, dass sämtliche 16 Länder im Rahmen von VERA in der 3. Jahrgangsstufe Vergleichsarbeiten durchführen. In die Vergleichsarbeiten fließen normierte, auf die Bildungsstandards bezogene Aufgaben ein. Inzwischen haben sich außerdem 15 Länder (außer Baden-Württemberg) darauf verständigt, in Zukunft Vergleichsarbeiten auch in der Klassenstufe 8 durchzuführen. Die Aufgabenentwicklung und -normierung wird zentral am IQB durchgeführt.

Mit der Teilnahme an den internationalen Schulleistungsuntersuchungen PISA, TIMSS und PIRLS/IGLU können die Bildungsstandards über sogenannte Ankeritems bezüglich Lesekompetenz, Mathematik und Naturwissenschaften an internationalen Maßstäben normiert und die Leistungsfähigkeit des deutschen Bildungswesen kann international verglichen werden. Gleichzeitig können in diesem Rahmen wesentliche Indikatoren für die Bildungsberichterstattung gewonnen werden. Auf diese Weise wird in Deutschland die bisher fehlende systematische Verknüpfung verschiedener Ebenen der Leistungsmessung/Evaluation gewährleistet.

Dazu müssen sich auch die Länder-Vergleichsarbeiten an den nationalen Bildungsstandards ausrichten. Dies kann durch eine inhaltliche Orientierung geschehen oder indem über Ankeritems zusätzlich eine empirische Koppelung dort hergestellt wird, wo Aufgaben existieren, die auf der Grundlage der Bildungsstandards normiert worden sind; Voraussetzung wäre damit, dass normierte Bildungsstandards auch für diejenigen Klassenstufen vorliegen, in denen die Vergleichsarbeiten geschrieben werden. Im Unterschied zu den internationalen Untersuchungen sind die Ergebnisse des Bildungsmonitorings für die einzelne Schule von größerer Bedeutung. Durch das *alignment* über die Ebenen hinweg können Schulen die Ergebnisse ihrer Schülerinnen und Schüler auf mehreren Ebenen vergleichen und erhalten vielfältige Informationen, die sie für ihre Qualitätsarbeit unmittelbarer nutzen können.

Die auf verschiedenen Wegen gewonnenen Bildungsindikatoren werden schließlich in einen regelmäßigen Bildungsbericht von Bund und Ländern einfließen. Um den Vergleich mit Entwicklungen in den Mitgliedstaaten der EU und der OECD zu ermöglichen, sollen, soweit sinnvoll, Anschlussfähigkeit und Kompatibilität mit internationalen Bildungsstatistiken wie *Education at a Glance* gesucht werden. Ein erster nationaler Bildungsbericht wurde 2006 vorgelegt.²¹¹

²¹¹ Bildung in Deutschland. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration. Hrsg. vom Konsortium Bildungsberichterstattung im Auftrag

Er steht „am Beginn einer auf *Dauer angelegten Berichterstattung*, die auf der Basis der amtlichen Statistik sowie ergänzender bundesweit repräsentativer Survey- und Panel-Daten [...] steuerungsrelevante Informationen zu allen Bereichen und Stufen des Bildungswesens liefern soll“ (S. 2). Der nächste Bericht mit dem Schwerpunktthema „Übergänge“ ist für 2008 geplant; die Festlegung von wechselnden Schwerpunktthemen erfolgt gemeinsam durch die KMK und den Bund.

Derzeit können verschiedenste Implementationsstrategien auf Länderebene beobachtet werden, wobei der Stand der Entwicklung durchaus unterschiedlich ist. Die am 30. Juni 2006 vom Bundestag beschlossene Föderalismusreform wird daran nichts ändern. Die Konsequenzen aus der Neufassung des Art. 91b des Grundgesetzes werden gegenwärtig in Bund-Länder-Arbeitsgruppen beraten. Betroffen sind u.a. die gesetzliche Regelung der Förderung von Wissenschaft und Forschung zwischen Bund und Ländern, das Hochschulrecht, aber auch die Verortung der Schulpolitik. Diese war in der Vergangenheit Sache der Länder und wird es auch in Zukunft sein. Ebenso ist das bisherige „Verwaltungsabkommen zwischen Bund und Ländern über die Errichtung einer gemeinsamen Kommission für Bildungsplanung“ (Bund-Länder-Abkommen) vom 25. Juni 1970 in der Fassung vom 17./21. Dezember 1990 dem neu gefassten Art. 91b GG anzupassen. Die in der darin angestrebten Form ohnehin nicht realisierbare Gemeinschaftsaufgabe „Bildungsplanung“ wird somit beendet und umfasst neu drei Elemente: Gemeinsame Feststellung der Leistungsfähigkeit des Bildungswesens im internationalen Vergleich, Bildungsberichterstattung und gemeinsame Empfehlungen. „Für Folgerungen aus diesem Zusammenwirken sind – unbeschadet eventueller gemeinsamer Empfehlungen – allein die Länder zuständig, soweit nicht der Bund konkrete Zuständigkeiten hat (außerschulische berufliche Bildung und Weiterbildung, Hochschulzulassung und Hochschulabschlüsse).“ Von Seiten des Bundes positiv ausgedrückt haben Bund und Länder „künftig die Möglichkeit, das Fundament für ergebnisorientierte Vergleichbarkeit unserer Bildungseinrichtungen zu verbessern sowie gemeinsame strategische Ziele für die Weiterentwicklung des Bildungs- und Wissenschaftssystems zu vereinbaren“ (BMBF: Chancen der Föderalismusreform nutzen, 2007²¹²).

der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland und des Bundesministeriums für Bildung und Forschung. Bielefeld: Bertelsmann 2006. <http://www.bildungsbericht.de/daten/gesamtbericht.pdf> (Stand 21.04.07).

²¹² <http://www.bmbf.de/de/1263.php> (Stand 23.04.07).

6.1.4 Aktueller Stand und Strategien der Umsetzung von Bildungsstandards

Gemäß einer Umfrage der Kultusministerkonferenz (KMK) zu den geplanten Strategien der Implementation von Bildungsstandards nannten die meisten Bundesländer drei Hauptbereiche, auf die sich die Maßnahmen zur Umsetzung konzentrieren sollen (KMK 2005, S. 19; Krapp 2004, S. 319):

- *Lehrplanarbeit*: In den meisten Ländern ist vorgesehen, Rahmenlehrpläne und Aufgaben für zentrale Prüfungen zu erarbeiten oder zu präzisieren und an die Bildungsstandards anzupassen bzw. entsprechende Kerncurricula zu erarbeiten. Bis dahin sollen die Kompetenzbeschreibungen in den Bildungsstandards die in gültigen Rahmenrichtlinien bzw. Lehrplänen enthaltenen inhaltlichen Vorgaben ergänzen, nicht jedoch ersetzen. Einige Länder entwickeln Leitlinien zur Umsetzung der Standards und darauf basierend „Erwartungshorizonte“ für bestimmte Klassenstufen der Sekundarstufe I.
- *Lehreraus- und -weiterbildung*: In den Ländern sind differenzierte Fortbildungsmaßnahmen für die Lehrkräfte und die Schulleitungen (z.B. über die Landesinstitute, zum Teil unter Mitwirkung der Mitglieder der Fachkommissionen) vorgesehen. Für die Lehrerausbildung werden in einigen Ländern die Studienseminare in den Implementationsprozess aktiv einbezogen.
- *Schul- und Unterrichtsentwicklung (Überprüfung der Standards)*: Die meisten Länder werden, zumeist über ihre Landesinstitute und in Zusammenarbeit mit dem Institut zur Qualitätsentwicklung im Bildungswesen (IQB), die Einhaltung der Standards überprüfen. Die Art und Weise der Überprüfung ist unterschiedlich. Eine Gesamtstrategie ist noch nicht erkennbar. In den Antworten dominieren zentrale Vergleichsarbeiten und Abschlussprüfungen, die Umstellung von Rahmenrichtlinien auf Kerncurricula sowie die Schulentwicklungsarbeit.

Neben den Konsequenzen für die Lehrplanarbeit und die Lehreraus- und -weiterbildung werden in den Ländern Unterstützungssysteme für die Implementation von Bildungsstandards aufgebaut, die sinnvollerweise auf vorhandene Strukturen und personelle Ressourcen zurückgreifen sollen. Eine eigentliche Implementationsstrategie ist nicht erkennbar. Es handelt sich um sehr verschiedene Maßnahmen, die gebündelt und unter einem einheitlichen Etikett kommuniziert werden müssten. Zudem sollten die Bedingungen klar formuliert sein.

Die Überprüfung der Standards auf den Ebenen Bildungsmonitoring und insbesondere Schulevaluation, aber auch die kompetenzbasierte Individualdia-

agnostik, setzen nicht nur die wissenschaftliche Entwicklung von je spezifischen Tests voraus, sondern im Sinne der Nutzung entsprechender Daten für die Schul- und Unterrichtsentwicklung auch die Etablierung von Rückmeldesystemen. Der professionelle und sachgemäße Umgang mit den entsprechenden Daten erfordert die Schulung der involvierten Lehrkräfte und Schulleitungen sowie des Schulaufsichts- und Beratungspersonals. Mit der Ablösung der traditionellen Lehrpläne durch Kerncurricula gewinnen schuleigene Curricula an Bedeutung, deren Entwicklung ebenfalls neue fachdidaktische Anforderungen an die Fachgruppen und -konferenzen stellt. Eine bedeutende Frage ist, ob die vorhandenen Ressourcen tatsächlich für die neuen Funktionen und Qualifikationsanforderungen genutzt werden können.

Im Folgenden soll entlang zweier zentraler Bereiche der Implementation ein Einblick in den Stand der Entwicklung in den deutschen Bundesländern gegeben werden. Die beiden Bereiche sind die Veränderung der Lehrplanarbeit durch Bildungsstandards auf der einen, der Aufbau von Unterstützungssystemen auf der anderen Seite. Die Darstellung ist über gewisse Strecken übersichtsmäßig, an anderen Stellen exemplarisch angelegt. Fragen etwa, inwiefern und auf welche Weise bei der Implementation von Bildungsstandards auf vorhandene Strukturen und Instanzen zurückgegriffen werden kann, können nur anhand eines Einzelfalls beantwortet werden.

Der Grundsatzbeschluss der KMK vom 24./25. Mai 2002, Bildungsstandards für ausgewählte Schnittstellen der allgemeinbildenden Schulen zu erarbeiten, traf die Bundesländer nicht gänzlich unvorbereitet. Bereits im Vorfeld dieses Beschlusses hatten einige Länder ihre Lehrpläne revidiert oder entsprechende Arbeiten in Auftrag gegeben (z.B. Thüringen, Baden-Württemberg, Brandenburg, Nordrhein-Westfalen). In vielen Ländern fand die Erarbeitung neuer Kern- oder Rahmenlehrpläne parallel zur Entwicklung der KMK-Standards statt. Inzwischen liegen neue Lehrpläne vor, die verbindliche Kernbereiche festlegen und zumindest begrifflich auf Kompetenzen rekurrieren. Diese Strategie rückt ab von den traditionellen Formulierungen der Lernziele und der dazu oft nicht passenden additiven Auflistung einzelner Inhalte, die im Fachunterricht in einem bestimmten Schuljahr gelernt werden sollen. Kerncurricula sind ungleich verbindlicher, ohne wie im englischen National Curriculum eine Prozentzahl vorzugeben. Die Verengung der Möglichkeiten führt auch dazu, eine bessere Zielsteuerung anstreben zu können.

Verschiedene Länder haben selbst explizite Standards formuliert, die beschreiben, welche Leistungen bzw. Kompetenzen von den Schülerinnen und Schülern zu einem bestimmten Zeitpunkt der Schullaufbahn erreicht werden sollen. Unterschiede zeigen sich zwischen den Ländern insbesondere im Hin-

blick darauf, ob die KMK-Standards direkt übernommen wurden, wie etwa in Sachsen ergänzend zu den neuen Lehrplänen (2004/05), ob ein Abgleich und die Anpassung der eigenen an die nationalen Vorgaben stattgefunden haben oder ob eigene Entwicklungsarbeit veranlasst wurde. Dabei sind verschiedene Wege eingeschlagen worden. Neben der Ableitung von Leistungserwartungen aus den bestehenden Curricula besteht ein Ansatz darin, Rahmenlehrpläne neu zu schreiben, Standards zu definieren und dann als Teil der Curricula Leistungserwartungen festzulegen. Das Land Brandenburg etwa weist in den neuen Rahmenlehrplänen für die Sekundarstufe I niveaudifferenzierte „Qualifikationserwartungen“ für die Jahrgangsstufe 10 aus; die Grundschullehrpläne enthalten separate Standards. In einem dritten Konzept sind Bildungsstandards auch neben den bestehenden Curricula entwickelt worden.

Klieme/Avenarius/Blum et. al. (2003) kamen angesichts dieser Ausgangslagen zu dem Befund, dass viele der in den Ländern erarbeiteten Dokumente noch eine sehr große Nähe zu herkömmlichen Lehrplänen aufweisen (S. 134). Neu ist vor allem, dass diese Arbeiten verbindliche Kerne beschreiben und durch Musteraufgaben, Orientierungsarbeiten etc. ergänzt werden, welche die Anforderungen konkretisieren und langfristig die Basis für Evaluationen von Schule und Unterricht bilden sollen (ebd.). Mehrere Lehrpläne greifen pädagogische und fachdidaktische Konzepte auf, die seit TIMSS und PISA verstärkt diskutiert werden und auch Eingang in die Entwicklung der KMK-Standards gefunden haben. Mathematik-Curricula verweisen in ihren Einleitungsabschnitten auf Unterrichtsprinzipien, wie sie auch in den *Standards and Principles* der amerikanischen NCTM beschrieben worden sind. In den Fremdsprachen sind die Niveaustufen des Gemeinsamen Europäischen Referenzrahmens wegleitend.

Deutlich wird aber auch, dass in diesen Dokumenten insgesamt recht unterschiedlich mit dem Begriff „Kompetenz“ umgegangen wird. Im Gegensatz etwa zur Definition in den neuen Bildungsplänen von Baden-Württemberg, die die Standards eher im Sinne von Schlüsselqualifikationen verwenden, geht die Expertengruppe um Klieme „von der pädagogisch-psychologischen Erkenntnis aus, dass *Kompetenzen* – verstanden als erlernte, anforderungsspezifische, Leistungs-Dispositionen – nur durch kontinuierlichen Aufbau von Wissen und Können in einem *Inhalts- und Erfahrungsbereich* entwickelt werden“ können (S. 134). Kompetenzmodelle beschreiben die Ziele, Strukturen und Ergebnisse fachlicher Lernprozesse, und erst „die Kompetenzmodelle geben den Standards eine Orientierungskraft für den Unterricht, indem sie [...] demonstrieren, welche Entwicklungs- und Niveaustufen fachliche Kompetenzen haben“ (S. 135). Sachsen verzichtet in den neuen Lehrplänen (2004/05) auf einen Bezug zu Kompetenzen und Standards, übernimmt jedoch die KMK-Bildungsstandards

in direkter Weise. Ihnen kommt die Funktion der Unterrichtsstrukturierung zu. Mit dieser Balance von Input- und Output-Steuerung soll eine Überforderung der Schulen und Lehrkräfte vermieden werden, da eine radikale, „von oben“ verordnete Umsteuerung an den Bedürfnissen der Betroffenen vorbeigehen und keine Leistungssteigerung mit sich bringen würde (Mannsfeld 2004, S. 303f.).

Auch in Thüringen formulieren die Lehrpläne Lernziele und Inhalte. Sie basieren auf „Kompetenzmodellen“ und nehmen den unterrichtlichen Prozess des Kompetenzerwerbs in den Blick, während die nationalen Bildungsstandards bestimmen, über welche Kompetenzen die Schülerinnen und Schüler bis zu einem bestimmten Zeitpunkt verfügen müssen. Demgegenüber orientieren sich die niedersächsischen Kerncurricula stark an den Kompetenzen der KMK-Standards, indem sie Kerninhalte benennen, fachspezifische Kompetenzen ausweisen und Anforderungsniveaus markieren.

Einen Überblick über die Entwicklungen insgesamt gibt der erste *Bildungsbericht für Deutschland* (Avenarius/Ditton/Döbert et al. 2003). Gemäß dem Bericht weisen die Lehrpläne der Länder gewöhnlich Standards im Sinne der am Ende einer Klassenstufe zu erreichenden Kompetenzen („Qualifikationserwartungen“, „Leistungserwartungen“ etc.) aus. Vor diesem Hintergrund gaben zum Berichtszeitpunkt sechs Länder an, mit der Entwicklung von „Standards“ für alle Fächer begonnen zu haben; die übrigen Länder arbeiteten nach eigenen Angaben an Standards für die Kernfächer. Als wegweisend wurden in den meisten Ländern die KMK-Standards genannt.

Festgestellt wird, dass sich die Länder bei der Entwicklung von Bildungsstandards als relativ innovativ erwiesen hätten. Das KMK-Konzept sei zügig aufgenommen und umgesetzt worden, jedoch gäbe es „eine große Vielfalt und Heterogenität bei der Entwicklung von Standards“ (ebd., S. 266). Einzelne Länder haben Allianzen gebildet. So haben Berlin, Brandenburg, Bremen und Mecklenburg-Vorpommern gemeinsame Rahmenlehrpläne für die Grundschule erarbeitet, die die zu erreichenden Kompetenzen als Standards ausweisen. Abgesehen davon aber scheine jedes Land seinen eigenen Weg zu gehen – eine Entwicklung, die die Etablierung nationaler Bildungsstandards erschweren würde.

Die Regelstandards der KMK sind der Beginn eines längeren Prozesses, in dessen Verlauf zumindest eine Annäherung der Inhalte und Verfahren zwischen den Ländern erwartet werden darf. Von den Ländern aufgenommen wurden offensichtlich grundsätzliche Impulse zur Ergebnisorientierung und zum Einbezug von Evaluationsergebnissen in Prozesse der Schul- und Unterrichtsentwicklung. Allerdings ist die Lehrplansteuerung immer noch gewichtig.

Wirklich schlanke Vorgaben wie z.B. in Schweden bestehen nicht. Ebenso wenig sind bislang zeitliche Flexibilisierungen absehbar, die die Schulautonomie sichtbar erhöhen würden. Dagegen ist die Entwicklung von Unterstützungssystemen weiter fortgeschritten.

Die Schulaufsicht nimmt in Deutschland traditionell die Doppelfunktion der Kontrolle und Beratung wahr. Vor diesem Hintergrund kann die Einführung externer Evaluation der Schulaufsicht die Chance bieten, die bisherige Diffusität in der Rollenzuschreibung im Sinne einer funktionalen Differenzierung zu überwinden. Allerdings muss, um Verunsicherung zu vermeiden, zugleich mit einer allfälligen Einführung einer neu gestalteten Schulinspektion eine Klärung ihrer Aufgaben, Verantwortlichkeiten sowie des zu ihrer Wahrnehmung erforderlichen Qualifikationsbedarfs erfolgen. Dabei wird zu entscheiden sein, welche Rolle der Schulaufsicht bei der Beratung und dem Controlling der durch externe Evaluation an der Einzelschule ausgelösten Folgeaktivitäten zukommen soll (Messner 2005).

Feststellbar in den verschiedenen Bundesländern ist im Moment eine Vielfalt neu entwickelter, pilotierter oder bereits eingeführter Konzepte der Aufsicht und Beratung, wenngleich anzumerken ist, dass die verschiedenenorts neu installierten bzw. geplanten Formen von Schulinspektionen oder -visitationen in anderen Varianten bereits früher existiert hatten. Verschiedene Beispiele lassen sich nennen. So ist etwa Sachsen dazu übergegangen, der Schulaufsicht vermehrt qualitätssichernde und zugleich beratende sowie unterstützende Aufgaben zuzuweisen; unterstützt wird sie dabei von der neu errichteten Sächsischen Evaluationsagentur (SEA), die zugleich die Brückenfunktion zum IQB wahrnimmt. Die SEA besitzt selber keine schulaufsichtliche und keine beratende Funktion, womit Aufsicht und Evaluation sowie Beratung und Evaluation getrennt werden; sie setzt sich im Kern aus Wissenschaftlern zusammen. Beabsichtigt ist, die Ergebnisse der Evaluation in den Schulporträts der Einzelschulen oder in anderer Form öffentlich zugänglich zu machen. Niedersachsen hat hinwieder eine von der Schulaufsicht unabhängige externe Inspektion nach niederländischem Vorbild eingerichtet. Die Inspektion nimmt unter anderem Bezug auf Ergebnisse der Selbstevaluation auf der Basis des *Orientierungsrahmens Schulqualität in Niedersachsen* (2003)²¹³ und des darauf abgestimmten SEIS-Instrumentariums der Bertelsmann Stiftung. Die Ergebnisse werden nicht im Sinne eines Rankings öffentlich gemacht, dennoch stellt die Schulinspektion ein Element der Qualitätssicherung vor dem Hintergrund eines erhöhten Wettbewerbs der Schulen dar, zusammen mit der Einführung von „open enrolment“ und

²¹³ http://cdl.niedersachsen.de/blob/images/C2677437_L20.pdf (Stand 05.03.06).

Pro-Kopf-Finanzierung. Auch Baden-Württemberg startete neben der Einführung von Instrumenten der Selbstevaluation im Schuljahr 2006/07 eine Pilotphase zur externen Evaluation.

Nach Abschluss der Pilotphase²¹⁴ mit rund 80 freiwilligen Schulen wird in Bayern seit dem Schuljahr 2005/06 schrittweise die externe Evaluation aller Schulen eingeführt. Organisatorisch zuständig für die Zusammenstellung der Teams ist die Schulaufsicht. Fachlich-inhaltlich wird die Arbeit der Teams von der *Bayrischen Qualitätsagentur* (Staatsinstitut für Schulqualität und Bildungsforschung) begleitet und koordiniert. Die Schulaufsicht wertet denn auch die Evaluationsberichte aus und trifft mit den Schulen Zielvereinbarungen. Zur Umsetzung heißt es von offizieller Seite, dass auch die bayrische Schulaufsicht „verstärkt in ihrer Rolle als Unterstützer, Berater und bisweilen Impulsgeber gefordert“ sei.²¹⁵ Auch das 2005 gegründete *Institut zur Qualitätsentwicklung* in Hessen beschäftigt sich mit der Umsetzung eines Konzeptes externer Schulevaluation; in Rheinland-Pfalz werden derzeit von der Agentur für *Qualitätssicherung, Evaluation und Selbstständigkeit von Schulen* an rund 50 freiwilligen Pilotschulen aller Schularten Instrumente und Verfahren der externen Evaluation erprobt. In Schleswig-Holstein nahmen 2004 externe Evaluationsteams unter Einschluss eines Vertreters der zuständigen Schulaufsicht ihre Arbeit auf; die Ergebnisse werden zusammen mit anderen Daten ab dem Schuljahr 2006/07 in den „Schulportraits“ im Internet publiziert und können so auch die eingeführte freie Schulwahl unterstützen.

In Nordrhein-Westfalen werden nach einer Pilotphase ab dem Schuljahr 2006/07 Schulinspektionen („Qualitätsanalysen“) durchgeführt, wobei die Inspektorenteams („Qualitätsteams“) aus speziell qualifizierten Schulaufsichtsbeamtinnen und -beamten zusammengesetzt sind. Nach dem Schulbesuch fassen die Qualitätsteams ihre Auswertungen in einem Bericht zusammen und leiten diesen an die Schulleitung und die zuständige Schulaufsicht weiter. Die Schule wertet den Bericht in ihren verschiedenen Gremien aus. Die sich daraus ergebenden Maßnahmen sind Grundlagen für die Zielvereinbarungen mit der Schulaufsicht. Ziel ist es, die Eigenverantwortung der Schulen zu stärken. Die ermittelten Informationen sollen einen wichtigen Beitrag zur schulischen Entwicklungsarbeit leisten.

²¹⁴ Zwischen Oktober 2003 und Juli 2005 wurden 80 bayerische Schulen, die sich sämtliche freiwillig gemeldet hatten, von 20 Teams evaluiert.

²¹⁵ Externe Schulevaluation.
<http://www.km.bayern.de/km/schule/qualitaetssicherung/massnahmen/schulevaluation> (Stand 03.08.06).

Das Zusammenspiel verschiedener Maßnahmen und Elemente der Qualitätssicherung und Qualifikation soll im Folgenden anhand zweier (willkürlich gewählter) Beispiele veranschaulicht werden: Brandenburg hat neu eine Schulvisitation eingeführt, die außerhalb der Schulaufsicht angesiedelt ist. Mit der Umstrukturierung der unteren Schulaufsichtsbehörden wurden in Brandenburg die regionale und schulinterne Fortbildung sowie die Beratung einschließlich der dafür zur Verfügung stehenden Ressourcen im Wesentlichen auf die staatlichen Schulämter übertragen. Die regionale Schulaufsicht ist so für die fachliche und pädagogische Entwicklung des Personals und der Schulen zuständig. Flankiert wurde diese Entwicklung seit dem Schuljahr 2002/03 durch die Zusammenfassung bisheriger Teilsysteme und die inhaltliche Neubestimmung zu einem *Beratungs- und Unterstützungssystem für die staatlichen Schulämter und Schulen* (BUSS). Das BUSS bildet zugleich das Rückgrat der regionalen Fortbildung der Lehrkräfte. BUSS umfasst *Fachberater* und *Schulberater* (*Schulentwicklungsberater* und *Evaluationsberater*), die vom pädagogischen Landesinstitut für Schule und Medien Berlin-Brandenburg (LISUM) qualifiziert und unterstützt werden.

Fach- und Schulberater, die innerhalb des BUSS tätig sind, sind besonders fähige und vom Landesinstitut qualifizierte Lehrkräfte. Den weitaus größten Anteil bilden die *Fachberater*. Sie stehen den Schulen für die Unterrichtsentwicklung, die Implementation der Rahmenlehrpläne, Lehrerfortbildung und interne Evaluation zur Verfügung. In den staatlichen Schulämtern arbeiten sie unter anderem bei der Entwicklung von Aufgabenbeispielen für schulische Prüfungen und Vergleichsarbeiten mit und unterstützen die Schulaufsicht. *Schulberater* unterstützen die Schulen unter anderem bei der Programmentwicklung und beraten sie bei der Vorbereitung, Durchführung und Auswertung von Evaluationsvorhaben. Ein Ziel besteht dabei darin, in Brandenburg die Beratergruppen auf die Implementation der Bildungsstandards zu fokussieren und die Beratungsanteile in Relation zur Fortbildung zu stärken.

Um das Verhältnis zwischen den neuen Rahmenlehrplänen und den nationalen Bildungsstandards zu klären, hat das LISUM *Implementationsbriefe* herausgegeben. Nachdem die Überprüfung der Kompatibilität beider Dokumente eine hohe Übereinstimmung gezeigt hatte, entschied man sich, die Implementation der Bildungsstandards mit der Weiterführung der Implementation der Rahmenlehrpläne zu verbinden. Der Unterricht wird in Brandenburg auf der Grundlage der Rahmenlehrpläne sowie der Schulcurricula erteilt. In der Anlage zu den Rahmenlehrplänen sind die KMK-Bildungsstandards aufgeführt. Sie sollen für die Planung von Lernprozessen und für die Überprüfung von Lernergebnissen und Kompetenzen beigezogen werden, soweit dies angesichts der Übereinstimmungen notwendig erscheint.

Im *Allgemeinem Implementationsbrief*, der die Einbeziehung der KMK-Bildungsstandards in die Rahmenlehrpläne der Sekundarstufe I erläutert, wird die allgemeine Bedeutung der Standards adressatengerecht verständlich gemacht. Es heißt hier, Bildungsstandards

- ermöglichen in der Fachkonferenzarbeit die gemeinsame Planung und Abstimmung von Zielen und Qualitätsmaßstäben auf der Schulebene,
- unterstützen in der Lerngruppe (Klasse) das Abstecken langfristig aufgebauter Lernziele,
- machen Bildungsansprüche für Schülerinnen und Schüler, Eltern und Bildungspartner transparenter und
- qualifizieren den Umgang mit Leistungen, insbesondere den Vergleich von abgesteckten Zielen und tatsächlich erreichten Lernergebnissen.

Ausgehend von diesen Funktionen ist es Aufgabe der Lehrkräfte, in den Fachkonferenzen den schuleigenen Lehrplan vor dem Hintergrund von im *Implementationsbrief* aufgeführten Prüfungsfragen mit den KMK-Bildungsstandards zu vergleichen und dort Anpassungen vorzunehmen, wo sie nötig sind. Neben dem *Allgemeinen Implementationsbrief* liegen fächerspezifische Implementationsbriefe vor, die die fachlichen Rahmenlehrpläne in ein Verhältnis zu den entsprechenden Standards setzen, die Kompetenzansätze sowie die Bildungsstandards mit den „Qualifikationserwartungen“ vergleichen und an Beispielen erläutern.

Auf die erhöhten fachdidaktischen Anforderungen an Lehrkräfte und Fachkonferenzen, die das Neben- und Miteinander von Rahmenlehrplänen, nationalen Standards und die zunehmende Bedeutung schulspezifischer Lehrpläne mit sich bringen, wird in Brandenburg neben dem Beratungs- und Unterstützungssystem BUSS auch mit neuen Schwerpunktsetzungen in der Lehrerfortbildung reagiert. Vor der Einrichtung von BUSS hatte sich die Fortbildung im Wesentlichen auf fächerübergreifende, fächerverbindende und didaktisch-methodische Fragestellungen konzentriert (Bericht der Landesregierung an den Bildungsausschuss des Landtags zur Lehrkräftefortbildung 2004, S. 13). Mit der Installierung von BUSS erfolgte eine inhaltliche Umsteuerung, die sich in neuen Schwerpunkten und einer Verzahnung von bildungspolitischen Inhalten, Qualifizierungsprogrammen und Umsetzungsstrategien niederschlägt.

Im Einzelnen sieht das so aus: Für jedes Schuljahr werden durch das Ministerium für Bildung, Jugend und Sport bildungspolitische Schwerpunkte festgelegt, die mit Fortbildungsmaßnahmen unterstützt werden und für die Schulämter verbindlichen Charakter haben. Der Fokus liegt, auch im Hinblick auf An-

förderungen der Implementation von Standards und neuen Rahmenlehrplänen, auf den Fächern und Lernbereichen der Schulstufen und ihrer Didaktik. Auf diese Implementationsanforderungen hin werden auch die Fachberater ausgebildet. Weitere Schwerpunkte der Fortbildung, auf die hin die Berater qualifiziert werden, bilden unter anderem die Bereiche Lernstandsdiagnostik und schulinterne Evaluation.

Mit der Berufung von *Schulentwicklungs-* und *Evaluationsberatern* wurde der Akzent auf die Qualitätsentwicklung der Schulen als eigenständige Systeme gerichtet. Dahinter steht der Gedanke, dass Qualitätsentwicklung von Schulen in hohem Maß davon abhängt, dass sich das gesamte Kollegium einer Schule Ziele setzt und eine Kultur der schulinternen Evaluation und Selbstreflexion entwickelt (ebd., S. 15). Das gelingt nur, wenn genügend Unterstützung zur Verfügung steht. Für die einzelne Schule stellen Forderungen nach Transparenz, Zielsteuerung und Rechenschaftslegung zusätzliche Anstrengungen dar, die nach eigenen Formen der Fortbildung und Beratung verlangen, wenn sie nachhaltig etabliert werden sollen.

Als ein zweites Beispiel für die Umsetzung und Sicherung von Bildungsstandards in den Schulen kann auf Schleswig-Holstein verwiesen werden. Hier liegen die strategischen und politischen Akzente etwas anders als in Brandenburg, gleichzeitig ergeben sich aber durchaus Überschneidungen. Auch die Philosophie ist weitgehend identisch: Zielklarheit für den Unterricht, Kontrolle der Zielerreichung durch (externe) Schülerleistungstests sowie Unterstützungsangebote für eine an Standards orientierte, die Schülerinnen und Schüler aktivierende Unterrichtsgestaltung gehören zu „den Schlüsselinnovationen in diesem Jahrzehnt“ (Institut zur Qualitätsentwicklung an Schulen Schleswig-Holstein (IQSH) 2005, S. 5).

Die gestärkte Eigenverantwortung der Einzelschulen verlangt in Schleswig-Holstein die Überprüfung der Leistungen der Schulen an klaren Standards sowie die Möglichkeit für die Eltern, sich über die Qualität der Schulen ins Bild zu setzen. Zu diesem Zweck werden externe Evaluationen durchgeführt (Externe Evaluation im Team EVIT), deren Ergebnisse neben anderen Informationen in die im Internet zugänglichen „Schulportraits“ eingehen. Die erhöhte Transparenz und Rechenschaftspflicht wird von Unterstützungs- und Qualifizierungsmaßnahmen begleitet. Wie in Brandenburg stellt die Einführung der Fachkonferenzleiterinnen und -leiter in die Arbeit mit den Bildungsstandards ein wichtiges Element bezüglich der Standards-Implementation dar. Ein spezielles Gewicht erhält hier zusätzlich die Qualifizierung von Führungskräften.

Die Öffnung der Lehrerfortbildung für die Implementation von Bildungsstandards bildet ebenfalls einen Schwerpunkt in Schleswig-Holstein. Ein Beispiel ist die Einführung der Leiterinnen und Leiter von Fachkonferenzen in die Arbeit

mit Standards. Ausgehend von einem Konzept, das Klarheit über die Anforderungen (Standards) mit einer gestärkten Eigenverantwortung der Schulen verbindet, werden neben den genannten externen Evaluationen folgende Maßnahmen getroffen:

- Deutliche Aufwertung der Landesfachberater als Impulsgeber für die Fachkonferenzarbeit und zur Erarbeitung fachspezifischer Fortbildungsangebote entlang den Standards,
- Hilfestellungen und Qualifizierungsangebote für die Leiterinnen und Leiter der Fachkonferenzen und
- Qualifizierungsangebote für Führungskräfte zur Frage der Unterrichtsentwicklung einschließlich der Fachkonferenzarbeit (ebd.).

Schleswig-Holstein hat seine Lehrpläne für die Grundschule und die Sekundarstufe I bereits 1997 revidiert. Seit dem Schuljahr 2004/05 arbeiten die Schulen zusätzlich mit den KMK-Standards. Der Beitrag, den die nationalen Bildungsstandards für den Unterricht leisten können, besteht darin, dass

- die Lehrkräfte sie zur Unterrichtsplanung, für eine differenzierte Diagnose des Leistungsstands ihrer Schülerinnen und Schüler sowie zur Evaluation der Unterrichtsergebnisse ihrer Klasse und an ihrer Schule nutzen können,
- in den Fachkonferenzen eine fachdidaktische Diskussion über Kompetenzen, geeignete Aufgaben für den Unterricht und Tests sowie über wirksame Fördermaßnahmen im Hinblick auf zu erreichende Standards geführt wird und
- die Arbeit an einem schulinternen Curriculum, das konkrete Umsetzungsziele für die einzelnen Fächer und Jahrgänge definiert, aufgenommen wird (ebd., S. 7).

Die tragende Rolle bei der klassenübergreifenden Einführung von Bildungsstandards kommt in diesem Konzept den Fachkonferenzen zu. Zur Unterstützung von deren Arbeit werden für die Leiterinnen und Leiter auch regionale „Qualitätsforen zu den Bildungsstandards“ angeboten, die für eine systematische Qualifizierung sorgen sollen. Diese Foren tagen zweimal jährlich. Sie sollen vor allem dazu beitragen, dass die Lehrkräfte in den Schulen über die Auseinandersetzung mit Aufgaben für den Unterricht lernen, welche Kompetenzen und Anforderungsbereiche für das jeweilige Fach wichtig sind, welche Differenzierungsmöglichkeiten es gibt und wie der Unterricht gestaltet werden muss,

um zu einem Aufbau von fachlichen Kompetenzen wirkungsvoll beitragen zu können (ebd., S. 8). Themen der Foren sind und waren unter anderem

- Entwicklung kompetenzorientierter Aufgaben für den Unterricht,
- Aufbau von Kompetenzen über die Jahre hinweg,
- Erstellung, Auswertung und Interpretation von standardorientierten Prüfungen/Tests,
- VERA: Vergleichsarbeiten für den Unterricht nutzen.

Einen anderen Schwerpunkt bildet die Qualifizierung von Führungskräften. In Anlehnung an die „Kieler Thesen“²¹⁶, gemäß welchen Unterrichtsentwicklung eine zentrale Führungsaufgabe darstellt, hat Schleswig-Holstein ein neues, systemisches Fortbildungskonzept eingeführt. Adressaten sind Lehrerinnen und Lehrer (inklusive der Leiterinnen und Leiter von Abteilungs-, Stufen- und Fachkonferenzen), die sich für Führungsaufgaben im Rahmen von Schulleitung und Schulentwicklung interessieren, sowie Schulleiterinnen und Schulleiter. Letzteren wird ein maßgeblicher Beitrag zum Erfolg von Prozessen der Qualitätsentwicklung in den Schulen zugesprochen. Daher soll ihre Managementfunktion von Verwaltungs- und Unterrichtsaufgaben abgehoben und innerhalb der eigenverantwortlichen Schule gestärkt werden. Zu ihren Kernaufgaben gehören Planungsaufgaben und Zielvereinbarungen, die Schaffung von Führungs- und Kooperationsstrukturen, die Zusammenarbeit mit Schulpartnern, die Förderung und Beurteilung der Kolleginnen und Kollegen sowie die Sicherung der Unterrichtsqualität.

6.2 Österreich²¹⁷

6.2.1 Diskussionen und Problemwahrnehmungen

Ähnlich wie für Deutschland bezüglich der vergangenen Jahrzehnte von einem Reformstau im Bildungswesen die Rede ist (Döbert/Klieme/Sroka 2004, S. 99), trifft auf die vergangene Dekade in Österreich die Feststellung einer kleinschrittigen Bildungspolitik zu. Die resultierenden Reformverzögerungen betreffen

²¹⁶ Kieler Thesen, vorgestellt auf dem Kieler Schulleitersymposium 24./25.09.2004: <http://www.infokumi.lernnetz.de/materials/jxRMTfH.pdf> (Stand 05.03.06).

²¹⁷ Die Darstellung zu Österreich rekurriert im Wesentlichen auf Lucyshyn (2006a, 2006b), ergänzt durch die Perspektive von Specht (2006).

besonders eine anstehende Strukturreform der Sekundarstufe I (Haider/Eder/Specht/Spiel 2003). Reformen der neunziger Jahre bezogen sich schwerpunktmäßig auf eine Erweiterung der Autonomie der Schulen, sowohl im pädagogischen Bereich (Lehrplan '99) als auch organisatorisch (Dienstrecht etc.) und auf innere Schulentwicklung. Vor diesem Hintergrund ist Bewegung in die Diskussion über die Qualitätssicherung und -entwicklung der österreichischen Schulen gekommen und auch das Thema einer stärkeren Kontrolle durch Rechenschaftslegung hat an Bedeutung gewonnen. Eine Intensivierung erfuhren diese Fragen, als das Land mit TIMSS und dann mit PISA erstmals an internationalen Vergleichsstudien über Schulleistungen teilgenommen hatte (Eder/Thonhauser 2002, S. 373).

Im Zusammenhang mit dem Konzept eines neuen Lehrplans für die Schulen der Sekundarstufe I ist erstmals die bildungspolitische Absicht zur Diskussion gestellt worden, Schulen im Rahmen der Arbeit an ihren Schulprogrammen zur (Selbst-)Evaluation zu verpflichten. Im Bundesministerium für Bildung, Wissenschaft und Kultur wurde mit „Qualität im Schulwesen“ (QIS) ein Projekt installiert, das die Schulen, insbesondere die Lehrkräfte, bei der Erfüllung der neuen bzw. verstärkten Ansprüche der Qualitätssicherung auf freiwilliger Basis unterstützen soll. Die auf der elektronischen Plattform laufend ergänzten Materialien (Erhebungsinstrumente, Hinweise für die Auswertung etc.) erfüllen einerseits eine motivierende Funktion, indem sie Schulen Mut für die Arbeit am Schulprogramm machen, andererseits eine kommunikative Funktion, indem sie die Diskussion und den Erfahrungsaustausch der Schulen untereinander fördern sollen. Gleichzeitig ist man sich bewusst geworden, dass unter diesen neuen Gegebenheiten auf die Schulaufsicht eine neue Rolle zukommt.

Hatten die Ansätze der standortbezogenen Schul- und Qualitätsentwicklung der neunziger Jahre in erster Linie organisationale Aspekte und Prozesse fokussiert, lenkten die internationalen Vergleichsstudien bezüglich Bildungssteuerung und Qualitätssicherung den Blick auf „Standards“ und Ergebnisse des schulischen Lernens. Bereits im Regierungsprogramm 2000 stand die Erarbeitung von „Leistungsstandards“ im Themenfeld *Qualitätssicherung und Qualitätsentwicklung* an oberster Stelle, es erfolgte jedoch zuerst ein relativ offener Umsetzungsauftrag, das heißt, politische Rahmenfestlegungen fehlten vorerst ganz.²¹⁸ Der

²¹⁸ Die Stadt Wien hat bereits im Jahr 2000 in einem Parallelprojekt zu den nationalen Bildungsstandards Standards entwickelt (vgl. Weihs-Dengg 2004). Sie sind abrufbar unter:

<http://www.schulen.wien.at/schulen/999999/Dokumente/WIENER%20BILDUNGSSSTANDARDS.pdf> (Stand 21.04.07). Die Wiener Bildungsstandards unterschei-

politische Wille aber war – und ist – deutlich erkennbar. Allein die Kosten des Bildungswesens legen Maßnahmen zur Qualitätssicherung nahe.

Österreich schnitt in PISA 2000 im Gegensatz zu Deutschland und der Schweiz in sämtlichen drei Domänen deutlich über dem OECD-Mittelwert ab, einen PISA-Schock wie in Deutschland gab es nicht. Auf der anderen Seite sind Mängel auch in Österreich unübersehbar und beeinflussen die politische Situation. Ein zu großer Teil der Jugendlichen verlässt die obligatorische Schule ohne über die grundlegenden Kompetenzen zu verfügen. Schulen unterscheiden sich in hohem Maße in ihren Fähigkeiten, Schülerinnen und Schüler zu fördern. Damit hängen teilweise eklatante Ungerechtigkeiten der Leistungsbeurteilung, der Zuweisung zu Bildungsgängen und der Vergabe von Abschlüssen und Berechtigungen zusammen. Eine unterschiedliche regionale Entwicklung der Allgemeinbildenden Höheren Schule (AHS) zwischen Stadt und Land führt dazu, dass in ländlichen Schulen für die gleiche Berechtigung erheblich mehr geleistet werden muss als in städtischen (Eder/Thonhauser 2002, S. 378). Heftige politische Debatten löste die Publikation der PISA-Erhebungen von 2003 aus: Insgesamt verloren die österreichischen Schülerinnen und Schüler in sämtlichen Kompetenzbereichen mehrere Rangplätze. Schülerinnen und Schüler aus Allgemeinbildenden Höheren Schulen konnten ihr Niveau in etwa halten, während Schülerinnen und Schüler aus berufsbildenden Schulen stark absackten. Aus diesen Befunden und aus der Tatsache, dass das österreichische Schulwesen bis jetzt keine objektive Ergebniskontrolle oder für alle Schulen eines Typs gemeinsame und verbindliche Abschlussprüfung kennt, ergeben sich insbesondere drei Problemfelder, bei deren Bearbeitung Bildungsstandards eine zentrale Rolle zukommen sollte: Leistungsförderung, Chancengleichheit und Gerechtigkeit im Schulwesen (Freudenthaler/Specht 2005, S. 6; Haider/Eder/Specht/Spiel 2003).

6.2.2 Konzeption und Funktion von Bildungsstandards

Kurz nach der Veröffentlichung des Regierungsprogramms 2000 wurden unter ministerieller Leitung Arbeitsgruppen (unter fast vollständiger Abstinenz der österreichischen Erziehungswissenschaftler) eingesetzt, die sich mit der Ausar-

den sich von den Bundesstandards in folgenden Punkten: Die Standards sind von der 5. bis 8. Schulstufe aufbauend formuliert; sie orientieren sich in stärkerem Maße an den gültigen Lehrplänen; sie weisen keine Abstufung in Leistungsniveaus auf; sie beinhalten zusätzlich fächerübergreifende Standards; konkrete Beispielaufgaben werden nicht beigegeben, um den Handlungsspielraum der Lehrerinnen und Lehrer nicht einzuschränken.

beitung der angekündigten Leistungsstandards auseinander setzten und erste Entwürfe vorbereiteten (Specht 2006).²¹⁹ Erst im Frühjahr 2003, nach dem Erscheinen der Expertise von Klieme/Avenarius/Blum et al. (2003), setzte die Bundesministerin eine „Zukunftskommission“ mit dem Auftrag ein, ein umfassendes Konzept von Strategien und Vorschlägen zur Reform und Weiterentwicklung des österreichischen Schulsystems vorzulegen, wobei die Frage der Präzisierung und Konkretisierung von Bildungsstandards einen besonderen Stellenwert einnahm.

Festgehalten wurden drei inhaltliche Themenschwerpunkte, die der Kommission als Minimalziele vorgegeben wurden (Huber/Späni/Schmellentin/Criblez 2006, S. 48):

- (1) Ausarbeiten von Vorschlägen für die Implementierung eines umfassenden Systems für ein nationales Qualitätsmanagement,
- (2) Vorschläge von Maßnahmen zur permanenten Verbesserung des Unterrichts und
- (3) Entwurf von Vorschlägen zur Stärkung der pädagogischen Aufgaben von Lehrerinnen und Lehrern und Schulen (Haider/Eder/Specht/Spiel 2003, S. 7).

Implizit war die Vorstellung vorhanden, die Kommission könne die Bildungsstandards innerhalb eines halben Jahres selbst entwickeln. Bereits im Oktober 2003 legte die Kommission in einer Studie (Haider/Eder/Specht/Spiel 2003) ein Reformkonzept vor, das fünf Prinzipien als zentral erachtet:

- systematisches Qualitätsmanagement an Schulen,
- mehr Transparenz durch Systemmonitoring,
- mehr Autonomie und Eigenverantwortung der Schulen,
- Professionalisierung des lehrenden Personals und
- deutlich mehr Ressourcen für Unterstützungssysteme sowie für Forschung und Entwicklung.²²⁰

Die fünf Prinzipien wurden in sieben „Handlungsbereiche und Reformmaßnahmen“ übersetzt, wobei die Einführung von Kompetenzstandards als Maß-

²¹⁹ Die jeweils neuesten Versionen der Bildungsstandards finden sich unter www.gemeinsamlernen.at (Stand 19. 07. 06).

²²⁰ Siehe dazu auch <http://www.klassezukunft.at/> (Stand 08. 09. 06).

nahme der Qualitätssicherung figuriert. Mit der Einführung von Kompetenzstandards sollten drei Ziele verfolgt werden:

- Das Ergebnis der Kompetenzprüfung soll zusammen mit dem Urteil der unterrichtenden Lehrpersonen die Grundlage für die Berechtigungsvergabe bilden.
- Das Ergebnis soll den einzelnen Schulen als Rückmeldung und zur weiteren Qualitätsentwicklung dienen.
- Das Ergebnis soll in Form einer anonymisierten Stichprobe Aufschluss über das Niveau der Kompetenzerreichung im Gesamtsystem geben.

Der Bericht empfahl, *Minimalstandards* zu definieren, die sich auf ausgewählte und grundlegende Kompetenzen beziehen, deren gesicherter Erwerb von besonderer Bedeutung ist. Die Tests sollten zentral entwickelt, jedoch an den Schulen von dazu eigens ausgebildeten Lehrerinnen und Lehrern durchgeführt werden. Die Daten sollten an der Schule verbleiben, abgesehen von solchen, die stichprobenweise zum Zweck eines Systemmonitorings anonymisiert weitergeleitet würden.

Die Empfehlungen der „Zukunftskommission“ wurden in der Folge nur entschärft umgesetzt, da sie einen pädagogischen Kulturbruch implizierten und als politisch nicht umsetzbar eingestuft wurden. Genauer: Die Empfehlungen wurden gar nicht erst abgewartet. Im Sommer 2002 waren von den Arbeitsgruppen erste Versionen von Standards für die Schnittstellen im Schulwesen für Deutsch, Mathematik (4. und 8. Schulstufe) und Englisch (8. Schulstufe) präsentiert worden. Und das Bundesministerium für Bildung, Wissenschaft und Kultur (BMBWK) hatte inzwischen eine Projektleitung eingerichtet, die von einer Steuerungsgruppe, bestehend aus Vertretern der Bildungspolitik, der Administration, der Bildungswissenschaft und der Schulpraxis, beraten wurde.

Aufgrund eines Mangels an Erfahrungen (nicht lediglich in Österreich, sondern in den hier zur Diskussion stehenden Ländern generell) und Know-how in Theorie, Konstruktion und Umsetzung von Bildungsstandards wurde vorerst nach Lösungen und Modellen gesucht, die mit den bestehenden Traditionen der Schule möglichst wenig in Spannung stehen. Entsprechend vage waren die Vorgaben, mit denen die zuerst eingerichteten Arbeitsgruppen jene ersten Versionen erarbeiteten. Die inhaltliche Orientierung erfolgte eng entlang den Vorgaben der geltenden Lehrpläne, aus welchen stufen- und fachspezifische Kernkompetenzen abgeleitet wurden. Überwiegend als *Regelstandards* formuliert, sollten sie in erster Linie eine Orientierungsfunktion für die Lehrkräfte erfüllen und ein Mittel zur Sicherstellung der Vergleichbarkeit und Einheitlichkeit des Schulsystems bieten. Die Standards wurden gelegentlich mit Beispielen illustriert, je-

doch nicht als Testitems vorgelegt, um die Sorge eines *teaching to the test* erst gar nicht aufkommen zu lassen. Die Überprüfung der Standards sollte schließlich durch die Lehrerinnen und Lehrer selbst geschehen.

In einem Review der Entwürfe im Frühjahr 2003 (Specht 2003) wurden die einzelnen Entwürfe kritisch geprüft, wurde die Inkonsistenz zwischen den Fächern herausgearbeitet und auf die Klärung einiger offener Fragen gedrungen. Dazu gehörte die Frage nach der prinzipiellen Funktion und dem Stellenwert bezüglich Qualitätssicherung und Steuerung der österreichischen Schulen, dem rechtlichen Status und dem Entwicklungszusammenhang zwischen Lehrplänen und Bildungsstandards. Als die Expertise von Klieme/Avenarius/Blum et al. (2003) erschien, war die Entwicklung von Standards also bereits in vollem Gang, gleichwohl trug die Expertise im Folgenden zur theoretischen und methodischen Präzisierung der Arbeiten bei.

Die auf der Basis des Reviews teilweise überarbeiteten Entwürfe wurden im Schuljahr 2003/04 in einer Pilotphase I an 18 Schulen der 8. Schulstufe und an mehr als 30 Volksschulen erprobt. Diese Phase wurde bezogen auf die 8. Schulstufe vom Zentrum für Schulentwicklung in Graz evaluiert (vgl. Freudenthaler/Specht 2005). Die Ergebnisse dieser Evaluation zeigten, dass die Lehrpersonen in einem gewissen Maß eine aufgeschlossen-kritische Einstellung gegenüber Bildungsstandards aufweisen, diese Innovation aber mit zahlreichen Unklarheiten, Ängsten und Befürchtungen (Rankings, Bildungsstandards als Kontroll- und Disziplinierungsinstrument) assoziieren. Dies hatte auch damit zu tun, dass ein Großteil der Lehrkräfte sich unzureichend unterstützt fühlte, wobei insbesondere bei den Hauptschullehrkräften ein großer Bedarf an zusätzlichen Informationen vorhanden war.

Zwar scheint vielen Lehrerinnen und Lehrern aus den Standards-Unterlagen relativ klar hervorgegangen zu sein, welche Kompetenzen von den Schülerinnen und Schülern erworben werden sollten, große Probleme zeigten sich aber bezüglich der konkreten Anwendung der Standards im Unterricht. Die fachliche Angemessenheit der Standards wurde, mit Unterschieden zwischen den Fächern, insgesamt positiv bewertet, wobei besonders von Seiten der Hauptschullehrkräfte unrealistisch hohe Anforderungen der Regelstandards bezogen auf ihre Schülerpopulationen moniert wurden. Obwohl die Standards der Mehrzahl der Lehrkräfte nützliche diagnostische Hilfsmittel anboten, wurde deren Orientierungsfunktion im Zusammenhang mit Leistungsbeurteilung eher kritisch eingeschätzt.

Auf der Basis der Rückmeldungen der Pilotschul-Lehrkräfte erfolgten Adaptierungen der Bildungsstandards unter Einbezug der universitären Fachdidaktik sowie von Praktikerinnen und Praktikern. Diese Versionen werden, mit Bei-

spielen versehen, in einer zweiten Pilotphase ab dem Schuljahr 2004/05 über zwei Jahre an ca. 140 ausgewählten Schulen im Unterricht erprobt. Damit soll einerseits die Erfahrungsbasis nochmals erweitert und eine Grundlage für die weitere Verbesserung der schriftlichen Unterlagen geschaffen werden, andererseits aber soll diese Phase bereits zur Entwicklung erweiterter Kompetenzen an den Schulen beitragen. Nach dem ersten Jahr der Pilotphase II, in der in erster Linie die Erprobung von Aufgabenbeispielen erfolgte, hat eine erneute Fragebogenstudie (Freudenthaler/Specht 2006) an rund 140 Pilotschulen eine gemischte Gesamtbilanz zutage gefördert: „Zwar bewertet eine Mehrheit die Pilotarbeit als eine ‚interessante Erfahrung‘, aber nur eine kleine Minderheit hat den Eindruck, ‚Wichtiges gelernt‘ zu haben oder ‚beruflich weiter‘ gekommen zu sein. [...] Diese Bilanzierung, die im Rahmen der Pilotphase I in gleicher Weise erhoben wurde, hat sich gegenüber damals sogar noch etwas verschlechtert“ (S. 53). Deutlich verbessert hat sich aber die Einschätzung der kommunikativen und unterstützenden Rahmenbedingungen. Die prinzipiellen Ziele der Bildungsstandards und deren potenzieller Nutzen scheinen den Lehrkräften klarer geworden zu sein; ermutigend ist auch das Ergebnis, dass die Arbeit an und mit Bildungsstandards offenbar stimulierend auf die Zusammenarbeit in Kollegien gewirkt hat. Als von zentraler Bedeutung haben sich die personalen und sozialen Faktoren erwiesen: Die Auswahl und Qualifikation der Betreuungspersonen und Multiplikatorinnen und Multiplikatoren erweist sich damit auch bezogen auf gruppenspezifische Prozesse der Meinungsbildung als mit ausschlaggebend. Schließlich ließen sich aus der Analyse zur Struktur der Beziehungen zwischen den Bewertungsdimensionen folgende wichtige Punkte im Hinblick auf Implementationsprozesse extrahieren:

- Von hoher Relevanz ist die Qualität der Information und Kommunikation von der Projektleitung zur Praxis;
- gerade die kritischen Rückmeldungen erweisen sich als Indikatoren konstruktiver Mitwirkung;
- kollegiale Zusammenarbeit scheint dort maßgeblich mit dazu beigetragen zu haben, wo die Arbeit mit den Standards zu wahrnehmbaren Verbesserungen im Unterricht führte.

Die verbindlichen Bildungsstandards als Regelstandards und die notwendigen gesetzlichen Regelungen für ihre Überprüfung auf der 4. und 8. Schulstufe werden nach Abschluss der Pilotphase II auf dem Verordnungsweg vom BMBWK kundgemacht. Dem Entscheid für Mindeststandards, wie von der „Zukunftskommission“ vorgeschlagen, stand die Differenzierung in Schulfor-

men und Leistungsgruppen entgegen, das heißt die Notwendigkeit, die Standards in diesem Fall tief genug für die unteren Leistungsgruppen der Hauptschule ansetzen zu müssen. Die österreichischen Bildungsstandards haben gemäß offizieller Bestimmung Orientierungscharakter und legen fest, welche Kompetenzen Schülerinnen und Schüler bis zu einer bestimmten Schulstufe in Bezug auf wesentliche Inhalte erworben haben sollen. Sie drücken eine normative Erwartung aus, auf die die Schule hinarbeiten soll; sie beeinflussen den Unterricht indirekt durch einen pädagogischen Orientierungsrahmen und den Blick auf Lernergebnisse (outcomes) (Lucyshyn 2006a).

Bildungsstandards werden in Österreich nicht primär als Kontroll- und Selektionsinstrumente vorgesehen, sondern sollen als Hilfsmittel für die *Selbstbewertung und Orientierung von Schulen und Lehrkräften* eingesetzt werden und Veränderungen des Unterrichts stimulieren (ebd.). Die Kompetenzbereiche des jeweiligen Fachs werden in einem Kompetenzmodell beschrieben und davon ausgehend Standards formuliert. Die Kompetenzen werden so konkret beschrieben, dass sie in Aufgabenstellungen umgesetzt und mittels Testverfahren erfasst werden können. Die Aufgabenbeispiele veranschaulichen die fachlichen Standards; dabei muss der Bezug zum Lehrplan gegeben sein. Die Beispiele sind nicht als Testformate für Abschlussprüfungen oder Berechtigungen gedacht. Spätere Tests zur Überprüfung der Bildungsstandards werden auf der Basis der Aufgabenbeispiele erstellt (ebd.). Im Gegensatz zu den Nachbarländern werden die Testitems erst in einem letzten Schritt konstruiert und erprobt. Zur Debatte steht auch die Möglichkeit eines Austauschs von Itempools mit Deutschland und der Schweiz.

Zwischen den verschiedenen Akteuren und Lagern im Bildungsbereich herrscht weitgehend Einigkeit in Bezug auf die oben genannten Problemfelder – ungleiche Verteilungen der Schülerleistungen zwischen Schulen, Ungerechtigkeiten der Leistungsbeurteilung, mangelnde Chancengleichheit – und den damit gegebenen Handlungsbedarf. Auch herrscht Konsens über die Notwendigkeit, das Schulwesen weiterzuentwickeln. Der Nutzen und die Notwendigkeit einer Einführung von Bildungsstandards, und so einer Outputorientierung des Bildungssystems als Antwort auf vorhandene Problemlagen, sind aber umstritten und werden zum Teil heftig kritisiert.

An den österreichischen Standards im Spezifischen wird zunächst und grundsätzlich kritisiert, dass sie – wie in Deutschland – an den Schnittstellen angesetzt werden. Weiter wird moniert, dass Unklarheit herrsche bezüglich ihrer Rolle in einem Gesamtkonzept schulischer Entwicklung und Steuerung, dies auch vor dem Hintergrund widersprüchlicher Botschaften hinsichtlich des Umgangs mit Testdaten (Gruber 2004). Es fehle auch an Gewissheit bezüglich der

Wirkungen, die Standards und Tests im komplexen (sozialen) System auslösen werden (Altrichter/Schratz 2004; Gruber 2004). Die bekannten Gefahren und negativen Auswirkungen von Standards und Tests in outputorientierten Systemen, wie sie sich in anderen (angelsächsischen) Ländern zeigen, würden in der österreichischen Standards-Konzeption zu wenig berücksichtigt. Weiter wird befürchtet, dass im Fall von leistungsschwachen Schülerinnen und Schülern der gesamte Unterricht (statt wie vorgesehen zwei Drittel der Unterrichtszeit) für die Vermittlung standardbasierter Kompetenzen aufgewendet werden müsste und kein Raum für andere wertvolle Inhalte, fächerübergreifende Kompetenzen und für individuelle Lernprozesse übrig bleibe.

Neuweg (2004) wertet die Möglichkeit, sich im Rahmen der Auseinandersetzung mit Bildungsstandards des Problems der Auswahl und Legitimation von Lernzielen wieder bewusster zu werden, als positiv, moniert aber die Blindheit gegenüber dem Problem, wer diese Standards in Prozessen welcher Art setzen soll. Es geht ihm also um die mangelnde Verfahrenslegitimation, aber auch um Probleme der Beschränkung der Inhaltsautonomie von Lehrpersonen, die er mit der Implementation von Bildungsstandards massiv beschnitten sieht (Huber/Späni/Schmellentin/Criblez 2006, S. 54). Die Delegation der Inhaltsdimension an Experten drohe die Auseinandersetzung der Lehrerinnen und Lehrer mit „ihren“ Fächern zu ersetzen (ebd.). Weiterhin macht er darauf aufmerksam, dass auch offen bleibe, welches Gewicht den zentralen Tests relativ zum Lehrerurteil zukommen solle. Einen Widerspruch sieht er des Weiteren in der Behauptung einer erhöhten Schulautonomie angesichts zunehmender Leistungskontrolle. Unterstützung findet die Kritik Neuwegs etwa von Altrichter/Posch (2004). In der Absicht der Bildungspolitik, die Einzelschule zunehmend zur zentralen Einheit der Qualitätssicherung zu erheben, sehen die Autoren die Tendenz hin zu einer Ablehnung der Verantwortung für Bildungsprozesse von Seiten des Staates: Die Leistungsziele werden zentral verlangt und geprüft, das *wie* der Leistungsproduktion wird aber der Autonomie der Schule anheim gestellt.

6.2.3 Überprüfung der Leistungen: Tests und Bildungsmonitoring

In Österreich werden voraussichtlich im Schuljahr 2007/08 erste Standardüberprüfungen in der 4. (Deutsch, Mathematik) und 8. (Deutsch, Mathematik, Englisch) Schulstufe stattfinden (Kubinger/Frebort/Holcoher-Ertl et al. 2006). Bis dahin ist die Pilotphase II abgeschlossen und sollen verbindliche Versionen der Bildungsstandards festgelegt und die Testinstrumente normiert sein. Pro-

betests fanden bereits 2005 und 2006 in der 4. und 8. Klasse statt.²²¹ Die betroffenen Schülerinnen und Schüler können ihre individuellen Leistungen aus diesem Testdurchlauf als differenziertes Stärken-/Schwächenprofil mit ihrem Zutrittscode auf einer Internetplattform²²² abrufen. Daneben wird auch eine Benchmark für den nationalen Vergleich ausgewiesen. In Vorbereitung ist eine Datenaufbereitung auch für Lehrpersonen, Schulleitungen und Schulverwaltung. Zusätzlich erfolgt eine persönliche moderierte Rückmeldung an die Schulleitungen und Klassenlehrpersonen durch eine eigens geschulte Person.

Die künftige Standardüberprüfung soll jährlich anhand einer Stichprobe von insgesamt 30% der Schulklassen vorgenommen werden, wobei sich diese 30% auf die drei getesteten Fächer verteilen, so dass pro Fach 10% getestet werden. Die Auswertung der Tests führen speziell geschulte Lehrerinnen und Lehrer an den Pädagogischen Instituten durch, die Testadministration und Datenverarbeitung erfolgt zentral am Pädagogischen Institut in Linz. Im österreichischen Implementationskonzept ist die Einführung der Lehrkräfte in eine „Kultur“ externen Testens von zentraler Bedeutung, umso mehr, als auch hier eine entsprechende Tradition externer Leistungsüberprüfung fehlt.²²³ Da die Form der Rückmeldung von Ergebnissen als besonders sensibel einzustufen ist, wird auch für das persönliche Rückmeldeverfahren an die Schulen ein Pool von Lehrkräften speziell qualifiziert.

Die Daten werden den Schülerinnen und Schülern individuell, den Lehrerinnen und Lehrern für ihre Klassen, den Schulleitungen für ihre Schule und den Schulverwaltungen anonymisiert für ihr Bundesland rückgemeldet. Die Er-

²²¹ Die Entwicklung der Testinstrumente in den Fächern Deutsch und Mathematik findet an der Universität Wien, Arbeitsbereich Psychologische Diagnostik, statt; für das Fach Englisch ging der Auftrag an das Language Testing Centre an der Universität Klagenfurt.

²²² Demoversion unter: <http://www.bildung-standards.at/demo/> (Stand 26. 05. 06).

²²³ Mit MATKOMP existiert in Österreich ein freiwilliges Instrument zur Selbstevaluation des Mathematikunterrichts sowie zur Erfassung des Kompetenzprofils einzelner Schülerinnen und Schüler, das zugleich in den Umgang mit Tests und die Auswertung und Interpretation von Ergebnissen einführt. Das standardisierte Testverfahren misst allgemeine und inhaltliche mathematische Kompetenzen am Ende der Sekundarstufe I. Die Aufgaben stammen aus dem TIMSS-Pool und wurden mit Bezug auf den Lehrplan '99 neu klassifiziert. Neben Auswertungshilfen stehen umfangreiche Interpretationsmaterialien zur Verfügung. Für die Anwendung ist eine kurze Einschulung erforderlich

(<http://www.gemeinsamlernen.at/siteVerwaltung/mOBibliothek/Bibliothek/MATKOMP%20I.pdf> (Stand 25. 05. 06).

gebnisse der Standardüberprüfungen sollen ausdrücklich nicht zu Vergleichen zwischen den Einzelschulen herangezogen werden, sondern einzig der Überprüfung des Gesamtsystems bzw. den Einzelschulen der eigenen Qualitätsentwicklung dienen. Ein nationales oder regionales Ranking der Schulen wie in England wird damit ausgeschlossen. Auf der Grundlage des Systemmonitorings wäre ein regelmäßiger nationaler Bildungsbericht zu erstellen; Vorarbeiten für ein Konzept sind am neu errichteten *Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des Bildungswesens* (Bifie) angesiedelt.

Hatte die „Zukunftskommission“ (Haider/Eder/Specht/Spiel 2003) noch gefordert, dass Standards- bzw. Kompetenzüberprüfungen zusammen mit dem Urteil der unterrichtenden Lehrperson zur Noten- und Berechtigungsvergabe herangezogen werden sollen, so wurde deren Funktion für die individuelle Leistungsbeurteilung inzwischen zurückgenommen. Betont wird nun, dass Bildungsstandards weder die Beurteilung durch die Lehrkräfte noch die Einstufung in Leistungsgruppen ersetzen würden, da sie lediglich Teilbereiche des Lehrplans abdecken. Standards sollen als Hilfsmittel für die Selbstbewertung und Orientierung von Schulen und Lehrkräften eingesetzt werden (Lucyshyn 2006a, S. 4). Damit – so kritische Stimmen – würde aber ihre Bedeutung für den Unterrichtsalltag weitgehend im Unklaren bleiben. Nicht angenommen wurde außerdem die Empfehlung, die Standards jährlich *flächendeckend* zu überprüfen und die Individualtestung auf diese Weise mit der Schul- und Unterrichtsevaluation sowie dem Systemmonitoring zu verbinden. Die „Zukunftskommission“ hielt zudem einen zusätzlichen Ressourcenaufwand zur Durchführung der Reform als unumgänglich, während in der Folge keine neuen Institutionalisierungen und Investitionen, sondern lediglich Umschichtungen vorgesehen wurden (Specht 2006).

6.2.4 Aktueller Stand und Strategien der Umsetzung von Bildungsstandards

Im Gegensatz zur föderalen Schulhoheit in Deutschland oder der Schweiz liegt in Österreich die Verantwortlichkeit für das Bildungswesen bei Schulgesetzen, abgesehen von gewissen subsidiären Rechten, grundsätzlich beim Bund. Die Fort- und Weiterbildungsinstitutionen (Pädagogische Institute) sind sowohl beim Bund wie bei den Ländern angesiedelt. Diese rechtliche und organisatorische Ausgangssituation wirkt sich auf die Strategien der Implementation von Bildungsstandards aus. Ausgehend von der zentralen Ebene des Ministeriums wurden in Österreich während der Pilotphase auf den Ebenen Bundesländer und Schulen Leitungs- und Koordinationsstellen eingerichtet. Über diese wer-

den die in den Arbeitsgruppen auf ministerieller Ebene entwickelten Standards an die Basis „diffundiert“, also in einen Top-Down-Prozess eingegeben.

Gleichzeitig existieren Feedback-Schlaufen. Erfahrungen der Pilotschulen mit den Standards und Aufgabenbeispielen werden in die Arbeitsgruppen rückgemeldet und dienen als Basis für die Überarbeitung. Das Implementationskonzept geht davon aus, dass primär die Lehrerinnen und Lehrer Träger des Reformprozesses sind und aktiv beteiligt werden müssen. Partizipationsmöglichkeiten ergeben sich abgesehen von den Feedbacks der Pilotschullehrkräfte über die Beteiligung von Praktikerinnen und Praktikern am Prozess der Standards- und Aufgabenentwicklung. Um das notwendige „commitment“ längerfristig zu gewährleisten, müssen nun auch die für die Implementationsprozesse notwendigen Unterstützungsstrukturen in den Bundesländern auf- und weiter ausgebaut werden.

Ebene Ministerium. Auf der ministeriellen Ebene besteht eine Projektleitung, der die leitenden Ministerialbeamten sowie externe Berater angehören. Die Projektleitung wird von einer Steuergruppe beraten, die sich aus Vertreterinnen und Vertretern des Ministeriums, aus Administration, Schulaufsicht und -praxis sowie aus der Wissenschaft und dem Zentrum für Schulentwicklung zusammensetzt. Ebenfalls auf Bundesebene sind fächer- und schulstufenspezifische Arbeitsgruppen zur Entwicklung der Kompetenzmodelle sowie zur Formulierung von Standards und Aufgabenbeispielen angesiedelt. Die Leiter der Arbeitsgruppen gehören ebenfalls der Projektleitung an. Die Mitglieder der Arbeitsgruppen rekrutieren sich aus Fachdidaktikern und -wissenschaftlern sowie Praktikern aus möglichst allen Bundesländern.

Ebene Bundesländer. In den einzelnen Bundesländern wurden als strategisch wichtige Schlüsselstellen Landeskoordinatoren eingesetzt, die die Verbindung mit der Projektleitung im Bundesministerium herstellen. Der Landeskoordinator oder die Landeskoordinatorin betreut die Pilotschulen und baut in Zusammenarbeit mit der Schulaufsicht und den zuständigen Abteilungsleitungen der Pädagogischen Institute Fach- und Expertenteams an den Pädagogischen Instituten und Schulen auf. Diese sind für die Implementierung der Bildungsstandards an den Schulen und die Unterstützung der Schulen zuständig. Ebenfalls in Zusammenarbeit mit den Pädagogischen Instituten und der Schulaufsicht konzipieren die Landeskoordinatoren Fortbildungsveranstaltungen zur Qualitätsentwicklung an den Schulen. Die Pädagogischen Institute sind dazu verpflichtet, 40% ihres Budgets für die Umsetzung der Standards einzusetzen.

Ebene Schulen. Als entscheidend für das Gelingen von Transformationsprozessen wie die Implementation von Bildungsstandards wird die Schulebene angesehen. Neben den Lehrpersonen kommt den Schulleitungen eine dominante

Rolle zu. Schulleiterinnen und Schulleitern wird seit 2004 vom Bildungsministerium unter dem Titel „Leadership Academy“ ein Qualifizierungsprogramm angeboten (vgl. Schley/Schratz 2005). Schulleitungen sollen damit befähigt werden, am Standort ein Schulprogramm mit den Schulpartnern zu erstellen und die Schul- und Unterrichtsentwicklung professionell zu steuern. Dafür müssen sie im Besonderen strategische Handlungskompetenz sowie Kompetenzen zur Kommunikation und Konfliktlösung erwerben. Das Augenmerk richtet sich speziell auch auf die Personalentwicklung der Mitarbeiterinnen und Mitarbeiter.

Um Informationen darüber zu erhalten, welche Unterstützungs- und Fortbildungsmaßnahmen Lehrpersonen bei der Umsetzung der Standards benötigen, wurden Workshops veranstaltet, in denen diese Fragen von Lehrerinnen und Lehrern unter Einbezug eigener Erfahrungen etwa an Pilotschulen selber bearbeitet werden konnten. Auf der Grundlage der Ergebnisse wurden Module für die Lehrerfortbildung entwickelt und ab Herbst 2006 interessierten Lehrpersonen angeboten.

Auf der Ebene Schule organisiert eine Lehrperson pro Fach als Schulkordinator die Standarderprobung und administriert die Standardtestung, wenn es an der Schule zu einer solchen kommt. Die Lehrkraft ist Ansprechpartner für das Pädagogische Institut und den Landeskoordinator und organisiert Fachkonferenzen und Maßnahmen zur Implementation von Standards an der betreffenden Schule. Den Lehrerinnen und Lehrern stehen für die Planung des Unterrichts und dessen verstärkte Ausrichtung an Ergebnissen (Outcomes) die Expertenteams der Pädagogischen Institute zur Verfügung. Die Lehrerinnen und Lehrer geben den Landeskoordinatoren und den Arbeitsgruppen auf Bundesebene Rückmeldungen über ihre Erfahrungen während der Pilotphase. Der Prozess wird vom Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des Bildungswesens (Bifie) evaluiert.

Um befähigt zu werden, aus Rückmeldedaten einen Qualitätsentwicklungsprozess in Gang zu setzen, brauchen die Schulen professionelle Begleitung. Dazu wurden bis Sommer 2006 150 Multiplikatorinnen und Multiplikatoren speziell in die Thematik der Bildungsstandards eingeführt, um sämtlichen Schulen unterstützend zur Verfügung zu stehen. Auch die Lehrpersonen selbst müssen auf den Einstieg in eine Testkultur vorbereitet werden. Deshalb werden seit Anfang 2006 erstmals Lehrkräfte aus sämtlichen Bundesländern auf den Umgang mit Rückmeldungen hin geschult.

Bis 2007/08, wenn die normierten Bildungsstandards voraussichtlich erstmals überprüft werden, sollen das Verständnis und die Akzeptanz der Bildungsstandards auch über die Pilotschulen hinaus so weit gewachsen sein, dass diese im Unterricht im Sinne der Qualitätssicherung und -entwicklung Wirksamkeit entfalten können. Bis dahin sollen auch die Unterstützungs- und

Rückmeldesysteme weitestgehend aufgebaut werden sowie die Lehrerfortbildung auf standardbasierten Unterricht umgestellt sein. Damit wird die Leitidee einer allmählichen Verlagerung der Steuerung von staatlichen Vorschriften hin zu einer Reform, die zunehmend von den Schulen selbst geleitet wird, verbunden. Um die Transformation von einer reinen Inputsteuerung zu einer Orientierung an Ergebnissen zu initiieren, ist zuerst eine Top-Down-Strategie mit stärkerer Lenkung des Prozesses vorgesehen, um sicherzustellen, dass das Projekt die Unterstützung eines kritischen Maßes an Lehrpersonen und der Schuladministration erhält und nicht in der Pilotphase stecken bleibt. Im Lauf der Jahre sollte es zu einer Rücknahme staatlicher Vorgaben kommen können, so dass der Prozess eine eigene Dynamik entfaltet.

6.3 Schweiz

6.3.1 Diskussionen und Problemwahrnehmungen

Mehrere der für Deutschland und Österreich genannten Ausgangsprobleme und Entwicklungstendenzen, wie Bestrebungen zur Erhöhung der Eigenständigkeit der Einzelschulen, kombiniert mit einem neuen outputorientierten Steuerungsmodell, die Notwendigkeit der Herstellung einer besseren Vergleichbarkeit von Noten, der Wille zur Hebung der Schülerleistungen und zur Verbesserung der Chancengleichheit im Anschluss an die PISA-Ergebnisse, haben auch Gültigkeit für die Schweiz. Gleichzeitig bilden Harmonisierungsbestrebungen zwischen den Bildungssystemen der Kantone in der föderalen Schullandschaft der Schweiz, wie sie bereits Jahrzehnte andauern, den speziellen Rahmen der schweizerischen Diskussion von Bildungsstandards.

Grundsätzlich liegt in der Schweiz die Zuständigkeit für das Schulwesen bei den Kantonen. Es handelt sich also um ein föderales System, das insofern mit Deutschland vergleichbar ist. Gleichzeitig müssen die unterschiedlichen Größenverhältnisse und Aufsichtsstrukturen innerhalb der Kantone sowie das Fehlen eines Bundesministeriums für Bildung, was insgesamt in der (schul-)politischen Tradition begründet ist, berücksichtigt werden. 1973 hatten die Kantone, im Gegensatz zu einer knappen Zustimmung des Soveräns, eine Ausweitung der Kompetenzen des Bundes im Bildungswesen abgelehnt. Inzwischen ist die Dominanz des föderalistischen Prinzips aufgrund verschiedener Entwicklungen mehr und mehr unter Druck geraten, ohne dass die grundsätzliche Zuständigkeit der Kantone für das Schulwesen in Frage steht. Vor dem Hintergrund einer zunehmenden Mobilität haben sich die Erwartungen und

Anforderungen von Seiten der Eltern und der Wirtschaft geändert, so dass kaum mehr auf Verständnis stößt, dass Mathematik auf kleinstem Raum unterschiedlich gelehrt und gelernt wird oder dass der Übergang in die Sekundarstufe I, die kantonal verschieden gegliedert ist, zu unterschiedlichen Zeitpunkten erfolgen soll. Zudem hat die OECD das schweizerische Bildungswesen als vergleichsweise teuer qualifiziert (OECD 1991), was zu Überlegungen hinsichtlich einer Effizienzsteigerung führen muss.

Vor diesem Hintergrund haben die Kantone ihre Kooperation im Bildungsbereich seit den neunziger Jahren beschleunigt und intensiviert. Die interkantonale Zusammenarbeit wird juristisch über Staatsverträge geregelt. Bereits mit dem Schulkonkordat von 1970 haben sich sämtliche Kantone außer dem Tessin zur Zusammenarbeit verpflichtet. Seitdem wurden weitere Vereinbarungen bezüglich zeitlicher Regelungen der obligatorischen Schule, zur Diplomanerkennung sowie Finanzierung und Freizügigkeit vor allem im tertiären Bildungsbereich getroffen. 2001 hat sich die *Schweizerische Konferenz der kantonalen Erziehungsdirektoren* (EDK) Leitlinien gegeben, welche die neuen Ziele der Bildungs Kooperation Schweiz und die Art der Zusammenarbeit umschreiben. Seither arbeitet die EDK nach einem Tätigkeitsprogramm, welches in Schwerpunkte gegliedert ist. Höchste Priorität haben

- die Harmonisierung der obligatorischen Schule durch die Festlegung landesweiter Leistungsstandards,
- die Weiterentwicklung des Sprachenunterrichts,
- die Stärkung der Professionalität der Lehrkräfte und
- der Aufbau eines schweizerischen Bildungsmonitorings in Zusammenarbeit mit dem Bund.

Am 21. Mai 2006 haben nun Volk und Stände der Revision der Bildungsartikel in der Bundesverfassung mit überwältigender Mehrheit zugestimmt. Die damit veranlassten Änderungen führen zu einer deutlichen Stärkung des Konkordatsrechts und der interkantonalen Koordination. Die Kantone sind in Zukunft zu einer umfangreicheren und wirksameren Zusammenarbeit verpflichtet, wobei der Bund interkantonale Verträge in Bereichen der Harmonisierung des Schulwesens allgemein verbindlich erklären oder Kantone zur Beteiligung an interkantonalen Verträgen verpflichten kann (Art. 48a, Art. 62).

Die Einführung von Bildungsstandards ist ein zentrales Element des aktuellen Konkordats zur Harmonisierung der obligatorischen Schule (HarmoS), welches im Juni 2002 von der EDK auf den Weg gebracht worden ist. Des Weiteren regelt die neue interkantonale Vereinbarung wichtige strukturelle Eckwerte der obligatorischen Schule, wie das Einschulungsalter (frühere und flexiblere Einschulung), die Dauer der Schulstufen und der obligatorischen Schulzeit insge-

samt. Das Vernehmlassungsverfahren in den Kantonen, einbezogen sind auch die nationalen Verbände der Lehrerschaft sowie die Elternorganisationen, konnte 2006 abgeschlossen werden. Nach der Bereinigung des Entwurfs soll das Konkordat spätestens im Oktober 2007 zu Händen der Kantone und im Rahmen der dort stattfindenden Beitrittsverfahren von der EDK-Plenarversammlung verabschiedet werden. Sobald zehn Kantone das Konkordat ratifiziert haben, voraussichtlich Anfang 2009, tritt es in Kraft. Die Kantone werden danach im Prinzip vier Jahre Zeit haben, um ihre Strukturen und Schulprogramme an die verbindlichen Vertragsbedingungen anzupassen.

Die Entwicklung und Einführung von Bildungsstandards bildet somit eine Komponente in diesen umfassenderen und längerfristigen Bestrebungen; sie ist im Kern bezogen auf curriculare Anpassungen. Harmonisierungsversuche bezogen auf die Ziele und Inhalte des Unterrichts sind übrigens bereits in den neunziger Jahren mit den kantonsübergreifenden „Treffpunkten“, welche curriculare Kerninhalte definierten, unternommen worden. Andererseits darf nicht übersehen werden, dass, abgesehen von diesen weiter zurückreichenden Bestrebungen, auch in der Schweiz die zum Teil enttäuschenden Ergebnisse aus PISA 2000 wenn nicht einen „Schock“, so doch heftige Diskussionen hervorgerufen hatten.

Im Gegensatz zu den offiziellen Dokumenten der EDK zeigen die Verlautbarungen aus den Kantonen, dass die Forderung nach nationalen Standards und deren Überprüfung durch Tests in einem direkten Zusammenhang mit PISA steht. Wegleitend dabei ist die Feststellung, dass die erfolgreichen Länder eine entsprechende Politik der Output-Steuerung eingeführt hätten (Strittmatter 2004). Als weitere „Triebkräfte“ zugunsten einer Einführung von Bildungsstandards sind Forderungen nach einer Klärung des Bildungsauftrages vor dem Hintergrund „inflationärer Lehrpläne“ und als Antwort auf Kritiken an den Zensur- und Zeugnispraktiken zu nennen. Mit Bezug auf letztere Problematik und im Sinne eines Misstrauens gegenüber den Zeugnisnoten haben Wirtschaftsverbände eigene Beurteilungsverfahren entwickelt („Basic-check“, „Multicheck“), die bei der Einstellung von Lehrlingen zum Zuge kommen.

6.3.2 Konzeption und Funktion von Bildungsstandards

In der Schweiz findet die Entwicklung und Einführung von Bildungsstandards, wie im vorangehenden Abschnitt erwähnt, im Rahmen des Konkordats zur *Harmonisierung der obligatorischen Schule* (HarmoS) statt. Das Konkordat umfasst folgende vier Zielsetzungen:

- Harmonisierung der Lerninhalte;
- Stärkung einer koordinierten Steuerung der obligatorischen Schule;
- Evaluation des Bildungssystems auf gesamtschweizerischer Ebene;
- Qualitätsentwicklung des Bildungssystems.

Mit dem Instrument der Bildungsstandards werden in Zukunft erwartete Leistungsergebnisse auf nationaler Ebene verbindlich vorgeschrieben. Bisher wurde nur in seltenen Fällen präzise definiert, welche schulischen Leistungen zu einem bestimmten Zeitpunkt erreicht werden sollen. Dies führt zu subjektiven und heterogenen Bewertungen von Schülerleistungen. Bildungsstandards sollen den Schulen und Lehrpersonen objektive Kriterien für die Beurteilung der Schülerinnen und Schüler sowie diagnostische Hilfe bei unterschiedlichen Leistungsständen bieten, aber auch als wirksames Mittel zur Hebung der Schulqualität und zur Steuerung der Schulentwicklung fungieren. Sie ermöglichen sowohl für die Schülerinnen und Schüler als auch für Eltern und Lehrkräfte eine wünschenswerte Transparenz hinsichtlich der erwarteten Fachleistungen.

Bei den schweizerischen Bildungsstandards handelt es sich um *Mindeststandards* und gemäß der Unterscheidung von Ravitch (1995) um *performance standards*, d.h. ergebnisorientierte Leistungsstandards, die mess- und überprüfbar sind (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) 2004). Sie sollen und können auch in der Schweiz Lehrpläne nicht ersetzen, diese aber wohl ergänzen und präzisierend zu einer Harmonisierung der Lerninhalte führen. Mindeststandards legen eine Stufe fest, unter die kein Lernender und keine Lernende zurückfallen darf. Dies schließt nicht aus, dass auch höhere Anforderungen ausgewiesen werden – als Ziele weiterführender Lernprozesse und zur Differenzierung von Leistungen. Die Konzentration auf Mindeststandards ist für die Qualitätssicherung im Bildungswesen aber von entscheidender Bedeutung. Sie zielt darauf ab, dass gerade die Leistungsschwachen nicht zurückgelassen werden. Der Bearbeitung des Befundes, dass die Schweiz im Vergleich mit anderen Industriestaaten Schwächen vor allem im unteren Leistungsbereich zeigt, kommt damit Priorität zu.

Die Standards werden in den Fächern Erstsprache, Fremdsprachen (Englisch und zweite Landessprache), Mathematik und Naturwissenschaften und zwar für das Ende des 2., 6. und 9. Schuljahres (= Ende Basisstufe, Ende Primarschule für die meisten Kantone, Ende obligatorische Schule) festgelegt. Bezogen auf die Sekundarstufe wird auf interkantonaler Ebene keine Differenzierung der Mindeststandards nach Schultypen vorgenommen, das bleibt den einzelnen Kantonen überlassen. Die nationalen Standards können in dieser Hinsicht also differenziert umgesetzt werden.

Im Gegensatz zu Deutschland werden in der Schweiz, der Expertise von Klieime/Avenarius/Blum et al. (2003) folgend, im ersten Schritt Modelle entwickelt, die Aspekte, Abstufungen und Entwicklungsverläufe von Kompetenzen sichtbar machen. Diese werden operationalisiert in Form von konkreten Aufgaben. Gleichzeitig wird die Progressionsspanne innerhalb eines Faches abgebildet, d.h. die kumulative Anordnung von Wissensinhalten und Kompetenzen mit den daran ablesbaren Schwierigkeitsstufungen. Auf der Grundlage der Aufgaben werden Testitems entwickelt und validiert. Die Entwicklung der Kompetenzmodelle und Bildungsstandards erfolgt im Anschluss an ein Ausschreibungsverfahren an verschiedenen Pädagogischen Hochschulen und der Universität Fribourg. Die Projektleitungen arbeiten jeweils mit Institutionen aus den anderen Sprachregionen zusammen. Die Arbeiten wurden im Laufe des Jahres 2005 aufgenommen und werden je nach Fach zwischen Ende 2007 und Herbst 2008 abgeschlossen sein.

Aus empirischer Sicht gibt es keine Gründe für die Wahl von Mindest-, Regel- oder Maximalstandards; schon gar nicht, wenn Standards auf der Grundlage von Kompetenzmodellen definiert sind. Dass Mindeststandards tatsächlich ausgehend von empirisch validierten Kompetenzmodellen mit präzisierten Niveaustufen festgelegt werden, erscheint aber besonders bedeutsam, weil die Verbindlichkeit von Mindestanforderungen *an alle* besonders hoch ist, eine Unterschreitung eine besondere Signalwirkung hat und speziellen Maßnahmen behoben werden sollten. Mindeststandards dienen also vor allem der gezielten Förderung.

Entsprechend werden die Schweizer Mindeststandards – im Gegensatz zu den deutschen Regelstandards – erst im Anschluss an die Entwicklung von Kompetenzmodellen und nach deren Überprüfung mittels Test an einer breiten repräsentativen Stichprobe festgelegt. Die Ergebnisse des Tests geben zudem Aufschluss über die statistische Verteilung der Schülerinnen und Schüler auf den Kompetenzniveaus. Aufgrund der Analyse der Ergebnisse werden Vorschläge für die Festlegung der Mindestkompetenzen unterbreitet und auf politischem Wege verabschiedet. Die Standards werden, wie sie im Herbst 2008 bzw. 2009 beschlossen werden, dem Konkordat im Sinne eines Anhangs angefügt.

6.3.3 Überprüfung der Leistungen: Tests und Bildungsmonitoring

In der Schweiz soll die landesweite Überprüfung der Bildungsstandards am Ende des 2., 6. und 9. Schuljahrs anhand einer nationalen repräsentativen Stichprobe künftig im Rahmen eines landesweit verbindlichen Bildungsmonitorings stattfinden. Verantwortlich für die Durchführung ist eine zu bestimmende nati-

onale Evaluationsstelle. Es spricht einiges für eine Ausweitung des Samples, so dass auch Aussagen bezüglich der einzelnen Kantone möglich sind. Das Bildungsmonitoring, welches voraussichtlich in Zyklen von vier Jahren durchgeführt wird, und der daraus hervorgehende Bildungsbericht sind von Bund und Kantonen gemeinsam veranlasst. Im Sinne der *evidence based policy* wird es ausgehend von den erhobenen Daten „sicher möglich sein, [...] begründete Entscheidungen für die Verbesserung des Bildungssystems zu treffen. Namentlich für die Strukturen, Stufen, Fächer, für bestimmte Schülergruppen, die Lehrerinnen- und Lehrerausbildung, Evaluationsmodalitäten, und gerade auch für die Festlegung der Mindestkompetenzen, die nach unten oder nach oben angepasst werden können“ (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) 2004, S. 14).

Ein erster Pilotbericht liegt seit Ende 2006 vor; aktuelle Schwerpunktthemen sind der Umgang mit Heterogenität und die Lehrerinnen- und Lehrerbildung. Generell werden im Bildungsbericht drei Dimensionen des Bildungssystems untersucht:

- die Effektivität (Wirkung),
- die Effizienz (Wirkung in Relation zum Aufwand, Verhältnis von Input und Output) und
- die Equity (Gerechtigkeit, Chancengleichheit).

Beurteilt werden diese Dimensionen anhand der politischen Vorgaben (Zielsetzungen), aufgrund zeitlicher Vergleiche (Längsschnitte, die mehrheitlich mit dem zyklischen Monitoring erst aufgebaut werden) und in kantonalen Vergleichen (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) 2006, S. 28). Aufbereitungen und Analysen der Daten dienen als Entscheidungsgrundlagen für die bildungspolitischen Behörden bei Bund und Kantonen. Was die Überprüfung der Bildungsstandards anbelangt, so werden auf der Ebene EDK keine Aussagen darüber gemacht, welche Konsequenzen sich aus dem Nichterreichen der Mindeststandards ergeben. Es ist Sache der Kantone, in diesem Fall und für die betroffenen Schülerinnen und Schüler die notwendigen Unterstützungs- und Differenzierungsmaßnahmen einzuleiten (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) 2006).

Gemäß Art. 7, Absatz 5 des HarmoS-Konkordats werden die EDK und die EDK-Regionalkonferenzen sich fallweise über die Entwicklung von Referenztests auf der Basis der festgelegten Bildungsstandards verständigen (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) 2006). Es spricht einiges dafür, die Entwicklung der Referenztests sprachregional anzugehen, insbesondere aufgrund der Nähe zur Lehrplanentwicklung, aber auch aus

Gründen von Kostenerwägungen. Es wird erwartet, dass auf diese Weise der Aufwand für die künftigen Leistungsmessungen wesentlich unter der Beteiligung an PISA mit 4,5 Mio. Franken²²⁴ pro Durchgang bleiben wird. Die Kantone bestimmen im Übrigen über den Einsatz von Referenztests im Rahmen eines kantonalen Monitorings, wobei der den Standards zugrundeliegende Referenzrahmen neben dem Systemmonitoring auch für die Entwicklung bzw. Anpassung von Tests auf Ebene Schulklasse und Schüler zur Anwendung kommen kann. Mit Klassenscockpit²²⁵, Stellwerk²²⁶ oder Check Five²²⁷ sind Testinstrumente zur Standortbestimmung der Klasse und Schule im Vergleich oder zur Analyse des Leistungsstandes individueller Schülerinnen und Schüler bereits vorhanden, die genutzt und mit Bezug auf die Standards weiterentwickelt werden können.

6.3.4 Aktueller Stand und Strategien der Umsetzung von Bildungsstandards

Die Arbeiten an der Entwicklung der Kompetenzmodelle sind in der Schweiz noch nicht abgeschlossen und es liegen zum momentanen Zeitpunkt noch keine konkreten Bildungsstandards vor. Im April/Mai und teilweise im Herbst 2007 werden die bisher entwickelten Kompetenzniveaus anhand einer großen Schülerstichprobe validiert. Ausgehend davon werden die wissenschaftlichen Konsortien der EDK Ende 2007 Bildungsstandards vorschlagen. Diese gehen Anfang 2008 in die Vernehmlassung und werden frühestens im Oktober 2008 durch die EDK-Plenarversammlung verabschiedet. Für die Naturwissenschaften verschiebt sich der Zeitplan um rund ein Jahr, da hier die Ende 2007 vorliegenden PISA-2006-Ergebnisse miteinbezogen werden sollen. Gemäß EDK soll das Vorgehen bezüglich Implementation ab Herbst 2007 in enger Zusammenarbeit mit den Regionen geplant werden; „[d]er Zeitplan sowie das konkrete Vorgehen sind noch zu bestimmen“ (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) 2004, S. 14). Die Standards werden gemäß EDK in erster Linie einen Einfluss auf die Lehrplanarbeit und die Lehrmittel haben. Über die Verwendung der offiziellen Lehrmittel, welche sich ihrerseits auf Standards-kompatible (Rahmen-)Lehrpläne stützen, werden die Lehrpersonen

²²⁴ Dies entspricht ca. 2,8 Mio. €.

²²⁵ <http://www.klassenscockpit.ch/> (Stand 26.05.06).

²²⁶ <http://www.stellwerk-check.ch/> (Stand 26.05.06).

²²⁷ <http://www.ag.ch/leistungstest/de/pub/check5.php> (Stand 26.05.06).

bei ihrer Arbeit unausweichlich die Standards mitberücksichtigen (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) 2006, S. 25).

Die Implementation der Bildungsstandards über neue Lehrpläne und Lehrmittel sowie Test- und Unterstützungssysteme wird in den Kantonen (und Regionen) vor sich gehen, bei denen die Zuständigkeiten liegen. Der neue Rahmenlehrplan für die Westschweizer Kantone (Plan cadre romand PECARO) befindet sich bereits jetzt, d.h. vor Vorliegen künftiger Bildungsstandards, in der Umsetzungsphase; Arbeiten an einem Lehrplan für die Deutschschweiz wurden im Dezember 2006 aufgenommen, mit dem Ziel, einen direkt einsetzbaren Lehrplan 2011 einführen zu können. Insbesondere auch die Notwendigkeit der Schaffung von Instrumenten der Evaluation und Selbstevaluation zur Überprüfung der Bildungsstandards macht eine verstärkte Zusammenarbeit zwischen Kantonen und Regionen sinnvoll. Das HarmoS-Konkordat sieht vor, dass die Entwicklung von Referenztests fallweise durch die EDK oder die EDK-Regionalkonferenzen erfolgen soll. Eine Initiative zur Einrichtung eines Kompetenzzentrums für Leistungsmessung, welches solche organisieren und durchführen würde, wurde von Deutschschweizer Regionalkonferenzen der EDK bereits gefasst. Der Entscheid, welche Instrumente der Evaluation und Selbstevaluation in welchem Ausmaß auf regionaler und/oder kantonaler Ebene zum Einsatz kommen sollen, steht aber momentan noch aus. Es ist an den Kantonen, für diejenigen Schülerinnen und Schüler, die die Basiskompetenzen nicht erreichen, Unterstützungs- und Differenzierungsmaßnahmen zu entwickeln (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) 2004; 2006).

Da in der Schweiz noch keine konkreten Standards-Dokumente vorliegen, mit denen Praktikerinnen und Praktiker in Berührung hätten kommen können, ist das Wissen um Sinn und Funktion bundesweit verbindlicher Mindeststandards bei den Lehrpersonen noch sehr begrenzt. Die kommenden Tests sind mit gewissen Ängsten verbunden, auch wird gelegentlich eine Einschränkung von Unterrichtsfreiheiten befürchtet; allgemein dominiert aber eher die Haltung des Abwartens. Dort, wo Erfahrungen mit Tests vorliegen, sind die Befürchtungen weitaus geringer, wie eine Studie zum Check Five zeigt (Moser/Tresch 2005).

Angesichts des gegebenen Kenntnisstandes wäre es sicher sinnvoll, die Funktion von *Mindeststandards* in Bezug auf Unterricht und ihr Verhältnis zu den vorhandenen Lehrplänen und dem neuen Lehrplan bzw. Rahmenlehrplan breit zu kommunizieren. Folgt man Äußerungen des Dachverbands Schweizer Lehrerinnen und Lehrer (LCH), so werden von dieser Seite Bildungsstandards durchaus mit der Chance verknüpft, Klarheit und Orientierung in die als überladen, diffus und unverbindlich wahrgenommene Lehrplansituation zu bringen.

Bei vielen Lehrpersonen trägt die Verfügbarkeit und in einigen Kantonen die breite Nutzung von freiwilligen Testsystemen wie etwa „Klass Cockpit“ noch zur Verunsicherung bei, auf welche Kernziele hin jetzt und in Zukunft unterrichtet werden soll. Auch ergaben sich zum Beispiel im Kanton Zürich unter den Lehrpersonen Unsicherheiten und Missverständnisse bezüglich der Relevanz solcher Testergebnisse für die Notengebung und Übertrittsentscheide.

Neben der „Lehrplankrise,“ welche sich auch als Wirksamkeitskrise äußert, ist es auch die „Notenkrise“²²⁸, die Bildungsstandards für Vertreter des LCH als begrüßenswert erscheinen lässt. Das Vertrauen der Schulabnehmer, vor allem der Lehrlingsausbilder, in die mitgebrachten Zeugnisnoten ist so weit gesunken, dass zur Einschätzung der Bewerberinnen und Bewerber oftmals Tests privater Anbieter eingesetzt werden.²²⁹ Zur Behebung der Krise verweist der LCH auf die Etablierung von Standards, die den wissenschaftlichen Kriterien der Expertise von Klieme/Avenarius/Blum et al. (2003) genügen. Gefordert wird zudem, dass im Lehrplan die heranzubildenden, d.h. kontinuierlich aufzubauenden, Kernkompetenzen und die dabei von den Lernenden zu erreichenden Mindeststandards klar und verbindlich festgelegt werden (Dachverband Schweizer Lehrerinnen und Lehrer 2005).

„Die Lehrerschaft würde zwar durch verbindliche, fokussierte Lehrpläne Freiheiten einbüßen; das Leiden an den Folgen der Beliebigkeit ist inzwischen aber so groß geworden, dass ein Paradigmawechsel zu verbindlichen Leistungsaufträgen mit einforderbaren Ressourcenallokationen nun fällig geworden ist“ (Strittmatter 2004, S. 40).

Vorausgesetzt, dass die Standards nicht zu Selektions-, Zertifizierungs- oder Rankingzwecken verwendet werden und dass sie mit den notwendigen Unterstützungssystemen ergänzt werden, kann die Haltung des Deutschschweizer Dachverbandes durchaus als offen-positiv bezeichnet werden; dies im Gegensatz zu den insgesamt skeptischeren und teilweise die Bildungsstandards grundsätzlich ablehnenden Stimmen aus der Romandie. Die Vernehmlassung des Konkordatsentwurfs im November 2006 machte deutlich, dass der LCH Bildungsstandards im Sinne einer erhöhten Auftragsklarheit durchaus wünscht, unter Vorbehalt, dass u.a. die qualifikationsmäßige und zeitlich-personelle Aus-

²²⁸ Der ursprüngliche Grund für die Einführung von Vergleichstests im Kanton St. Gallen (heute „Klass Cockpit“) im Jahr 1995 war die Unzufriedenheit des Erziehungsrats mit der Aussagekraft der Noten.

²²⁹ Vgl. „Basic-check“: www.basic-check.ch; „Multicheck“: www.multicheck.ch (Stand 04. 06. 06).

rüstung der Schulen für den Umgang mit Standards und diagnostischen Instrumenten gewährleistet wird. Zudem müssen Lehrpläne, Bildungsstandards, Lehrmittel und Tests in Zukunft kohärent aufeinander abgestimmt sein, sich keinesfalls aber zu sich gegenseitig konkurrenzierenden Parallellehrplänen entwickeln.

6.4 Luxemburg

6.4.1 Diskussionen und Problemwahrnehmungen

Luxemburg hat im Jahr 2000 mit PISA erstmals an einer internationalen Schülerleistungsstudie teilgenommen und erreichte dabei in sämtlichen drei Domänen lediglich den zweitletzten Platz. Drei Jahre später gewann das Land zwar überall Ränge, die Ergebnisse blieben jedoch enttäuschend. Als Grund für das schlechte Abschneiden wird auf die besondere Sprachsituation verwiesen; Luxemburg bildet mit den drei Sprachen Luxemburgisch, Französisch und Deutsch eine multilinguale Einheit, dies im Gegensatz etwa zu den drei getrennten Sprachgebieten der Schweiz. Kennzeichnend ist schließlich eine hohe Zahl an Einwanderern.

Dominierte in den vergangenen Jahren das Thema einer Erhöhung der schulischen Eigenverantwortung, welche besonders mittels Ausarbeitung und Durchführung von Schulprojekten gefördert wird,²³⁰ hat PISA die bildungspolitische Aufmerksamkeit auf die Leistungen des Bildungssystems gelenkt; Luxemburg steht bezüglich Bildungsausgaben pro Schülerin bzw. Schüler im Primar- und Sekundarbereich an der Spitze der OECD-Staaten (OECD 2005). Der soziokulturelle Hintergrund der Schülerinnen und Schüler hat sich in PISA als starke Einflussgröße im Hinblick auf den Schulerfolg erwiesen. Deutlich schlechter fielen die Leistungen im technischen Bildungsgang (Enseignement Secondaire Technique) gegenüber dem klassischen (Enseignement Secondaire) aus. Ein großes Problem stellt zudem die Tatsache dar, dass ein Fünftel der Schülerinnen und Schüler die Schule ohne Abschluss verlässt; außerdem weist Luxemburg eine hohe Zahl an Klassenwiederholungen auf.

²³⁰ Diese Schulprojekte sind aufgrund des Gesetzes zur Reform des technischen Unterrichts entstanden und stellen ein neues Element der Dezentralisierung im Schulsystem dar. Die Projekte werden in der Regel von den einzelnen Schulen ausgearbeitet und sollen den Lehrpersonen und Schülerinnen und Schüler der Sekundarbereiche die Möglichkeit bieten, außerhalb von vorgeschriebenen Programmen neue Forschungsvorhaben und Realisierungsstrategien zu erproben (Pull 2002).

Das Ungenügen bezüglich messbarer Outcomes ist begleitet von Ergebnissen, die auf eine vergleichsweise negative Einstellung zur Schule und mangelnde Motivation der Schülerinnen und Schüler verweisen. Im Vergleich zum OECD-Durchschnitt schätzen die Schülerinnen und Schüler Mathematik als weniger wichtig ein für ihren späteren Beruf (Ministère de l' Education Nationale, de la Formation Professionnelle et des Sports (MENFP) 2004), was sich auf die Lern- und Leistungsmotivation auswirkt und die Frage aufwirft, inwiefern dies an den in diesen Fächern vermittelten Inhalten und der Unterrichtsmethodik liegt. Damit in Zusammenhang dürfte die Beobachtung stehen, dass die Schülerzahlen in mathematisch-naturwissenschaftlich ausgerichteten Klassen der Sekundarstufe, obwohl in diesem Feld gute Berufsaussichten bestehen, rückläufig sind.

Insgesamt diagnostiziert man in Luxemburg eine mangelnde Vorbereitung auf die sich schnell wandelnden Kompetenzanforderungen der „Wissengesellschaft“. Diese scheinen im Rahmen der althergebrachten, extrem gestrafften und wenig flexibeln Organisationsformen von Schule und Unterricht besonders auf der Sekundarstufe nicht mehr erreichbar. Die Forderung nach Bildungsstandards bzw. Kernlehrplänen steht auch im Zusammenhang damit, dass die Lehrpläne in der Vergangenheit zunehmend an Umfang gewonnen haben, besonders in der technischen Sekundarschule als überladen wahrgenommen werden und eine Entschlackung gewünscht wird.

Vor diesem Hintergrund wendet sich der Blick auf erfolgreiche PISA-Länder, die mittels neuer Modelle der Steuerung auf die anstehenden Herausforderungen besser zu reagieren im Stande sind. Zu den erfolgversprechenden Elementen gehören Bildungsstandards und eine ergebnisorientierte Steuerung, kombiniert mit einer teilautonomen Organisation der Schulen; Letztere, inkl. Erweiterung der Gestaltungsfreiräume der Lehrkräfte im Unterricht, wurde mit dem Gesetz vom 25. Juni 2004 für die Lyzeen (Lycées und Lycées Techniques) bereits eingeführt.

6.4.2 Konzeption und Funktion von Bildungsstandards

Die geringe Größe des Großherzogtums Luxemburg (2,6 km², 430 000 Einwohner) selbst im Vergleich zur Schweiz bringt es mit sich, dass hier die eigenen Ressourcen und die Expertise bezüglich der Entwicklung von Bildungsstandards eng begrenzt sind und man auf Vorarbeiten und die fachwissenschaftliche Zusammenarbeit mit den Nachbarländern angewiesen ist. Eine Universität existiert in Luxemburg erst seit 2003. Dort wurde allerdings mit der Forschungsabteilung EMACS (Educational Measurement and Applied Cogni-

tive Science) eine Stelle u.a. für schulische Kompetenzmessung, computergestützte Tests und das nationale Bildungsmonitoring eingerichtet; das dort angesiedelte Projekt „La place de l'école dans la société luxembourgeoise de demain“ verfolgt das Ziel, internationale und nationale Daten und Reporte auszuwerten, aufgrund international vergleichender Studien und wissenschaftlicher Forschung Konzepte für eine Bildungsreform und die Schulentwicklung auszuarbeiten und schließlich die Einstellungen von Schülerinnen und Schülern und Eltern gegenüber der Schule zu untersuchen.

Im Zentrum steht momentan die Entwicklung des Mathematikunterrichts; in diesem Fach existieren inzwischen kompetenzorientierte Bildungsstandards im Sinne vom Mindeststandards (socles de compétences) für die 2. Klassen der Sekundarschulen²³¹; diese formulieren vier Kompetenzen bezogen auf mathematische Prozesse und fünf inhaltsbezogene Kompetenzen. Jeder Kompetenzbereich definiert sich über bestimmte Fähigkeiten, Fertigkeiten und Einstellungen bzw. Kenntnisse. Ausgearbeitet werden die Standards, welche des Weiteren für Sprachen und Naturwissenschaften vorgesehen sind, von verschiedenen Arbeitsgruppen im Ministerium und auf der Ebene der Programmkommissionen (Commissions Nationales des Programmes)²³². Die Standards beziehen sich auf das Ende der Primarschule (6. Klasse) und auf die Sekundarstufe, aber auch Standards für die Vorschule befinden sich in Ausarbeitung. Die Arbeitsgruppen werden begleitet von Wissenschaftlern der Universitäten und Hochschulen Salzburg, Zürich, Kiel und Freiburg.

Das Kompetenzmodell Mathematik/Sekundarstufe und der Kernlehrplan sind im Auftrag des Erziehungsministeriums von der Projektgruppe „BiStaMath“ erstellt worden; die Projektgruppe wird von einem Mitarbeiter des Ministeriums geleitet und setzt sich aus Lehrpersonen der verschiedenen Schulformen zusammen. Bei der Entwicklung des Kompetenzmodells Mathematik stand die Festlegung der mathematischen Prozesskomponenten, die das Wesen des mathematischen Tuns beschreiben, am Anfang. In einem zweiten Schritt wurden Inhaltsbereiche beschrieben, in denen prozessbezogene Kompetenzen erworben und weiterentwickelt werden. Das Zusammentreffen von

²³¹ URL: [http://www.men.public.lu/publications/postprimaire / brochures_information/ kompetenzen/Bildungsstandards_Mathematik.pdf](http://www.men.public.lu/publications/postprimaire/brochures_information/kompetenzen/Bildungsstandards_Mathematik.pdf) (Stand 20.04.07).

²³² Die CNP setzen sich u.a. aus Delegierten der Fachlehrerschaft der Schulen und des Inspektorats zusammen. Der Präsident und der Sekretär werden direkt vom Minister gewählt, die übrigen Mitglieder aufgrund des Vorschlags der jeweiligen Fachschaft, die sie vertreten. Die CNP sind wichtige Einrichtungen für den Informationsfluss zwischen Ministerium und Schulen/Lehrkräften.

Prozesskomponenten und Lerninhalten findet in Lernsituationen statt, die exemplarisch im Anhang der Kernlehrpläne beschrieben sind. Hauptbestandteil des Kernlehrplans sind die erwarteten Schülerkompetenzen am Ende der obligatorischen Schulzeit (Klassen 9/10), die sich nicht nur auf fachliche Inhalte beziehen, sondern auch dazu beitragen, die Jugendlichen für den Übertritt in das Erwachsenen- und Berufsleben vorzubereiten. Sie sollen selbstreguliertes, d.h. verständnisförderndes und motiviertes Lernen begünstigen.

Die Arbeit der Projektgruppe Mathematik umfasst auch das Sammeln von Aufgaben, die bestimmte Anforderungen des Lehrplans darstellen und das Kompetenzmodell erläutern und dabei praktische Erfahrungen berücksichtigen. Die erarbeiteten Dokumente werden immer auch den zuständigen Programmkommission (Commission nationale pour les programmes (CNP)) vorgelegt und deren Rückmeldungen bei der Überarbeitung berücksichtigt. Die Projektgruppe macht zudem Vorschläge für Items für die Lernstandserhebungen und gibt Anregungen für die Unterrichtsentwicklung und Qualitätssicherung.

Ganz allgemein betrachtet man es bei der Entwicklung fachorientierter Bildungsstandards als vorteilhaft, zuerst ein schulformübergreifendes Konzept zum Erwerb übergeordneter Kompetenzen auszuarbeiten. Diese übergeordneten Kompetenzen sollen aus allgemeinen Bildungszielen hergeleitet werden (Bertemes o.J.). Man will dabei gewisse Felder von Fähigkeiten und Einstellungen definieren, die inhaltlich eine sinnvolle Einheit ergeben, aber (teilweise) unabhängig von fachspezifischen Bereichen sind. Diesen transversalitätfördernden Kompetenzfeldern (z.B. Problemlösen) sollen Ressourcen entsprechen, die die Schülerinnen und Schüler in allen Lebens- und Arbeitssituationen mobilisieren können (ebd.). Die Beschreibung fachlicher Kompetenzen schließt sich an die Ausarbeitung übergeordneter Kompetenzfelder an.

Die daraus abgeleiteten Bildungsstandards sollen im luxemburgischen Verständnis das Lehren und Lernen auf verbindliche Ziele ausrichten, wobei der Befähigung zum lebenslangen und selbstgesteuerten Lernen zentrale Bedeutung zukommt. Bildungsstandards betonen die pädagogische Verantwortung für Lehren, Fördern, Fordern und Bewerten und sind ein wichtiges Instrument der Unterrichtsentwicklung. Sie leisten aber auch einen Beitrag zur Schulentwicklung, da das Umsetzen von Bildungsstandards in der Schule die notwendige Reflexion über bestehende Organisationsformen und Lerninhalte nach sich zieht. Sie legen fest, über welche fachlichen, aber auch methodischen und sozialen Kompetenzen die Schülerinnen und Schüler bis zum Ende einer bestimmten Klassenstufe verfügen müssen. Betonung findet neben Motivation und Nachhaltigkeit bei der Aneignung von Fähigkeiten und Fertigkeiten auf der Schülerseite auch die Forderung, dass Standards die Lehrpersonen methodisch nicht

einengen, sondern mittels der Ergebnisorientierung Gestaltungsfreiheiten ermöglichen sollen. Dies soll durch die Schaffung neuer Kernlehrpläne auf der Sekundarstufe gewährleistet werden. Ausgehend von den Ergebnis- und Zielvorgaben von Bildungsstandards/Kernlehrplänen werden schuleigene, jahrgangsbezogene Lehrpläne erstellt, wie dies aktuell für das Fach Mathematik an den Schulen geschehen ist. Die Kernlehrpläne beanspruchen maximal zwei Drittel der Unterrichtszeit, d.h. die Stofffülle der einzelnen Fächer wird reduziert, damit Freiraum für die Differenzierung von Inhalten und Methoden durch die Lehrperson und die Umsetzung schuleigener Zielsetzungen geschaffen werden kann.

6.4.3 Überprüfung der Leistungen: Tests und Bildungsmonitoring

Bildungsstandards bilden auch in Luxemburg die Grundlage für eine Kompetenzüberprüfung mittels Klassenarbeiten, Tests oder zentraler Prüfungen. Für die Primarstufe wurden standardisierte Tests für das Ende der drei Zyklen, d.h. für das Ende der 2., 4. und 6. Klasse erarbeitet; obligatorisch ist die Durchführung in der 6. Klasse als Bestandteil der Orientierungsprozedur. Dieser Plan wird jedoch im Ministerium erneut überdacht, da auf diese Weise keine Möglichkeit der Förderung im Hinblick auf den Übertritt in die Sekundarstufe mehr besteht. In der Sekundarschule wurden die Klassenarbeiten inzwischen zu kompetenzorientierten Lernstandserhebungen weiterentwickelt. Am Ende des Schuljahres 2004/05 haben bereits erste flächendeckende Lernstandserhebungen im Fach Mathematik der technischen Sekundarschule und im Herbst 2006 für die Schülerinnen und Schüler der 3. Klassen sämtlicher Sekundarschultypen stattgefunden. Für die Vorschule wird zurzeit das Salzburger „Beobachtungsmodell“ erprobt.

Die Ergebnisse der Lernstandserhebungen können mit anderen Klassen und Schulen sowie mit einem nationalen Durchschnitt verglichen werden. Im Sinne der Messung kumulativ aufgebauter Kompetenzen beziehen sich die Aufgaben nicht unmittelbar auf den vorausgegangenen Unterricht. Sie orientieren sich stattdessen am mathematischen Wissen und Können, über das Schülerinnen und Schüler gegen Ende der Sekundarstufe I verfügen sollen. Die Items wurden so gewählt, dass aus den Ergebnissen Stärken und Schwächen des fachlichen Lernens in den jeweiligen Klassen erkennbar werden und sich Hinweise für eine entsprechende Weiterentwicklung des Unterrichts, aber auch auf den Förderbedarf einzelner Schülerinnen und Schüler ableiten lassen. Zu letzterem Zweck lassen sich insbesondere die „Diagnoseaufgaben“ verwenden. Die Ergebnisse dienen der Lehrperson damit zur eigenen Standortbestimmung bezüglich der Klasse sowie zur individuellen Lernstandsdiagnose. Die Ergebnisse sol-

len jedoch nicht in die Benotung einbezogen werden, wie auch keine Schulrankings vorgesehen sind. Gleichzeitig mit dem ersten Probelauf fand eine Testung verschiedener Formate der Datendarstellung und -rückmeldung statt.

6.4.4 Aktueller Stand und Strategien der Umsetzung von Bildungsstandards

Die Verantwortung für die allgemeinbildenden Schulen, inkl. Lehrpläne, Stundentafeln, Lehrmittel etc., liegt im Großherzogtum Luxemburg beim Erziehungsministerium und ist durch staatliche Gesetze geregelt. Eine bedeutende Rolle bei der Definition der nationalen Curricula kommt den Commissions Nationales des Programmes (CNP)²³³ zu, die fach- und stufen-/schultypspezifisch organisiert sind und die zuständige Ministerin bzw. den Minister beraten. Die Vorschulen und Primarschulen werden lokal verwaltet. Zwischen der nationalen Ebene und den Gemeinden bestehen keine weiteren politischen und administrativen Einheiten, was sich zusammen mit der geringen Größe des Landes auf strukturelle und organisatorische Fragen der Implementation auswirkt.

Die Vermittlung der Bildungsstandards an die Schulen und Lehrpersonen geschieht in erster Linie über die Zusammenarbeit des Ministeriums mit den CNP, von wo aus Informationen über die Vertreterinnen und Vertreter in den CNP an das Kollegium ihrer Schulen herangetragen werden und Feedbacks von dort zurück in die CNP gelangen. Der Fachvertreterin bzw. dem Fachvertreter kommt somit als Bindeglied zwischen Ministerium und Schule, wo sie/er die jeweilige Fachgruppe leitet, im Implementationsprozess als „Koordinatorin“ bzw. „Koordinator“ eine essentielle Rolle zu. Diese koordinierende Funktion ist auch von Bedeutung bei der Erstellung des schuleigenen Mathematik-Profiles in der Fachgruppe sowie bei der Planung der fachlichen Weiterbildung auf Schulebene. Der Implementation der Bildungsstandards dienen zudem die Aufgabenbeispiele, wie sie für die Mathematik in der Projektgruppe BiStaMath gesammelt werden. Schließlich sollen exemplarische Lernarrangements erstellt werden, die als Anregung zur standardbasierten Unterrichtsentwicklung dienen können.

Zur Schaffung von Akzeptanz und zur Einführung in die Arbeit mit Standards werden die Lehrkräfte in den Prozess der Erstellung der Bildungsstandards eingebunden. Das Kompetenzmodell zur Mathematik (Sekundarstufe) wurde den Lehrerinnen und Lehrern in zwei Phasen vorgestellt: In einer ersten

²³³ Vgl. Fußnote 206.

Zusammenkunft in der CNP wurden den Vertreterinnen und Vertretern der Schulen der Verlauf der Entwicklung der Bildungsstandards sowie die dabei einbezogenen überfachlichen Kompetenzen präsentiert. Die Feedbacks aus den Schulen wurden dann zusammengetragen und der Projektgruppe übermittelt. Zusätzlich haben die Fachlehrkräfte wiederum über die Vertreterinnen und Vertreter in der CNP Stellung zu nehmen zur Frage, welche fachspezifischen Inhalte, Fähigkeiten und Fertigkeiten Bestandteil des Kernlehrplans sein sollen und welche optional bleiben, d.h. den schulspezifischen Lehrplänen überlassen werden können. Die definitiven Bildungsstandards wurden durch eine Sammlung von Aufgaben ergänzt, die das Kompetenzmodell erläutern, bestimmte Anforderungen des Lehrplans darstellen und praktische Anregungen geben. Dasselbe Prozedere ist inzwischen ebenfalls im Hinblick auf den kompetenzbasierten Sprachenunterricht (Französisch, Deutsch) angelaufen; die Definition der Sprachkompetenzen nimmt Bezug auf den Gemeinsamen europäischen Referenzrahmen für Sprachen.²³⁴

Den Bedürfnissen u.a. eines kompetenzorientierten Unterrichts entsprechend arbeitet das Ministerium ebenfalls unter Konsultation der CNP/ Lehrkräfte ein modulares Weiterbildungskonzept aus; ausgehend davon können sich die Koordinatorinnen und Koordinatoren und Lehrkräfte den eigenen Bedürfnissen entsprechend weiterbilden.

Die Ergebnisse aus PISA werden in Luxemburg, wo die Tests flächendeckend eingesetzt wurden, vom *Service de coordination de la recherche et de l'innovation pédagogiques et technologiques* SCRIPT schulspezifisch aufgeschlüsselt und zur kritischen Reflexion vor dem Hintergrund eines nationalen Durchschnitts an die Schulen weitergeleitet. Intensive Debatten fanden nach den Analysen und dem Vergleich mit erfolgreichen Ländern im Anschluss an die Erhebungen von 2000 und 2003 auch im Parlament statt. Die Regierung wurde aufgefordert, ein Schulevaluations- und Begleitverfahren zu entwickeln. Darauf wurde mit der Initiative *Protocole d'Action Qualité Scolaire* (PAQS)²³⁵ ein System externer Schulevaluation für die Sekundarschulen beschlossen. Diese Qualitätssicherungsinitiative steht auch im Zusammenhang mit der erhöhten Eigenverantwortung, die die Sekundarschulen erhalten haben.

Es handelt sich bei PAQS um ein Verfahren der Schulevaluation, das in einen individuellen Schulbericht mündet, von dem ausgehend die Ziele der Schulentwicklung für einen Aktionsplan abgeleitet werden, dessen erfolgreiche

²³⁴ Vgl. Arbeitspapier „Deutschprofile VII“, URL: http://focoweb.script.lu:8080/script/pdf/Spec_comp/Deutschprofile_7_Juli_2006.pdf (Stand 20.04.07).

²³⁵ <http://www.script.lu/activinno/paqs/projet-paqs.pdf> (Stand 11. 03. 06).

Umsetzung dann Gegenstand der nächsten Evaluation ist. Die Evaluations- und die Begleitfunktion wird vom SCRIPT ausgeübt, die Aufsicht obliegt dem Ministerium. Dieser Bericht soll es jeder Schule ermöglichen, einen Schulentwicklungsplan zu erstellen und die Schulqualität zu verbessern. PAQS setzt besonders auf Kommunikation und Zusammenarbeit zwischen den beteiligten Schulpartnern. Schulqualität orientiert sich damit, abgesehen von Schülerleistungen und Effizienzfragen, an der Zufriedenheit sämtlicher Schulpartner. Entsprechend kommt ab 2006/07 für die Erfassung qualitativer Aspekte ein Fragebogen hinzu.

PAQS ist im Herbst 2005 angelaufen, seit Januar/Februar 2006 liegen nun die Berichte zu den ersten acht visitierten Schulen vor. Die Initiative beschränkt sich auf Gymnasium und technisches Gymnasium, während die Primarschulen abgesehen von staatlichen Inspektoraten bis jetzt keine externe Evaluation kennen. Die Schulen werden von einem Evaluationsteam, zusammengesetzt aus zwei Mitgliedern des SCRIPT und zwei Mitgliedern des Erziehungsdepartements, besucht. PAQS beansprucht, summative wie formative Bestandteile zu umfassen: Der Bericht erstellt aufgrund von Daten zu Schulbudget, Schulverwaltung und Leistungsergebnissen eine Diagnose, es wird eine Bilanz gezogen; die Evaluation ist formativ, insofern Prozesse ebenso Berücksichtigung finden wie Ergebnisse, wobei dem Dialog und dem Meinungsaustausch Bedeutung beigemessen wird. Daten aus der Selbstevaluation der Schule werden ebenfalls berücksichtigt. Der Schulbericht enthält in einem ersten Teil eine Beschreibung der Schule und der vorhandenen Ressourcen (opportunities to learn und Input, sozio-ökonomische Zusammensetzung der Schülerschaft etc.). Der zweite Teil gibt die erreichten Resultate wieder (Anzahl Schulabschlüsse, Noten – output). Einen wichtigen Indikator bilden die Schulnoten, was problematisch ist: Sie sind zumeist nicht objektiv, sondern orientieren sich am Klassendurchschnitt und geben lediglich ein statisches Bild vom Leistungsstand eines Schülers. Das Evaluationsteam diskutiert den Bericht mit der Schule, und gemeinsam werden die Ziele für die nächsten drei Jahre, d.h. bis zum nächsten Besuch, festgesetzt. Diese werden in einem Aktionsplan auf Schulebene festgehalten und von Seiten des Ministeriums sowie der Schule in der Folge umgesetzt.

7 Zusammenfassung und Schlussfolgerungen

Die wesentlichen Erkenntnisse der Expertise werden abschließend nochmals zusammenfassend dargelegt (7.1). Die Aussagen orientieren sich am Aufgabenspektrum der Implementation von Bildungsstandards und verdeutlichen die damit verbundenen Chancen und Probleme.

Die Ausführungen stützen sich auf

- die internationale Forschung,
- die Erfahrungen mit Implementationen in vier ausgewählten Ländern,
- die von uns geführten Expertengespräche,
- die bisherigen Entwicklungen und damit verbundenen Diskussionen in den Auftragsländern,
- die Klieme-Expertise und
- die Würdigung des historischen Kontextes und der kulturellen Voraussetzungen der Systementwicklung.

Die Schlussfolgerungen (7.2) richten sich an die Bildungspolitik und beziehen sich auf Aspekte, bei denen die politische Steuerung, strukturelle Weichenstellungen, die Allokation von Ressourcen sowie die Kommunikation der Reform nach außen betroffen oder gefordert sind. Wir geben keine länderspezifischen Empfehlungen, weil dies nicht Teil des Auftrages war.

7.1 Zusammenfassung zentraler Aussagen der Expertise

Ausgangspunkt der Expertise ist der internationale Wandel der Bildungspolitik, der inzwischen häufig als „Paradigmenwechsel“ von der Input- zur Outputsteuerung bezeichnet wird. Bildungsqualität soll danach nicht nur gute Absicht, sondern auch Orientierung am Resultat bedeuten. Der Expertise liegt ein Modell zugrunde, das zu Beginn des zweiten Kapitels vorgestellt wird. In diesem Modell wird eine Outputorientierung unterstützt, nicht jedoch als *alleinige* Option. Bildungssysteme müssen wohl neu stärker evidenzbasiert vom Ertrag her betrachtet werden, aber auch vom Angebot und von den Prozessen her, wobei zwischen Angebotsqualität, Prozessqualität und Ergebnisqualität kein linearer Zusammenhang besteht. Die Theorie der Expertise ist bestimmt von Annahmen, die Angebot und Nutzung in ein Verhältnis setzen und dabei ressourcenorientiert verfahren. Das Ergebnis des Unterrichts ist abhängig von der Art und Weise sowie den gegebenen Möglichkeiten der Nutzung. Beide sind in mehreren Hinsichten variabel.

Die zentrale Prämisse der Expertise stellt Bildungsstandards in einen größeren Zusammenhang. Die Steuerung durch den *Output*, also durch die Resultate des Unterrichts, wird grundsätzlich mit *Unterrichts- und Schulentwicklung* in Verbindung gebracht, steht also nicht für sich. Die Lehrgänge in den Schulen bringen nicht einfach „Ergebnisse“ hervor, sondern gestalten komplexe Lehr- und Lernprozesse. Die Qualität dieser Prozesse muss entwickelt werden und Bildungsstandards müssen dazu beitragen; sie sind daher kein Selbstzweck. Das zentrale Medium der Qualitätssicherung ist der *Unterricht*, was inzwischen als gesichertes Resultat der Forschung gelten kann. Die Erträge eines Schulsystems zu kennen, ist eine zwar hilfreiche, jedoch nicht hinreichende Bedingung zu einer Qualitätsverbesserung. So wächst unser Wissen über die Leistungsergebnisse unserer Bildungssysteme derzeit schneller als das Wissen darüber, wie diese Erkenntnisse für eine Verbesserung der Angebots- und Prozessqualität genutzt werden können.

Aus der Analyse der internationalen Forschung geht des Weiteren und ebenfalls eindeutig hervor, dass die Rückverwandlung von Output in (verbesserten) Input – was mehr bedeutet als die bloße Rückmeldung der Ergebnisse an die Akteure – eine, wenn nicht *die* zentrale Aufgabe und Herausforderung der Implementation darstellt. Eindeutig ist schließlich auch der Befund, dass diese Aufgabe nur an der Basis und unter Berücksichtigung der regionalen und der lokalen Verhältnisse gelöst werden kann. Die vorliegenden Studien zur Implementation von Bildungsprojekten zeigen, dass Politikstrategien nicht direkt umgesetzt werden können. Der Begriff der Implementation verweist auf einen Mehrebenenprozess. Wo auf zentrale Top-down-Steuerungen gesetzt wurde, ohne lokale und regionale Besonderheiten und Prozesse der Willensbildung zu beachten und ohne nachhaltige Lernprozesse vor Ort anzuregen und zu unterstützen, sind die Ergebnisse in aller Regel dürftig.

Wir gehen aus von einem historisch gewachsenen System, das trotz gewisser Tendenzen zur Globalisierung immer noch weitgehend nationalkulturell geprägt ist. Die Schulsysteme der vier Auftragsländer unterscheiden sich in erheblichem Maße und wir denken nicht, dass sie sich einem einheitlichen Modell annähern werden. Das schließt die gemeinsame Bearbeitung vergleichbarer Probleme jedoch nicht aus. Die Frage der Qualitätssicherung durch Bildungsstandards stellt sich in *allen* Ländern, wengleich die Lösungen durchaus verschieden sind. Das erklärt sich nicht zuletzt durch die Tatsache, dass die politischen Kompetenzen und die Verfahren der Systemabstimmung unterschiedlich sind.

Es macht einen Unterschied, ob die Bildungssysteme in einer föderalen Struktur entwickelt werden oder von einer Zentrale aus. Deutsche Bundeslän-

der werden zentral gesteuert, bilden untereinander aber ein föderatives System. Die viel kleineren Kantone der Schweiz sind ebenfalls föderal verbunden, werden aber vor allem auf der Ebene der Gemeinde gesteuert. Österreich und Luxemburg wiederum setzen ein nationales Ministerium voraus, das es in der Schweiz nicht gibt. In Deutschland ist Bildung nach dem politischen Entscheid zum Föderalismus mehr denn je Ländersache, zugleich ist aber auch ein Bundesministerium tätig. In jedem Land sind zudem verschiedene Ebenen der Implementation zu beachten.

Die bildungspolitische Diskussion ist, wenngleich in den Auftragsländern in je unterschiedlichem Ausmaß, durch die Ergebnisse der PISA-Lesestudie aus dem Jahre 2001 beeinflusst worden. Das zwei Jahre später vorliegende Gutachten *Zur Entwicklung nationaler Bildungsstandards* (Klieme/Avenarius/Blum et. al. 2003) hat die Konzeptentwicklung in den Auftragsländern maßgeblich bestimmt. Die internationale Diskussion selbst geht zurück auf die siebziger Jahre, als die ersten Modelle zur Entwicklung der Bildungssysteme durch Evaluation und Feedback vorgelegt wurden. Die beiden Elemente der Leistungsüberprüfung und Rückmeldung sind grundlegend für die anschließende Entwicklung, die von Verantwortung der Schulen für ihre Resultate sowie von Rechenschaftslegung (Accountability) bestimmt war.

Standards in Bildungssystemen sind historisch nichts Neues, versteht man darunter Verhaltenserwartungen, Raum- und Zeitnormen, staatliche Reglemente oder auch etablierte Medien für den Unterricht. In der Literatur wird der Standard-Begriff unterschiedlich gebraucht. Pragmatisch lässt sich zwischen einem engen Begriff von Bildungsstandards (als Ergebnisstandards) und einem weiten (Strukturen und Prozesse einschließenden) Verständnis von Standards unterscheiden. Jede fixierte und für eine bestimmte Dauer gültige Regelung kann „Standard“ genannt werden. Und man kann immer auch Niveaus unterscheiden, also einen höheren oder tieferen Standard festlegen. Typisch für die Beschreibungen von Lernleistungen sind unterschiedliche Niveaus. Auch Kompetenzstufen sind aus dieser Sicht nichts Neues, das historische Notenschema ist nichts anderes als eine abgestufte Leistungsbeschreibung.

Bildungsstandards im engeren Sinne sind in der Regel curriculare Gehalte, die für bestimmte Schulfächer und Lerndomänen als Ziele festgelegt wurden. Sie stellen einen Kern dar und sollten von allen Schülerinnen und Schülern zu einem bestimmten Zeitpunkt erreicht sein. Der gesamte Umfang eines Faches geht darüber hinaus und testfähige Bildungsstandards in der Form fachlicher Kompetenzstandards werden in den Auftragsländern auch nicht für alle Fächer entwickelt. Neben Content Standards werden auch Performance Standards und Opportunity-to-Learn Standards bestimmt. Die curricularen Gehalte werden durch Leistungen gelernt, die zu unterschiedlichen Ergebnissen führen, und die

Qualität des Lernens hängt maßgeblich von den Ressourcen oder Lernmöglichkeiten ab.

Für die europäische und insbesondere auch für die deutschsprachige Diskussion ist nach dem Gutachten von Weinert (1999) der Kompetenzbegriff maßgeblich geworden. Die schulischen Inhalte werden so gelernt, dass sich damit am Ende bestimmte Kompetenzen verbinden, die sich wiederum nach ihrer Performanz, also den Leistungen und Ergebnissen, unterscheiden. Der Kompetenzbegriff ist schulnah und präzisiert das Hauptgeschäft des Unterrichts, nämlich den Erwerb von Wissen und Können. Fachliche werden inzwischen von überfachlichen Kompetenzen unterschieden. Auch überfachliche, nicht-kognitive, indirekt leistungsbezogene Kompetenzen werden auf Schule und Unterricht bezogen. Es handelt sich dabei etwa um Lernstrategien, Diskursfähigkeiten, Lernmotive oder Lernhaltungen, nicht um vage „Schlüsselqualifikationen“.

In der Expertise wird der Begriff „Bildungsstandards“ in zwei Bedeutungen verwendet. Im engeren Sinne handelt es sich um die Erwartung von Ergebnissen oder um *Outputstandards*. Im weiteren Sinne bezieht sich der Ausdruck auf allgemeine Schulstandards, die sich auch als *Prozess- oder Inputstandards* verstehen lassen. Lehrpläne zum Beispiel beschreiben in aller Regel Standards für den Input, Lehrmittel und Lehr-Lerntheorie bzw. Didaktik fassen genauer, wie der Prozess gestaltet werden soll. Davon zu unterscheiden sind Standards, an denen sich das Ergebnis messen lassen soll. In der Schule ist jeder Input ergebnisbezogen, aber nicht jedes Ergebnis ist messbar.

Kompetenzstufenmodelle, die der Komplexität fachlicher Wissens- und Lernprozesse gerecht werden, sind wünschbar, aber auch schwierig zu erstellen. Hier bleibt die Entwicklung der nächsten Jahre abzuwarten, die vom Zusammenspiel zwischen Psychometrik, Kognitionspsychologie und Fachdidaktik geprägt sein wird. Generell spiegeln Bildungsstandards ein verändertes oder erweitertes Verständnis von Bildungszielen. Die Stoff- wird durch die Kompetenzorientierung erweitert. Inwieweit sich eine Orientierung an den Inhalten oder Stoffen der schulischen Bildung durch eine an den Kompetenzen ersetzen lässt, bleibt ebenfalls abzuwarten. Die Expertise löst dieses Problem vorläufig dadurch, dass Lehrpläne in ihrer Funktion erhalten bleiben und Bildungsstandards eine deutliche Output-Funktion erhalten. Von Vorteil ist, dass der Kompetenzbegriff domänen- oder fachgebunden verwendet wird.

Der politische Zweck der Bildungsstandards ist beim heutigen Stand der Diskussion nur bedingt klar und in den Auftragsländern durchaus verschieden, wie sich am Verhältnis von *Regelstandards* und *Mindeststandards* zeigen lässt. Das Schweizer Projekt HarmoS stützt sich auf Mindeststandards in Kernfächern,

die alle Schülerinnen und Schüler erreicht haben müssen. Der Erreichungsgrad wird an drei Zeitpunkten flächendeckend getestet, die Testresultate sind ein wichtiger Indikator, um den Förderbedarf zu ermitteln. In Deutschland sind für verschiedene Fächer Regelstandards entwickelt worden, die eine inhaltliche Norm darstellen und in Form von Kerncurricula formuliert wurden. Auch in Österreich läuft die Entwicklung auf Regelstandards hinaus, wobei der Begriff etwas anders verwendet wird als in Deutschland (Lucyshyn 2006a, S. 5f.).

Bildungsstandards formulieren gegenüber Lehrkräften, Eltern, Schülerinnen und Schülern, Bildungsadministration und Öffentlichkeit verbindliche Erwartungen in bestimmten inhaltlichen Lernbereichen. Das ist für die Lehrplangestaltung neu. Bislang wurden wohl Ziele vorgegeben, deren Erreichung wurde aber nicht überprüft. Auch wenn Teile des Lehrplans für verbindlich erklärt wurden, so blieb offen, ob und wie sie den Unterricht bestimmten. Eine Theorie und ein praktisch wirksames Konzept der Verbindlichkeit gab es nicht. Dabei ist aus der Forschung seit langem bekannt, dass Lehrkräfte nicht „nach Lehrplan“ unterrichten und daher auch nicht einfach dessen Verbindlichkeiten beachten. Zudem sind Lehrpläne oft so formuliert, dass sie leicht und mit Vorteil für die Praxis umgangen werden können.

Die Verteilung des Stoffes, also der Gegenstände und Themen des Unterrichts, wurde und wird in den Lehrplänen mit Hilfe von Beispielsammlungen beschrieben, die die Auswahl letztendlich offen lassen. Das Ergebnis war und ist eine hohe Beliebigkeit bei formal gleichen Anforderungen. Hieran vor allem entzündete sich die Diskussion der Bildungsstandards. Am Ende kann mit ganz unterschiedlichen Lernleistungen ein- und derselbe Abschluss erreicht werden, zumal auch die Notensysteme ohne Bezugsnormen arbeiten und lediglich den Klassendurchschnitt beschreiben. Mit Bildungsstandards wird in Teilen des Unterrichts versucht, Leistungserwerb und Beurteilung zu objektivieren.

Historisch neu für die Schulsysteme und die Akteure in allen vier Auftragsländern ist vor allem die Einführung von regelmäßigen Leistungstests. Unter „Output-Orientierung“ wird vielfach die Messung der je erreichten Leistungen am Ende einer bestimmten Dauer verstanden. Basis der Messung sind Tests, die die Bearbeitung von Aufgaben durch die Schülerinnen und Schüler vorsehen. Die Aufgaben kommen von außen, sie sind für diesen Zweck konstruiert worden und haben eine Validierung erfahren. Die Lehrkräfte beeinflussen diesen Prozess nicht, wohl aber können sie ihre Schülerinnen und Schüler auf Tests vorbereiten, was umso mehr der Fall ist, je wichtiger der Test wird. Es ist daher kein Zufall, dass im Mittelpunkt der Kritik an den Bildungsstandards die Tests stehen.

Nach Analyse der Literatur lassen sich grob drei Richtungen der Kritik unterscheiden: Eine fundamentale Kritik bezweifelt, dass mittels flächendeckend

eingeführter Bildungsstandards im Verbund mit Tests die Qualität des Schulsystems verbessert werde. Im Gegenteil wird befürchtet, dass die Schule unter dem Einfluss eines reduktionistischen Bildungsbegriffs in eine „technokratisch“ geprägte Passform zentral gesteuerter Tests geraten könnte und der kollektive Leistungsdruck (anstelle verstärkter individueller Förderung der Schülerinnen und Schüler) dadurch immer früher überhand nehmen werde. Diese Kritik gibt es vor allem in Deutschland, wo Bildungsstandards überhaupt noch nie getestet wurden, weniger in den Vereinigten Staaten, wo Tests zum normalen Schulalltag gehören und an die Stelle von (teils fundamentalistisch anmutenden) Grundsatzdiskussionen über das Für und Wider von Bildungsstandards empirisch gestützte Diskussionen über ihre spezifischen Funktionen und Qualitäten getreten sind. Eine eher moderate Kritik spricht den Tests eine bestimmte Funktion zu und akzeptiert sie im Sinne einer Leistungskontrolle, weist aber direkte schulbezogene Konsequenzen (z.B. öffentlich einsehbare Rankings, negative finanzielle und personelle Konsequenzen) zurück, die sich mit den Resultaten verbinden. In den Vereinigten Staaten ist dies zum Beispiel die Kritik am High-Stakes-Testing. Eine dritte Kritik ist technischer Natur. Sie bezieht sich auf schlecht konstruierte Tests. Bestimmte Tests testen nicht, was sie testen sollen, und oft übersteigen die politischen Erwartungen die Möglichkeiten auch gut konstruierter Tests.

Die Forschung zur Implementation von Tests zeigt vor allem zwei Resultate. Die Lehrkräfte nutzen die Tests und deren Ergebnisse zur Verbesserung des Unterrichts und zur Objektivierung der Leistungsbeurteilungen. Was sie ablehnen, sind Tests, deren Ergebnisse Folgen haben über das hinaus, was sie für ihre Praxis nutzen können. Wenn das Verfahren fair ist, keine Lohnfolgen drohen und genügend Spielraum vorhanden ist, mit den Daten umzugehen, dann haben Lehrkräfte, wie etwa Studien aus der Schweiz zeigen, keine Schwierigkeiten mit Leistungstests, und dies umso weniger, wenn gute Aufgaben den Test auszeichnen. Wenn Tests dagegen als repressiv oder als Bedrohung erfahren werden, lehnen die Lehrkräfte sie ab.

Die Länderberichte, vor allen die Auswertung der Forschungsliteratur aus den Vereinigten Staaten, aber auch aus England zeigen, dass eine Orientierung der Bildungspolitik mehr oder weniger *allein* an den Testresultaten Nebenwirkungen haben kann, die nicht ungewollt sind, aber unbedingt einer Risikoanalyse unterworfen werden müssen. Die Testdaten werden politisch gedeutet und sind so keineswegs neutral. Und sie sind nicht nur mit Konsequenzen für die Lehrkräfte verbunden, sondern haben breite Folgen, die bei der Einführung der Tests beachtet werden müssen. Davon sind nach Stand der Literatur vor allem drei Bereiche betroffen: die freie Wahl der Schulen, die Koppelung der Leh-

rerlöhne an den getesteten Outcome sowie die curriculare Vernachlässigung aller Bereiche, die nicht getestet werden.

Das Thema der Schulwahl (school choice), ein Thema auch in England und Schweden, impliziert einen testbasierten Qualitätsindex, der etwa in England auch mit einem öffentlichen Ranking verbunden ist. Die Eltern erkennen nicht nur die Qualität einer Schule, gemessen an deren Testresultaten und anderen Indikatoren, sondern auch deren Rang im Vergleich mit anderen Schulen. Wenn ein Testsystem eingeführt wird, ist die Folge davon eine politische Diskussion der freien Schulwahl. Der Länderbericht zu England zeigt auf, wie sich damit Forderungen nach Demokratisierung verbunden haben, welche Probleme sich stellen, eine faire Lösung zu erreichen, und welche konkreten Möglichkeiten gesucht werden müssen, um in der Problembearbeitung voranzukommen.

In den Vereinigten Staaten und in Anfängen auch in England ist das Thema der Leistungslöhne für Lehrkräfte (performance pay) inzwischen weit fortgeschritten, begleitet von scharfen Kontroversen insbesondere zwischen Schulträgern und den verschiedenen Professionen der Lehrerschaft, die sich allerdings nicht alle grundsätzlich ablehnend verhalten, sondern eher wiederum die Fairness des Verfahrens bezweifeln. Unabhängig davon: In bestimmten amerikanischen Schuldistrikten werden Lehrkräfte bereits heute nach ihren Leistungen bezahlt, ein System, das in England 1898 abgeschafft wurde. Der Grund für die Wiederkehr der Leistungslöhne sind High-Stakes-Tests. Die Lehrkräfte werden mit einem Teil ihres Lohns danach bezahlt, welche Resultate ihre Schülerinnen und Schüler in den nationalen oder bundesstaatlichen Tests erzielen.

Wenn die getesteten Fächer in der Ressourcennutzung eine Priorität erhalten, dann hat das fast automatisch einen Bedeutungsverlust für die anderen Fächer zur Folge. Das deckt sich auch mit den Ergebnissen von Studien zur Motivation der Schülerinnen und Schüler, die dann nur noch solche Tests und mithin Fächer ernst nehmen, bei denen es um etwas geht. Damit entsteht die Gefahr, dass Tests eine Umgewichtung des Curriculums bewirken. Eine solche Verschiebung der Schwerpunkte lässt sich auch durch die Neukonstruktion der Lehrpläne erreichen, wie am National Curriculum in England gezeigt werden kann. Nur bestimmte Fächer, und nicht zufällig solche, mit denen sich besonders hohe Nützlichkeitsersparungen verbinden, werden auf allen vier Key Stages angeboten.

Aus der vergleichenden Länderanalyse ergeben sich unbeschadet aller Unterschiede zwei allgemeine Erfolgsbedingungen für Strategien der Qualitätssicherung: Eine in sich konsistente Bildungspolitik auf der einen und eine sie tragende Gesetzgebung auf der anderen Seite. Beide Bedingungen treffen auf unterschiedliche Systeme und Politiken gleichermaßen zu. Wer einen Systemwan-

del will, muss die Gesetze und die damit verknüpfte Verteilung der Ressourcen verändern. Das konnte ausgehend vom englischen *Education Reform Act* von 1988 in allen vier untersuchten Ländern gezeigt werden, wenngleich die Intentionen und Folgen der Gesetze ganz unterschiedlich waren und sind.

Die vier Länder zeigen starke und schwächere Formen der Outputsteuerung. Sehr starke Formen wie in England und in einigen amerikanischen Bundesstaaten sind mit einem folgenreichen Gegensatz zwischen Politik und Profession verbunden. Dieser Gegensatz ruft alternative Entwicklungs- und Forschungsstrategien hervor. Bewiesen werden soll jeweils, dass die andere Seite Unrecht hat; die Folge sind anhaltende Spannungen zwischen zwei grundlegenden Paradigmen, die sich nicht aufeinander zurückführen lassen. Letztlich geht es in der angelsächsischen, vor allem in der amerikanischen Diskussion um die Frage, ob andere als statistische Verfahren überhaupt zulässig sind. Die Optik der Profession favorisiert gemischte Ansätze, die Tests in einen Zusammenhang mit anderen Formen der Leistung und Leistungsüberprüfung stellen.

Solche gemischten Formen gibt es in den Niederlanden und in Schweden, wobei sie generell charakteristisch sind für kontinentaleuropäische Bildungssysteme. Sie kennen bislang weder High-Stakes-Testing noch härtere Formen der Accountability, ohne deswegen in ihren Leistungen zurückzubleiben. Das politische Ziel der Verbesserung der Schülerleistungen oder in Systemen mit Best Practice der Bewahrung des hohen Leistungsstandes lässt sich offenbar auf verschiedenen Wegen erreichen. Die skandinavischen Länder erzielen durchgehend sehr gute Werte in den internationalen Vergleichen, benötigen dazu aber weder Incentives für die Lehrkräfte noch Sanktionen für die Schulen. Das System basiert auf hohem Ressourceneinsatz und einer diesem angemessenen Kontrolle, die kein grundsätzliches Misstrauen der Politik voraussetzt.

Die internationalen Erfahrungen mit Bildungsstandards lassen sich im Hinblick auf Qualitätsverbesserung folgendermaßen zusammenfassen: In angelsächsischen Ländern zeigen verschiedene Befunde, dass die Einführung von Standards und Leistungsüberprüfung durch Tests zur Verbesserung der Bildungsqualität beitragen kann. Schwächere Schulen profitieren in besonderem Maße, bei starken Schulen ist der Effekt fraglich, aber es gibt wohl keine negativen Effekte. In den Niederlanden und in den skandinavischen Ländern sind die Leistungsüberprüfung und der Kontrollcharakter von Tests weniger ausgeprägt. Die Systeme setzen stärker auf Autonomie, in Schweden inzwischen aber auch auf Wettbewerb zwischen den Schulen. Es gibt Untersuchungen, die auch in dieser Hinsicht einen Zusammenhang mit Leistungsverbesserung in bestimmten Fächern feststellen konnten.

In den Niederlanden sind externe Inspektorate seit langem die Regel. Die Schulqualität wird so regelmäßig extern überprüft. Das Standardkonzept ist allerdings wie in Schweden stärker verbunden mit Verantwortungsübernahme durch Schulen und Gemeinden. Diese Form ist anschlussfähig an die Schweiz, nicht jedoch an Deutschland und Österreich. Hier fehlen Formen der Aufsicht vor Ort, wenngleich in verschiedenen deutschen Bundesländern neu Inspektorate aufgebaut werden. Ein Kontrollsystem wie in England, also die regelmäßige Überprüfung durch eine unabhängige Behörde (Ofsted), gibt es in den Auftragsländern nicht. In der Schweiz laufen allerdings Versuche in diese Richtung. Im Blick auf die Übertragung der Resultate des Ländervergleichs muss zwischen der angelsächsischen und der skandinavischen Kultur des Umgangs mit Standards unterschieden werden.

Im vierten Teil der Expertise wird ein Modell der Implementation einer schulischen Innovation, wie sie die Einführung von Bildungsstandards darstellt, entwickelt, das interdependente Ebenen, Akteure und Prozesse unterscheidet. Ausgangspunkt ist ein Verständnis von Implementation als einem komplexen Anpassungs- oder Adaptationsprozess auf den verschiedenen Ebenen des Bildungssystems. Jenseits eines einfachen „Transfers“, einer „Verordnung“ oder der „Übernahme“ von Vorgaben geht es darum, eine komplexe Neuerung in einem mehrstufigen, ko-konstruktiven Lern- und Entwicklungsprozess unter Berücksichtigung lokaler Gegebenheiten zu rekontextualisieren und daraus sich ergebende neue Kompetenzen, Routinen, Einstellungen und Identitäten aufzubauen. Im Mittelpunkt des Prozesses, in dem Phasen der Entwicklung, Anwendung, Reflexion und Evaluation ineinander übergehen, steht die Verknüpfung von Bildungsstandards mit professionellem Lernen und Schulentwicklung. Das bestehende System soll dabei nicht künstlich herausgefordert und auf eine Weise verändert werden, die auf großen Widerstand stößt. Vielmehr sollen das System und die Akteure zum Lernen angeregt werden. Dies kann nach Meinung dieser Expertise weder mit einer hierarchisch-interventionistischen „Outside-in“- oder „Top-down“-Strategie der Bildungssteuerung noch mit einer „Bottom-up“- bzw. „Inside-out“-Strategie der Selbstregulierung und der Selbstentwicklung gelingen. Vorgeschlagen wird eine partizipationsbezogene Mischform der Regulierung bzw. von Educational Governance, gestützt auf Erkenntnisse, wonach es sich bei der Aneignung einer den Kernbereich einer Profession tangierenden Innovation um einen durch Kernideen, klare Ziele, Verfahren und Instrumente geleiteten anregungs- und unterstützungsbedürftigen individuellen und kollektiven *Lern- und Entwicklungsprozess* handelt.

Das vorgeschlagene Modell schenkt vor allem den Akteuren an der Basis große Beachtung, weil sie die Reform langfristig internalisieren und realisieren müssen. Sie verfügen de facto über große Macht, die sich die Reform, wenn sie

gelingen soll, zunutze machen muss. Dabei sind die Adressaten nicht nur die Lehrkräfte und die (lokalen) Bildungsverwaltungen, sondern auch die Eltern, die Schülerinnen und Schüler, die Medien und nicht zuletzt die lokale Öffentlichkeit. Die Kommunikation mit diesen Akteursgruppen bringt in der Regel divergierende Interpretationen mit sich, die sich auf gemeinsame Begriffe und ein bestimmtes Reformverständnis beziehen müssen. Da öffentlich-politische Reformen jedoch nicht nach kommunizierbaren einfachen Regeln, sondern in Mentalitäten ihrer Akteure, die nur zum Teil professionell geformt sind, erfolgen, reichen vorgefertigte Reformpakete und Verwaltungshandeln allein dafür nicht aus.

Hinter dem Modell steht eine Theorie der Veränderung, des individuellen und kollektiven „conceptual change“, die davon ausgeht, dass der zu einem grundlegenden Wandel erforderliche Problemdruck individuell und örtlich erzeugt wird (und werden muss) und mit Materialien und arbeitsplatzbezogenen Anreizen zum „Sich-darauf-Einlassen“ verbunden sein muss. Einsichten in die Notwendigkeit und den Sinn des Wandels müssen zuallererst aufgebaut werden, was einen wichtigen Teil der Anstrengung zur Implementation einer erweiterten Qualitätssicherung darstellt. Das in der angelsächsischen Forschungsliteratur anzutreffende sogenannte „Sense-Making“ stellt eine bildungspolitische – und lehrerbildungspolitische – Herausforderung dar, die auch damit zu tun hat, dass die Akteure oft über gegenteilige Erfahrungen und nur schwer modifizierbare Überzeugungen („Beliefs“) verfügen und jede Reform unterlaufen können. Die Einsicht zum Wandel ergibt sich nicht aus dem politischen Appell, sondern verlangt Kommunikation sowie zur Selbsterfahrung und zum Umlernen beitragende Systeme der Unterstützung.

Strategien der Implementation von Bildungsstandards zielen auf verschiedene, voneinander wechselseitig abhängige Akteursgruppen: die Lehrkräfte, die Schulleitungen, die Behörden sowie die Anbieter in der Aus- und Weiterbildung. Die Gruppen werden das neue System der Outputsteuerung unterschiedlich interpretieren, wobei ein Konzept der Kommunikation gefunden werden muss, um die Unterschiede möglichst gering zu halten. Das neue System muss verständlich dargestellt und mit überzeugenden Beispielen versehen werden. Dabei müssen verschiedene Ebenen beachtet werden. Als besonders produktiv haben sich Akteursgruppen in kooperativen Netzwerken erwiesen.

Wie besonders skandinavische und niederländische Beispiele zeigen, stellt eine erweiterte Schulautonomie mit Verantwortungsübernahme und Rechenschaftspflicht der Einzelschule eine wesentliche Voraussetzung eines standardbasierten Bildungssystems dar. Zur Autonomie gehören dabei der flexible Einsatz der zeitlichen und materiellen Ressourcen, das eigene Curriculum oder

auch die Personalrekrutierung und -entwicklung. Die Zielsteuerung steht in keinem Widerspruch dazu, sofern für effektive, transparente und faire Formen der Leistungsüberprüfung gesorgt ist. Hier bieten sich gemischte und gemäßigte Formen an, die von den Schulen produktiv angenommen werden können. Wir verweisen hier auf die OECD-Studie zum Assessment, die genau das bestätigt (OECD 2005a). Ein zentrales Problem der Leistungsüberprüfung von Schulen ist nicht die Zielsteuerung an sich, sondern die Fairness der verwendeten Verfahren.

Die Breite der Orientierung am Output spielt für die Akzeptanz eine entscheidende Rolle. Wenn sich alles nur noch auf die Resultate konzentriert, ist wenig gewonnen, aber viel verloren. Erst wenn neben den basalen Literacy-Standards und Fähigkeiten Raum bleibt für kognitives und nicht-kognitives Lernen, das sich auch auf andere Inhalte und auf multiple Ziele bezieht, ergibt der Begriff einer bildenden und zugleich autonom ihr Profil mitbestimmenden Schule inhaltlich und nicht bloß rhetorisch Sinn. Die verschiedenen Arten des Assessments müssen zur Schule als einer bildenden Institution passen und dürfen nicht als Fremdkörper erfahren werden. Inspektorate etwa müssen wie in den Niederlanden selbstverständlicher Bestandteil der Schulhauskultur sein und zugleich *in* den Schulen möglichst starke Beachtung finden.

Die Sicherung von Standards und der Umgang mit Differenz setzen eine Konkretion in der Mikropolitik voraus, die etwa von der deutschen Pädagogik selten gesehen wird. Es ist zur Implementation von Bildungsstandards nicht nur nötig, Handlungsebenen und Akteursgruppen zu unterscheiden, sondern diesen Ebenen und Gruppen auch Instrumente und Verfahren zur Verfügung zu stellen, mit denen sich das Know-how des praktischen Handelns bestimmen lässt. Daher wird in dieser Expertise größter Wert auf Werkzeuge oder Tools gelegt, die in der didaktischen Forschung inzwischen einen hohen Stellenwert erlangt haben. Die Beschreibung und Zuordnung der Werkzeuge wird im fünften Teil dieser Expertise vorgenommen. Dieser konkretisiert erstmalig, was es genau heißt, Bildungsstandards zu implementieren.

Die Palette ist bewusst breit gewählt. Die Umsetzung der Reformpolitik kann nicht mit vereinzelt Strategien erfolgen, sondern verlangt ein Maßnahmenbündel, das in seiner Systematik und seinen praktischen Schlussfolgerungen dargestellt wird. Unterschiedliche, ergebnis- und prozessbezogene Instrumente müssen als Teil eines Gesamtkonzeptes auf mehreren Ebenen des Bildungssystems im Sinne eines feedbackgesteuerten Lernens eingesetzt werden. Internationale Assessments, nationale Bildungsmonitorings und lokale Schulevaluationen dürfen dabei keine isolierten Elemente darstellen, sondern müssen aufeinander abgestimmt sein. Wie das geschehen kann und welche Forschungsdaten diesen Weg stützen, wird ausführlich dargelegt.

Der Aufwand wird sich langfristig dann lohnen, wenn es gelingt, die Ergebnisse von Tests und Schulevaluationen nicht nur als neue Form der Überprüfung, etwa durch die Benennung von Stärken und Schwächen, sondern auch zur Ingangsetzung empiriegestützter Prozesse der Systementwicklung und Weiterbildung zu nutzen. Die Implementation von Bildungsstandards ist mehr als eine wissenschaftliche und technische Herausforderung. Das Sense-Making zeigt sich hier. Die Lehrerinnen und Lehrer sowie alle anderen Akteure müssen vom Nutzen und von der Fruchtbarkeit derartiger, für sie selbst wie für das System als Ganzes bislang fremder Ansätze zuerst überzeugt werden.

Leistungsmessungen sind für eine Schule dann produktiv, wenn sie ins Feld zurückübersetzt werden und dort neuen Input und hochwertigere Prozesse auslösen. Produkt- und Prozessqualität sind wechselseitig aufeinander bezogen. Einerseits erlaubt die Qualität von Lernergebnissen diagnostische Rückschlüsse auf die Qualität vergangener Unterrichtsprozesse. Andererseits entwickelt sich Qualität immer nur in sorgfältig gestalteten Lernsituationen. Man erkennt die Qualität schulischer Lernprozesse ja nicht erst am Testresultat. Viele Qualitätsmerkmale lassen sich an der Sorgfalt schulischer Maßnahmen und damit an vielen einzelnen Prozessen und Abläufen schon sehr viel früher ablesen. Und viele Fehler könnten behoben werden, wenn den Zusammenhängen zwischen Input-, Prozess- und Outputqualitäten in der konkreten Unterrichtsarbeit, etwa durch eine bessere Lernprozessdiagnostik, sensibler Rechnung getragen würde.

Die entscheidende Gelenkstelle zwischen den beiden Seiten Output und Input im Projekt der Entwicklung und Einführung von Bildungsstandards ist eine professionell aufgebaute und gestaltete *Rückmeldekultur*, welche die einzelnen Lehrpersonen, im Prinzip auch die Schülerinnen und Schüler sowie die Eltern, erreicht. Ein systemweites Bildungsmonitoring, in dem mittels Tests, statistischer Auswertungen sowie weiterer Indikatoren der Ist-Zustand eines Bildungswesens regelmäßig und über die Zeit erfasst wird, ist für die politische Rechenschaftslegung unabdingbar, ersetzt aber keine pädagogische Prozessdiagnostik und auch keine Individualdiagnostik. Das systemische Monitoring sollte zudem durch regionale Lernstandserhebungen und lokal durchgeführte Parallelarbeiten ergänzt werden. Auch und gerade hier können vergleichende Rückmeldungen Basisinformationen zur Verbesserung der Unterrichtsqualität liefern.

Ein entscheidender Punkt bei der Gestaltung von Rückmeldungen ist, dass faire Vergleiche ermöglicht werden. Das heißt, die Testergebnisse müssen nach für die Schulleistungen bedeutsamen Kontext- oder Inputparametern (soziale Herkunft, Migrationshintergrund, Lektionenzahl) statistisch bereinigt werden. Noch besser wäre es, längsschnittlich Lernfortschritte zu messen und die unter-

schiedlichen fachspezifischen Lehr-Lernvoraussetzungen ebenfalls statistisch zu berücksichtigen. Schließlich zeigen die Erfahrungen, dass Lehrerkollegien und einzelne Lehrpersonen auch bei optimaler Differenziertheit und Individualisierung von Rückmeldungen selten in der Lage sind, diese ohne fremde Hilfe und Fortbildung handlungsbezogen in solcher Weise zu interpretieren, dass sie unmittelbar für die eigene Unterrichtsentwicklung genutzt werden können.

Was sich als Kernproblem eines *idealen Outputmodells* der Steuerung herauskristallisiert, ist – im Wechselspiel von Zielen, Tests, Rückmeldungen und Entwicklungsmaßnahmen – die Rückübersetzung von Output in Input und Prozesse. Man kann von einem Rückverflüssigungsproblem sprechen. Der Kern der Professionalität von Lehrpersonen besteht darin, Ergebnisse oder fertigen Stoff in Prozesse zu verwandeln. Deshalb ist die Implementation von Bildungsstandards zentral auf professionelles Lernen und Schul- bzw. Unterrichtsentwicklung zu beziehen. Bildungsstandards, die für sich genommen lediglich eine neue Form der Inhalts- und Zielformulierung darstellen, sowie Tests und daran geknüpfte Ergebnismeldungen können erst dann einen Beitrag zur Qualitätsentwicklung leisten, wenn ihre zielklare politische und administrative Kommunikation mit pädagogischen Initiativen zur Schulentwicklung und zum professionellen Lernen von Lehrkräften verbunden wird. Damit sind verbindliche Maßnahmen zur Unterrichtsentwicklung und hinsichtlich professioneller Fortbildung angesprochen, deren innovatives Potenzial in Modellversuchen und Forschungsvorhaben hinreichend nachgewiesen werden konnte.

Ein weiterer wichtiger Punkt ist die Messlatte oder der Umgang mit starken und schwachen Schülerinnen und Schülern. Dieser Punkt wird etwa auch durch die DESI-Studie bestätigt. Bildungsstandards in den Kernfächern sollen in ihrem Umfang so schlank sein, dass sie keinen Druck erzeugen, der die ganze Lernzeit der Schülerinnen und Schüler in Anspruch nimmt. Genau diese Gefahr besteht aber. Die Expertise verweist mehrfach auf Erfahrungen mit testbezogenem Unterricht, der leicht zur Haupterfahrung werden kann, wenn die Tests in der Dichte zunehmen und eine vorwiegend selektive Funktion erhalten. So wie es kaum gute Lehrmittel für schwache Schülerinnen und Schüler gibt, so besteht auch in den Tests die Gefahr, einen idealen Schüler zu konstruieren, der für die Schwächeren kein Orientierungspunkt sein kann.

Viele Lehrplanautoren – wie auch viele Lehrpersonen und Fachdidaktiker – überschätzen die Leistungsfähigkeit der Schülerinnen und Schüler systematisch, und das, was im Unterricht unter realistischen Bedingungen machbar ist. Es fehlt ihnen ein Gefühl dafür oder das Wissen darüber, was Schülerinnen und Schüler umfangmäßig lernen können. Genau deswegen besteht bei dieser Implementation von Bildungsstandards die Gefahr, dass es zu wenig Aufgaben für die Schwächeren gibt und das Zielniveau für die mittleren bis schwächeren

Schülerinnen und Schüler allgemein zu hoch angesetzt wird. Analysen von Tests zeigen einen ähnlichen Befund. Sie schaffen es oft nicht, auf die tatsächliche Bandbreite der Leistungsfähigkeit einzugehen. Minimalziele und Kompetenzstufen könnten hier Abhilfe schaffen.

Die Einführung von Bildungsstandards hat auch Folgen für die Entwicklung der Lehrmittel. Dieser Bereich ist in den Auftragsländern unterschiedlich geregelt. In der Schweiz gibt es staatliche Lehrmittelverlage, die auf Auftrag hin gezielt Lehrmittel entwickeln; in Deutschland herrscht das Marktprinzip, wobei Lehrmittel verschiedener Verlage eine staatliche Zulassung erlangen müssen, wenn sie im Unterricht eingesetzt werden sollen. Die Kontrolle im Hinblick auf die tatsächliche Verwendung der Lehrmittel ist eher gering, was einen Teil der Ursachen für die viel beklagte Beliebigkeit ausmachen dürfte. Andererseits sind Freiheiten und Spielräume für die Lehrkräfte ein wichtiger Faktor für das Zustandekommen von Unterrichtsqualität, die stark von der geschickten Anpassung an die gegebene Situation abhängt.

Es ist unklar, wie Verlage und Lehrbuchautoren auf Bildungsstandards und Tests reagieren werden. Das Hauptproblem wird auch hier sein, eine Verengung des Curriculums zu verhindern. Schon aus diesem Grunde sind Lehrpläne, die den Zusammenhang des gesamten schulischen Angebots darstellen, auch in Zukunft unverzichtbar. Eine offene Frage ist jedoch diejenige nach der zukünftigen Form der Lehrpläne, die möglichst schlank und informativ sein müssen, wenn sie orientieren sollen. Dies gilt auch für die Lehrmittel, deren Umfangssteigerungen, besonders in Bezug auf die Handreichungen für die Lehrkräfte, in keinem Verhältnis zur tatsächlichen Nutzung stehen. Der Effekt von Standards, Aufgaben und Tests auf die Lehrmittelproduktion bleibt abzuwarten, sollte aber politische Beachtung finden.

Der Stand der Entwicklung in den vier Auftragsländern wird im sechsten Teil der Expertise dargestellt. Angesichts des Tempos der Reformen und der unterschiedlichen Ansätze nur allein schon in den deutschen Bundesländern kann dies lediglich anhand von Beispielen erfolgen, nachdem jeweils die Problemwahrnehmung und die Grundmerkmale der Steuerungspolitik dargestellt wurden. Das größte Land, Deutschland, hat naturgemäß die größte Differenz erzeugt. Die Bundesländer gehen in der konkreten Ausgestaltung der Qualitätssicherung durchaus eigene Wege, auch wenn sich zum Teil kooperative Lösungen abzeichnen. Die grundlegende Richtung der Politik ist aber auf KMK-Ebene unumstritten.

Offene Probleme sind der Einsatz und die Funktion der Tests, die Verknüpfung der Bildungsstandards mit Schulentwicklung und die Umschichtung der Ressourcen infolge veränderter Zuständigkeiten. In Österreich wird es zu ei-

nem sehr limitierten Einsatz von Tests kommen, ebenso in Luxemburg, während in der Schweiz ähnlich wie in Deutschland zu verschiedenen Zeitpunkten flächendeckend getestet werden wird, wobei die Modalitäten noch nicht genau bestimmt sind. Es ist auch noch nicht entschieden, welche Folgen sich mit den Tests verbinden sollen und welche nicht.

Weder die freie Schulwahl noch Leistungslöhne für Lehrkräfte sind in den Auftragsländern bislang ein Thema. Die kritische Diskussion des Einsatzes von Tests konzentriert sich in Deutschland und Österreich auf die Frage, ob damit die Selektion der Schülerinnen und Schüler besser begründet werden kann oder ob die Testresultate vor allem für die Förderung eingesetzt werden sollen. Auf dieser Linie ist dann auch der Zeitpunkt der Tests umstritten. In der Schweiz ist diese Frage auf EDK-Ebene entschieden, wobei natürlich auch hier die Entwicklung in der Praxis abgewartet werden muss.

Die Befunde lassen sich so summieren: Unterschiedliche Länder haben auf der Basis ihrer geschichtlichen Voraussetzungen die Frage des anzustrebenden Bildungssystems und seiner Mechanik, aber auch der Vision seiner Weiterentwicklung sehr verschieden beantwortet. Bestehende Kulturen in einem bestimmten System lassen sich nicht einfach auflösen. Auch zwischen den Auftragsländern bestehen neben Gemeinsamkeiten zum Teil erhebliche Unterschiede. Daraus erwächst die Einsicht, dass sich kein Prinzip und keine Regel, die sich in einem Land als produktiv und wirksam erweist, kontextlos auf ein anderes Land übertragen lässt. Bildungssysteme sind historisch gewachsene soziokulturelle Konfigurationen, an die sich Bildungsreformen anschließen müssen. Es geht um Modernisierung, ohne die eigene historisch geformte Identität zu verlieren.

Bildungsstandards sind als ein zentrales Element der Qualitätssicherung anzusehen. Sie sind aber kein Selbstzweck. Ohne ko-konstruktive und kultursensible Kontextualisierung in bestehenden Systemen wird die Implementation nicht gelingen oder auf halber Strecke stehen bleiben. In optimaler Gestaltung verlangen Bildungsstandards die Verklammerung und Vernetzung von Systemebenen, Prozessen Akteuren und Werkzeugen. Der Forschungsstand lässt viele Fragen offen und ist nur in Teilen zufriedenstellend.

7.2 Schlussfolgerungen

Die Schlussfolgerungen richten sich an die Bildungspolitik. Sie benennen Punkte und Aspekte, bezüglich derer die Expertise Handlungsbedarf sieht im Hinblick auf die Gesamtstrategie der Implementation von Bildungsstandards, die Entwicklung von Werkzeugen ihrer Umsetzung, die Allokation von Kom-

petenzen und Ressourcen sowie die Kommunikation der Reform nach außen – kurz: auf eine die Implementation umgreifende Qualitätspolitik auf der Makro-, Meso- und Mikroebene des Bildungssystems. Die Schlussfolgerungen gelten in Abstufungen für alle Auftragsländer.

- (1) *Die Einführung von Bildungsstandards verlangt nach einer neuen Qualitätspolitik im Bildungswesen. Diese soll kohärent sein und sich auf alle Ebenen der Bildungssteuerung beziehen. Qualitätsentwicklung, Orientierung an Standards und eine Kultur der Evaluation sollen in ihr zusammenwirken.*

Die durch Bildungsstandards angestrebte, evidenzbasierte Qualitätspolitik umfasst nicht nur Funktionen der Kontrolle und der Rechenschaftslegung im Sinne einer Stärken-Schwächen-Analyse von Bildungssystemen, sondern auch der Qualitätssicherung und der pädagogischen Qualitätsentwicklung. Ihre Instrumente sollen sowohl angebots- als auch nachfrageorientiert sein, wobei es in einem System von „Checks and Balances“ auf ein alle Systemebenen einbeziehendes Zusammenwirken von verpflichtenden Vorschriften und Anreizen („push and pull“), von externen Vorgaben und der Stimulation professioneller Selbstverpflichtung, von kluger Führung und Selbstregulation ankommt. Dabei geht es nicht nur um neue Instrumente. Eine zentrale Größe der künftigen Schulentwicklung ist die Übernahme von Verantwortung. In den angelsächsischen Ländern ist hier von „Accountability“ die Rede. Schule und Lehrkräfte übernehmen in einem inneren Konsens Verantwortung für den Leistungsstand ihrer Schülerinnen und Schüler und legen Rechenschaft ab über Erreichtes und Nicht-Erreichtes.

- (2) *Periodisch überprüfte Bildungsstandards sind innovativ und sinnvoll, weil sie erstmals in den deutschsprachigen Ländern die Qualität der Bildungssysteme auch am Output bemessen und weil sie im Rahmen einer umfassenden Qualitätspolitik das Potenzial zu einer sichtbaren Verbesserung der Schulkultur und des Unterrichts besitzen.*

Offene demokratische Zivilgesellschaften verlangen Rechenschaft über das Funktionieren ihrer Bildungseinrichtungen. Öffentliche Schulen und Lehrerschaft als deren wichtige Glieder dürfen sich nicht länger einem Denken in Kategorien empirisch festgestellter Wirksamkeit und Zielerreichung entziehen. Schule muss eine echte Ziel- und Feedback-Steuerung ausbilden und sich damit auch an ihren Ergebnissen orientieren, was bisher nicht annähernd der Fall ist. Durch Bildungsstandards und Kompetenzorientierung werden bei einer reflektierten Umsetzung Ziele und Anforderungen an das

Lehren und Lernen in der Schule konkretisiert und als klar definierte Ergebniserwartungen in Leistungserhebungen überprüfbar gemacht. Wie internationale Forschung und Erfahrungen zeigen, kann die Einführung von Standards, standardbasierten Leistungstests und entsprechenden Schulreformen zur Verbesserung der Bildungsqualität beitragen. Periodisch überprüfte Standards fördern dabei insbesondere den Blick auf langfristige, kumulativ angelegte Lernprozesse und -ergebnisse. Die Gesamtstrategie der Implementation muss dabei von einem komplexen Feld ausgehen, das durch Bildungsstandards nicht neu erfunden, wohl aber durch ihre verbindliche Gültigkeit und ihre Überprüfung nachhaltig verändert wird. Sie muss die Eigenart bestehender Mentalitäten und eingespielter Handlungsmuster ebenso in Rechnung stellen wie kulturelle Differenzen in den Auftragsländern berücksichtigen.

- (3) *Die Implementation von Bildungsstandards soll weder ausschließlich noch primär auf die Etablierung einer Testkultur als Steuerungsmittel setzen, sondern muss in erster Linie der Entwicklung von Schule und Unterricht verpflichtet sein. Die Länder müssen sich überlegen, durch welche Unterstützungsmaßnahmen sie die Weiterentwicklung der Schulen vor allem auf der Unterrichtsebene fördern können.*

Tests sind ein wichtiges Hilfsmittel, jedoch weder der zentrale Hebel zu einer Bildungsreform noch das Fundament von Qualitätsentwicklung. Tests sind sinnvoll, wenn sie zur Qualitätsentwicklung beitragen, nicht jedoch, wenn sie zum Lernziel selbst werden. Bildungsstandards sollen nicht in erster Linie im Dienst einer immer sophistischeren Mängeldiagnose stehen, sondern zur Einleitung von Entwicklungsprozessen und zur Verbesserung der Lernergebnisse im System beitragen. So müssen Schulen mit besonderen Rahmenbedingungen (z.B. einem hohen Migrantenanteil) die Möglichkeiten, Ressourcen und Werkzeuge haben, um auf problematische Befunde reagieren zu können – z.B. durch verstärkte Inanspruchnahme von Unterstützungsmaßnahmen. Die Erreichung dieses Ziels setzt über leistungsfähige Tests hinaus den Einsatz einer breiten Palette von Werkzeugen und Steuerungsmitteln voraus. Die Politik muss der Gefahr entgegenwirken, dass im Zuge des Aufbaus eines leistungsfähigen Bildungsmonitorings umfangreiche Ressourcenverschiebungen von der Schulentwicklung hin zu Diagnostik, Controlling, Evaluation und Monitoring stattfinden. Schul- und Unterrichtsentwicklung dürfen nicht als nachgeordnete Größen oder Anhängsel einer auf Tests fixierten Qualitätspolitik betrachtet werden. Bisher wurde bei der Implementation von Bildungsstandards vor allem auf den technischen Ebenen der Standardformulierung und Testentwicklung gearbeitet. Die erforderlichen Aktivitäten auf der Schul- und Unterrichts-

ebene werden vielerorts noch unterschätzt oder kaum gesehen. Sie müssen, ebenfalls in einer langfristigen Perspektive, dringend verstärkt werden, wenn das zu etablierende Testsystem nicht wirkungslos bleiben oder sogar kontraproduktiv werden soll.

- (4) *Die Philosophie der Reform muss einem breiten, mehrdimensionalen Begriff von Bildungsstandards verpflichtet sein, der neben einem multikriterialen Verständnis von Zielen und Ergebniserwartungen auch Ressourcen, Verfahren und Lerngelegenheiten umfasst, also neben dem Output auch die Bildungsprozesse und den Input einschließt.*

Einer eindimensionalen Fachleistungsorientierung einer bildenden Schule ist ebenso entgegenzuwirken wie einer Marginalisierung von nicht durch Standards erschlossenen Fächern bzw. einer Verschärfung der klassischen Fächerhierarchie und -dominanz. Im Kern dessen, was durch Tests überprüft wird, sollen sich Bildungsstandards in erster Linie auf basale, in Schulfächern zu erreichende Ziele und Kompetenzen richten. Standards sollen dabei nicht umgeschriebene Lehrplaninhalte sein. Sie formulieren einen Kompetenzkern und nicht „alles“. Zur Erfassung „weicherer“, nicht minder wichtiger, auch überfachlicher Bildungsziele muss auf andere Assessment-Methoden, insbesondere auf Verfahren der bilanzierenden und formativen (Selbst-)Evaluation zurückgegriffen werden. Zu vermeiden ist, dass Bildungsstandards zu einer Verengung des Bildungsbegriffs und der Zielperspektive im Sinne abfragbaren Wissens und kognitiver Könnensanforderungen, zu einer wertbezogenen Hierarchisierung des Fächerkanons zu Ungunsten nicht „getesteter“ Fächer sowie zu einer an zentralen Leistungstests orientierten „Standardisierung“ des Unterrichts führen. Standards sind zu verstehen als verbindliche Maßstäbe, die nicht unterlaufen werden dürfen, sondern erreicht werden müssen. Dies erzeugt sicherlich Druck, der aber nicht so groß werden darf, dass kein Raum für anderes mehr besteht und der Unterricht auf ein Teaching-to-the-Test reduziert wird. Basiskompetenzen in den Kernfächern dürfen nicht das Einzige sein, worauf es in der Schule ankommt. Das Lernen in den Kunstfächern, in Geschichte und Literatur, Politik oder Geografie muss seine Bedeutung ebenso behalten wie die Einübung in gemeinsame Werthaltungen der Gesellschaft. Schulfächer stehen für verschiedene originale Weltzugänge und Modi des Wahrnehmens, Wissens und Denkens. Keiner dieser Zugänge darf (noch mehr) an den Rand gedrängt werden.

- (5) *Die Formulierung von auf Kompetenzmodellen beruhenden Bildungsstandards sowie die Entwicklung darauf bezogener Tests und Aufgabenpools benötigt wissenschaftlichen Sachverstand, eine entsprechende Infrastruktur und eine mehrjährige Entwicklungszeit. In diesem Sinne sind auch die bisher vorliegenden (KMK-)Bildungsstandard-Kataloge keinesfalls als Endprodukte zu betrachten, sondern müssen forschungsgestützt weiterentwickelt werden.*

Bei der Entwicklung kompetenzorientierter Fachstandards und ihrer Umsetzung in Tests zur Diagnose von Lernständen und von längerfristigen Trends handelt es sich um ein ehrgeiziges Unternehmen, das mit den bisher erreichten Entwürfen längst nicht abgeschlossen ist. Als „pièce de résistance“ hat sich dabei die auf Identifizierung und Niveaustufung von Teildimensionen gerichtete Kompetenzmodellierung erwiesen. Wieweit eine solche Aufschlüsselung in einem theoretisch strengen Sinne und auf systematische Weise in inhaltlich verschiedenen Fächern überhaupt möglich sein wird, ist eine wissenschaftlich offene Frage. Erforderlich ist zudem, dass die formulierten Bildungsstandards vor ihrer Normsetzung für sämtliche Schulen im Praxisfeld auf ihre Brauchbarkeit hin evaluiert werden. In die Evaluation müssen auch die Lehrerinnen und Lehrer systematisch einbezogen werden. Ihre Rückmeldungen zu Fehlern, Inkonsistenzen, Nutzbarkeit von Unterlagen, Verständnis- und Praxisproblemen müssen in die Überarbeitungszyklen einfließen. Wichtig ist sodann die Überprüfung der den Fachstandards beigefügten Beispielaufgaben durch die Praktikerinnen und Praktiker. Ohne systematisches und evaluatives Feedback durch die Nutzer bereits in den Entwicklungsphasen ist eine längerfristige Akzeptanz der Standards nicht zu erwarten.

- (6) *Die Implementation von Bildungsstandards soll hohe Ziele auf mehreren Ebenen in einem großen und weitläufigen System erreichen, was ohne eine langfristige Perspektive und ohne wirksame Maßnahmen des „scaling-up“ nicht möglich sein wird.*

Das Projekt „Bildungsstandards“ wird bis zu zehn Jahre Dauer in Anspruch nehmen und es verlangt einen kontinuierlichen politischen Willen während des gesamten Zeitraums. Dies gilt auch dann, wenn es Rückschläge gibt. Die Politik muss akzeptieren, dass das Reformvorhaben Zeit braucht und messbare Resultate zu Tiefenwirkung, Nachhaltigkeit, Akzeptanz in der Breite und Identifikation der Adressaten mit den Reformideen nicht innerhalb einer Legislaturperiode zu erwarten sind. Man muss den riesigen Aufwand vor Augen haben, der erforderlich ist, um mehr als eine Million Lehrkräfte in den Auftragsländern, von denen etwa die Hälfte in den nächsten zwölf Jahren durch neu auszubildende ersetzt werden

muss, mit den professionellen Handlungsfolgen der politischen Weichenstellung und dem damit verbundenen Kulturwechsel in der Steuerung vertraut zu machen.

- (7) *Einem ertragsorientierten Bildungsmonitoring und einer externen Evaluation und Zielerreichungskontrolle auf Einzelschulebene müssen inhaltlich gefüllte und rechtlich gesicherte Formen von regionaler und lokaler Schulautonomie entsprechen. Von bisherigen Formen engmaschiger bürokratischer Detailsteuerung und „Top-down“-Kontrolle, die an Kontextbedingungen vor Ort oft wenig angepasst ist, muss abgerückt werden zugunsten einer horizontalen und partizipationsbezogenen Kontextsteuerung im Bildungswesen und damit der Verankerung der Kernprozesse von Bildung und Erziehung bei denjenigen Einrichtungen, denen die konkrete Aufgabenerledigung obliegt.*

Die Philosophie einer „neuen Rechenschaftslegung“ verlangt die Kopplung einer stärkeren Ertragsorientierung schulischer Bildungsprozesse (durch verbindliche Leistungsvorgaben) mit der Gewährung von zureichenden schulorganisatorischen, curricularen und pädagogischen Gestaltungsräumen, verstanden als Freiheit, die verbindlichen Ziele auf unterschiedlichen Wegen und im Rahmen eines eigenen pädagogischen Profils zu erreichen (Prozessautonomie). Einzelschulen müssen dazu mit gesetzlich abgesicherten Kompetenzen und Ressourcen ausgestattet werden. Nur wenn eine sachgemäße Zuordnung von Verantwortlichkeit zusammen mit der Befähigung und den Mitteln, diese auch wahrzunehmen, gesichert ist, kann Eigenverantwortung tatsächlich gelebt werden. Voranzutreiben ist ein Kulturwandel weg von einer dichten, hierarchiebetonten und verwaltungsorientierten Aufsicht und Detailkontrolle hin zu mehr Vertrauen und Eigenverantwortung und damit einer stärkeren Handlungsfähigkeit der Schulen. Dies bedeutet, dass Befugnisse von administrativ vorgesetzten Stellen aufgegeben und an die einzelnen Schulen übertragen werden. Dieser Kulturwandel hin zu weniger zentraler, obrigkeitstaatlicher Detailsteuerung und -kontrolle sowie zu mehr lokaler Leitung und Autonomie ist für Deutschland vermutlich gewöhnungsbedürftig und verlangt Umbaumaßnahmen und einen langen Atem. Aus den internationalen Daten lässt sich schließen: Nur das Zusammenspiel von Output-Orientierung mit einer weitreichenden Autonomie auf den operativen Ebenen wird die Implementation einer „neuen“ Bildungssteuerung und eine damit erhoffte Steigerung des Bewusstseins von „Accountability“ in der Profession voranbringen. Die Expertise unterstützt alle politischen Entwicklungen in diese Richtung.

- (8) *Testergebnisse dürfen nicht ohne zusätzliche Informationen für die Schullaufbahnplanung genutzt und schon gar nicht mit der Mittelvergabe an Schulen oder gar mit der Beförderung der Lehrpersonen verbunden werden, weil dadurch der Nutzeffekt von Tests für die Qualitätsentwicklung und Qualitätssicherung untergraben wird. Einer Ranking- und Wettbewerbs- mentalität auf Schulebene ist entgegenzuwirken. Produktiver als zur Demotivierung der Lehrerschaft führende negative Sanktionen sind Entwicklungsanreize und die Kommunikation von „best practice“ in der Profession.*

Die Funktion von Standards darf nicht darin bestehen, den Leistungs- und Selektionsdruck auf Schülerinnen und Schüler kollektiv zu erhöhen und dabei auch die Schulen unter Konkurrenzdruck zu setzen und unter Fachleistungskriterien in Verlierer und Sieger einzuteilen. Öffentlich gemachte Schulrankings halten wir zum jetzigen Zeitpunkt deshalb für nicht produktiv. Dies aus pädagogischen und wissenschaftlichen Gründen: Je wichtiger ein quantitativer Indikator für Entscheidungen im sozialen Bereich wird, desto größer ist die Wahrscheinlichkeit, dass der soziale Prozess verzerrt und korrumpiert wird. Geradezu fatal wäre es, wenn durch die Einführung von Bildungsstandards die Konkurrenz um begehrte Abschlüsse bereits in der Grundschule einsetzen würde. Die politische Funktion der Tests muss deshalb eindeutig bestimmt sein. Eine Verbesserung der durchschnittlichen Schülerleistungen verlangt eine frühe individuelle Förderung aller Schülerinnen und Schüler. Wenn Tests dazu beitragen sollen, dann verlangt das eine klare bildungspolitische Position, welche ausschließt, dass die Resultate von Monitorings und flächendeckenden Lernstandserhebungen als Selektionsinstrumente genutzt werden. Damit (spezielle) zentrale Tests auf der Basis von Bildungsstandards auch bei Selektionsentscheidungen und für Abschlüsse eine Hilfe darstellen können, ist ein längerer Prozess der erfahrungsbezogenen Annäherung erforderlich.

- (9) *Die Lehrkräfte als wichtigste Akteure und Erfolgsgaranten der Reform müssen gewonnen werden. Dies erfordert nebst einer gemeinsam geteilten Problemwahrnehmung insbesondere ein hinreichend elaboriertes, gemeinsames Verständnis der Ziele der Reform und der durch Bildungsstandards und die damit verbundenen Tests zur Verfügung gestellten Information.*

Wenn die Standards nicht die Köpfe der Lehrerinnen und Lehrer sowie der Schulleiterinnen und Schulleiter erreichen und die Mitverantwortung zu ihrer Erreichung sich nicht ebenfalls ein Stück weit auf die Schülerinnen und Schüler sowie deren Eltern überträgt, werden sie ein rhetorisches Unternehmen und ihre Wirkung oberflächlich bleiben. Lehrpersonen werden den Nutzen von Bildungsstandards daran messen, ob diese für ihre Unter-

richtsarbeit taugen. Das gelingt nur dann, wenn klar ist und sichtbar wird, was der Zweck eines an Standards orientierten Unterrichts ist und wo die Vorteile gegenüber der bisherigen Steuerung über den Input liegen. Lehrkräfte, die zeigen alle neueren Studien, arbeiten problemzentriert und unter Handlungsdruck in einem sehr komplexen und heterogenen Berufs- und Aufgabenfeld. Wer ihnen keine besseren Lösungen für ihre bestehenden Probleme anbietet und es versäumt, sie in deren Erarbeitung und produktive Weiterentwicklung einzubeziehen und dabei gleichzeitig den Mehrwert überzeugend zu kommunizieren, wird sie nicht erreichen. Bildungsstandards müssen als produktiver Beitrag zur Verbesserung des Kerngeschäfts erscheinen, was nicht allein einen hohen technischen sowie fachdidaktisch-theoretischen, sondern vor allem einen gewaltigen fortbildnerischen (sozialen) Aufwand verlangt. Allein schon sicherzustellen, dass die durch Bildungsstandards und darauf bezogene Tests zur Verfügung gestellten Informationen von den Lehrkräften verstanden und verarbeitet werden, ist eine schwierige Aufgabe. Noch weit anspruchsvoller ist es, die daraus gezogenen Erkenntnisse in Unterrichtsqualität rückzuverflüssigen und umzusetzen.

- (10) *Ein massiver Paradigmenwechsel wie die Änderung der Steuerungsphilosophie im Bildungswesen bedarf auch massiver Unterstützung. Ohne zusätzliche Investitionen und ohne Bündelung der maximal verfügbaren Ressourcen für die Verwirklichung der Vision und für die Schaffung der dazu notwendigen Voraussetzungen wird die Implementation von Bildungsstandards keinen Erfolg haben. Die Politik wird hier Entscheidungen über Umschichtungen in den Haushalten zu treffen haben.*

Ohne für die Aufgabe neu oder durch Umlagerung bereitgestellte umfangreiche Ressourcen wird es keine Reform mit nachhaltiger Tiefenwirkung geben, die das System auch in der Breite erfasst und mit deren Kernideen sich die Akteure identifizieren. Schulen und Lehrpersonen werden nur dann zu einem produktiven Umgang mit Tests und darauf bezogenen Rückmeldungen finden und ihr Unterrichtshandeln auf Bildungsstandards einstellen, wenn leistungsfähige Unterstützungssysteme es ihnen erlauben, mit gezielter Fortbildung und Förderung auf die Ergebnisse zu reagieren. Das Idealmodell einer Outputsteuerung – verstanden als Kreisprozess von Zielbestimmung, Überprüfung, Rückmeldung und Entwicklungsmaßnahmen – hat seine größte Schwachstelle dort, wo es darum geht, Ergebnisse und Rückmeldungen in verbesserten Input rückzuverwandeln. Wenn dazu nicht umfangreiche Ressourcen bereitgestellt oder umgelagert

werden, ist mit keiner Tiefenwirkung in der Umsetzung der Reform zu rechnen. Das heißt, nur wenn zusätzlich investiert werden kann, ergibt das systematische Aufdecken von Mängeln Sinn. Die Schulen wollen guten Unterricht, dazu brauchen sie aber auch gute Unterrichtsbedingungen und Professionalisierungsangebote zu ihrer Weiterentwicklung.

- (11) *Nicht nur das Bildungssystem als solches, sondern auch seine Reformen brauchen Instrumente. Neben einer Vision einer produktiven Reformsteuerung und der Bereitschaft zur Investition von Ressourcen bedarf es zur Implementation von Bildungsstandards geeigneter Verfahren und Werkzeuge des Wissenstransfers und der Personalentwicklung. Dazu gehören auch Verfahren zur Überwindung der Isolierung der Lehrkräfte und zur Ausrichtung der Schulentwicklung an gemeinsamen Zielen.*

Ohne geeignete „Tools“, die zu den Aufgaben und zu den Schulen passen, wird wenig bewirkt. Das verlangt eine Entwicklungsarbeit, die wiederum klare politische Vorgaben erfordert. Werkzeuge, die sich nach internationalen Erfahrungen im Schulbereich bewährt haben, sind unter anderem eine unterstützende Schulleitung, die Arbeit in Fachgruppen und kooperativen Netzwerken sowie eine offene und vielfältig instrumentierte Kultur der Selbst- und Fremdevaluation. Unter den neueren Verfahren der Personalentwicklung, die sich auf die für Bildungsstandards entscheidende *Unterrichtsebene* beziehen, gehören fachpädagogische und -fachdidaktische Coaching-Angebote sowie die Arbeit mit Unterrichtsvideos zu den vielversprechenden Instrumenten des professionellen Lernens und der Unterrichtsentwicklung. Die innovativen Instrumente können zum Teil aus den Ressourcen bestehender Dienste neu zusammengestellt und aufgebaut werden. Das heißt, die organisierte Beliebigkeit bestehender Fortbildungsangebote muss reduziert und es muss in innovative Formen der Fortbildung investiert werden.

- (12) *Damit Standards ihre Wirksamkeit auf den Zielstufen entfalten, müssen auch die Unterstützungssysteme, administrative Strukturen, die Landesinstitute sowie die Anbieter in der Aus- und Weiterbildung in den Transformationsprozess einbezogen werden. Das Ziel der Politik muss es sein, für eine möglichst kohärente Abstimmung („alignment“) der Arbeit der Kontextsysteme mit den Zielen der Reforminitiative zu sorgen, ohne damit die Ausrichtung der Unterstützungssysteme allein auf diese Aufgabe konzentrierten zu wollen. Was es insbesondere braucht, sind innovative Formen einer pflichtmäßigen und zielbezogenen pädagogisch-didaktischen Fortbildung der Lehrkräfte.*

Im Blick auf Deutschland scheint ein Zusammenführen von Fortbildung und zweiter Phase der Lehrerbildung (wie beim IQSA in Schleswig-Hol-

stein) angezeigt. In der Schweiz haben die Pädagogischen Hochschulen Aufträge auch in der Weiterbildung der Lehrpersonen, die für die Implementation der Bildungsstandards genutzt werden können. Die Vorteile und das Erbe einer einphasigen, seminaristischen Lehrerbildung mit ihren ehemals engen Theorie-Praxisbezügen können dabei genutzt werden. Auch die Universitäten können einen Beitrag leisten, müssten dann aber einen erweiterten Auftrag erhalten. Allgemein sind die bestehenden Professionalisierungs- und Unterstützungssysteme wie die schulpsychologischen Dienste daraufhin zu überprüfen, welche neuen Funktionen sie im Rahmen der Implementation von Bildungsstandards übernehmen können. Hier ist eine politische Abklärung erforderlich. Aufgabe der regionalen Unterstützungs- und Kontextsysteme muss es sein, gemeinsam mit einer der Beliebigkeit entzogenen, innovativen Lehrerfortbildung die Einzelschulen mit guten Anregungen aus der Vereinzelung zu holen und durch regelmäßigen Austausch auf regionaler und überregionaler Ebene Kohärenz und Kontinuität in der Qualitätsentwicklung zu sichern.

- (13) *Da es sich bei der Einführung von Bildungsstandards um einen komplexen Designvorgang handelt mit dem Ziel, eine neue Steuerungsphilosophie in die bestehenden Bildungssysteme und deren Untersysteme verträglich zu integrieren, sind authentische Modelle und Beispiele auf allen Ebenen und bezüglich aller Facetten des Unternehmens notwendig. Insbesondere bestehende Reformprojekte von Ländern und Regionen zur Qualitätsentwicklung sollten in einen Zusammenhang mit der Implementation von Bildungsstandards gebracht, neue Projekte von Anfang an darauf ausgerichtet werden.*

Internationale Erfahrungen, aber auch Erfahrungen in Deutschland, etwa im Projekt „Selbstständige Schule“, legen es nahe, neben Einzelmaßnahmen wie dem Bildungsmonitoring auf Bundesebene auch größere qualitative Programme auf Länderebene oder in den Regionen zu fördern, die wie SINUS/SINUS-Transfer Qualitätsentwicklung und Professionalisierung unterstützen. Dies geschieht zum Beispiel durch kooperative Netzwerkbildung und Entwicklung einer Aufgabekultur. Mehrjährige regionale Reformprojekte und Modellversuche können nach empirischem Forschungsstand wirksame Instrumente der Innovation darstellen.

- (14) *Die Politik sollte beim Unternehmen „Bildungsstandards“ sowohl ihre eigenen Grenzen als auch jene der empirischen Forschung mit in Rechnung stellen. Das in einem analytisch nur schwer auflösbaren Bedingungs-, Ursachen- und Wirkungszusammenhang stehende Bildungssystem lässt sich weder durch Bildungspolitik beherrschen noch durch*

Wissenschaft einfach kalibrieren. Das Gegenteil zu meinen entspräche einer Kontrollillusion. Auf die Ebene des pädagogischen Handelns bezogen: Schulisches Lernen lässt sich immer nur sehr begrenzt steuern. Lehrkräfte stellen durch professionellen Unterricht und gute Lernbedingungen ein Angebot bereit; ob dieses ertragreich genutzt wird, hängt jedoch nicht allein von ihnen, sondern auch von den Lernenden und einem ganzen Kranz von Kontext- und Systemvariablen ab.

Nach dem in dieser Expertise begleitenden Konzept eines prinzipiell nur sehr begrenzt steuerbaren Bildungssystems sowie der Vorstellung schulischer Wirkungsmöglichkeiten im Rahmen eines Angebots-Nutzungs-Modells schulischen Handelns kommen Schulleistung und Lernerträge immer nur durch ein mehr oder weniger optimal synchronisiertes Wechselspiel von Angebots- und Nutzungsqualitäten zustande. Während die Schule für die professionelle Gestaltung ihres Angebots weitreichende Gestaltungsmöglichkeiten besitzt, sind ihr in Bezug auf die Bereitschaft zu dessen produktiver Nutzung durch eine heterogene Schülerschaft Grenzen gesetzt. Diese lassen sich immer nur „vor Ort“, in den Klassen, durch die einzelnen Lehrkräfte im Kontakt mit individuellen Schülerinnen und Schülern bearbeiten. Bildungsqualität, verstanden als Qualität der Lehr-Lernangebote, wird in der Schule und also im Unterricht hergestellt von ausgebildeten Lehrpersonen, die sich für das Lernen ihrer Schülerinnen und Schüler engagieren. Diese Motivation und Kompetenz gilt es, mit der Initiative „Bildungsstandards“ zu stärken. Im Übrigen sollte das Unternehmen aber nicht überschätzt werden. Auch wenn Bildungsstandards gegenwärtig als Dreh- und Angelpunkt einer Vielzahl von Qualitätsinitiativen betrachtet werden und sich im politischen Raum auch – durchaus nachvollziehbare, jedoch nicht überzeugende – „Kontrollillusionen“ bemerkbar machen, sind sie kein Universalproblemlöser für alle Probleme unserer Bildungssysteme. Weder übertriebene Heilserwartungen noch mancherorts geschürte Ängste und Abwehr rhetoriken sind angebracht. Bei allen Professionalisierungschancen, die Bildungsstandards ermöglichen und bei allen Impulsen, die sie zurzeit sichtbar auslösen, sollten ihre Grenzen dennoch klar gesehen werden.

- (15) *Die Implementation von Bildungsstandards verlangt nach einer strukturell starken und von der Politik unabhängigen wissenschaftlichen Bildungsforschung, die nicht nur evaluativ und nicht nur analytisch-diagnostisch ausgerichtet, sondern auch pädagogisch-entwicklungsorientiert ist und überdies imstande ist, auf fehlgehende Entwicklungen kritisch hinzuweisen.*

Was es neben einer auf Tests und Systemmonitoring ausgerichteten, von der Politik in Auftrag gegebenen Evaluationsforschung in Zukunft für die Etablierung von Wissensgrundlagen für eine Steuerung und Weiterentwicklung des Bildungswesens im Dienste eines geteilten Verständnisses öffentlicher Bildung braucht, ist eine unabhängige – ohne durch die Politik auferlegten Rechtfertigungsdruck arbeitende –, international vernetzte und konkurrenzfähige empirie- und theorieorientierte Bildungsforschung. Was die Implementation einer schulischen Neuerung, wie sie Bildungsstandards darstellen, betrifft, wissen wir z.B. immer noch zu wenig darüber, worin – innerhalb der Grenzen des Möglichen – die Klugheit einer erfolgreichen Reformsteuerung bzw. eines erfolgreichen „change managements“ besteht, konkret: mit welcher Art von Educational Governance, Schulentwicklung und Fortbildung eine nachhaltige Unterrichtsentwicklung, die in die Breite des Systems wirkt und dabei tief verwurzelte Einstellungen und Gewohnheiten ihrer langjährigen Akteure herauszufordern und zu verändern in der Lage ist, ausgelöst und gefördert werden kann. In diesem Zusammenhang gesucht ist überdies (vgl. auch Terhart 2002) ein Forschungstypus, welcher Analyse und Intervention, Systembeschreibung und handlungstheoretische Bestimmung der Möglichkeiten der Veränderung von Handlungs- und Systemstrukturen längsschnittlich verbindet und sich forschungsmethodologisch mit der Frage beschäftigt, mit welchen Strategien, Verfahren und Werkzeugen (z.B. der Lehrerbildung) sich in Bezug auf angestrebte Reformziele eine optimale Tiefen- und Breitenwirkung erreichen lässt. Zu einem solchen Forschungstypus gehören neben fachdidaktischer Entwicklungsforschung auch die Durchführung von längsschnittlich angelegten Design-Experimenten und von Interventionsstudien sowie die Unterstützung von Forschungsprojekten, die den Transfer von (analytischem) Wissen in nutzbares Handlungswissen, vorzugsweise in der Lehrerbildung, zum Gegenstand haben. Eine wichtige Bedeutung wird dabei künftig ebenfalls der Aufarbeitung verfügbarer wissenschaftlicher Erkenntnisse zu spezifischen bildungspolitischen Fragestellungen zukommen, verbunden mit der Klärung der Bedingungen ihrer Nutzbarkeit und Umsetzbarkeit in einem gegebenen Systemkontext. In diesem Zusammenhang empfehlen wir neben der Förderung von (politiknaher) Begleitforschung die vermehrte öffentliche (auch internationale) Ausschreibung von Projekten und Programmen zur *Transfer- und Innovationsforschung, einschließlich Projekten zur Governance im Bildungsbereich*, und damit zum Wechselspiel zwischen Wissensgenerierung in (grundlagen)wissenschaftlichen Kontexten und bedürfnisorientierter Wissensnutzung in bildungspolitischen Entscheidungs- und

Steuerungs- sowie fortbildungsbezogenen Lern- und Entwicklungskontexten. Schließlich besteht eine wichtige allgemeine Voraussetzung für jede im Kontext politischer Rahmensetzungen durchgeführte Bildungsforschung darin, dass die in politikgesteuerter Bildungs- und Evaluationsforschung (z.B. in Lernstandserhebungen und Bildungsmonitorings) erhobenen, zunehmend umfangreicher und exzellenter werdenden Datenbasen ohne hemmende Auflagen für Re-Analysen im Rahmen selbstdefinierter Fragestellungen zugänglich sind.

Literaturverzeichnis

- Abs, H.J./Klieme, E. (2005): Standards für schulbezogene Evaluation. In: Gogolin, I./Krüger, H.-H./Lenzen, D./Rauschenbach, Th. (Hrsg.): Standards und Standardisierung in der Erziehungswissenschaft. 4. Beiheft der Zeitschrift für Erziehungswissenschaft. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 45-62.
- Ackeren, I.v. (2003a): Evaluation, Rückmeldung und Schulentwicklung. Erfahrungen mit zentralen Tests, Prüfungen und Inspektionen in England, Frankreich und den Niederlanden. Studien zur International und Interkulturell Vergleichenden Erziehungswissenschaft, Bd. 2. Münster: Waxmann.
- Ackeren, I.v. (2003b): Nutzung großflächiger Tests für die Schulentwicklung. Bonn: Bundesministerium für Bildung und Forschung.
- Ackeren, I.v. (2004): Internationale Erfahrungen für die Entwicklung in Deutschland nutzbar machen! In: Schlömerkemper, J. (Hrsg.): Bildung und Standards. Zur Kritik der „Instandardsetzung“ des deutschen Bildungswesens. 8. Beiheft der Deutschen Schule. Weinheim: Juventa, S. 250-261.
- Ackeren, I.v. (2007): Zentrale Abschlussprüfungen. Entstehung, Struktur und Steuerungsperspektiven. In: Pädagogik 59, Nr. 3, S. 12-15.
- Adam Smith Institute (1984): Omega Report: Education Policy. London: Adam Smith Institute.
- Adamina, M. (2004): Bottom Up und Top Down – die Verschränkung von schulpraktischen und grundlegenden fachdidaktischen Anliegen bei der Entwicklung von Lern- und Lehrmaterialien. In: Aeberli, C. (Hrsg.): Lehrmittel neu diskutiert. Ergebnisse des 1. Schweizerischen Lehrmittelsymposiums vom 29. und 30. Januar 2004. Zürich: Lehrmittelverlag.
- Adey, P./Hewitt, G./Hewitt, J./Landau, N. (2004): The professional development of teachers: Practice and theory. Dordrecht: Kluwer.
- Advies van de commissie kerndoelen basisonderwijs (2002): Verantwoording delen. Herziening van de Kerndoelen basisonderwijs met het oog op beleidsruimte voor scholen. URL: <http://www.minocw.nl/documenten/brief2k-2002-doc-1342a.pdf> (Stand 8. Juli 2006).
- Aebli, H. (1951): Didactique psychologique. Application à la didactique de la psychologie de Jean Piaget. Neuchâtel: Delachaux & Niestlé.
- Aebli, H. (1961/1983): Zwölf Grundformen des Lehrens. Stuttgart: Klett.
- Ahlin, A. (2003): Does School Competition Matter? Effects of a Large-Scale School Choice Reform on Student Performance. Unpubl. paper. Uppsala University: Department of Economics.
- Aitken, M./Bennett, N./Hesketh, J. (1981): Teaching Styles and Pupil Progress. In: British Journal of Educational Psychology 52, S. 170-186.
- Akker, J.v.d./Letschert, J. (2002): Schulentwicklung in den Niederlanden. In: Pädagogik 10, S. 47-51.

- Aldrich, R. (2000): Educational Standards in Historical Perspective. In: Goldstein, H./Heath, A. (Hrsg.): Educational Standards. Oxford: Oxford University Press, S. 39-65 (= Proceedings of the British Academy Vol. 102).
- Allen, M. (2003): Eight Questions on Teacher Preparation: What Does the Research Say? Denver: Education Commission of the States.
- Allgemeiner Implementationsbrief (2005): Bildungsstandards der KMK in die Implementation der Rahmenlehrpläne für die Sekundarstufe I einbeziehen. Mittlerer Schulabschluss. LISUM Landesinstitut für Schule und Medien Brandenburg. URL: http://www.bildung-brandenburg.de/fileadmin/bbs/unterricht_und_pruefungen/rahmenlehrplaene/lisum_cd/pdf/S1-ImpBrief.pdf (Stand 12. März 2006).
- Altrichter, H. (2004): Unterrichtsentwicklung durch forschende Lehrerinnen und Lehrer. In: Gschwend, R./Claude, A. (Hrsg.): Unterrichtsentwicklung – zum Stand der Diskussion. Bern: EDK.
- Altrichter, H./Brüsemeister, Th./Wissinger, J. (Hrsg.) (2007): Educational Governance. Handlungskoordination und Steuerung im Bildungswesen. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Altrichter, H./Eder, F. (2004): Das „Autonomie-Paritäts-Muster“ als Innovationsbarriere? In: Holtappels, H.-G. (Hrsg.): Schulprogramme – Instrumente der Schulentwicklung. Juventa: Weinheim, S. 195-221.
- Altrichter, H./Posch, P. (Hrsg.) (1996): Mikropolitik der Schulentwicklung. Innsbruck: Studien Verlag.
- Altrichter, H./Posch, P. (2004): Die Diskussion um Bildungsstandards in Österreich. In: Journal für Schulentwicklung 8, Nr. 4, S. 29-38.
- Altrichter, H./Schley, W./Schratz, M. (Hrsg.) (1998): Handbuch der Schulentwicklung. Innsbruck: Studienverlag.
- Altrichter, H./Schratz, M. (2004): Bildungsstandards und die Weiterentwicklung von Unterricht und Schule. In: Erziehung und Unterricht 154, Nr. 7/8, S. 630-645.
- Altrichter, H./Wiesinger, S. (2004): Der Beitrag der Innovationsforschung im Bildungswesen zum Implementierungsproblem. In: Reinmann, G./Mandl, H. (Hrsg.): Psychologie des Wissensmanagements. Perspektiven, Theorien und Methoden. Göttingen: Hogrefe, S. 220-233.
- American Educational Research Association (AERA)/American Psychological Association/National Council on Measurement in Education. (1999): Standards for Educational and Psychological Testing. Washington, DC: AERA.
- American Federation of Teachers (AFT) (1985): Critical Thinking: It's a Basic. In: American Teacher, September, S. 21.
- American Federation of Teachers (AFT) (1995): Making Standards Matter: A Fifty-State Progress Report on Efforts to Raise Academic Standards. Washington, DC: AFT.
- American Psychological Association (2001): How Should Student Learning and Achievement be Measured? URL: <http://www.apa.org/releases/testing.html> (Stand 7. Juli 2006).

- Amrein, A.L./Berliner D.C. (2002): High-Stakes Testing, Uncertainty, and Student Learning. In: Education Policy Analysis Archives 10, Nr. 18. URL: <http://epaa.asu.edu/epaa/v10n18/> (Stand 12. März 2006).
- Anderson, L.W./Shirley, J.R. (1995): High School Principals and School Reform: Lessons Learned from a Statewide Study of Project. In: Educational Administration Quarterly 31, S. 405-423.
- Apple, M. (1992): Do the Standards Go Far Enough? Power, Policy, and Practice in Mathematics Education. In: Journal of Research in Mathematics Education 23, S. 412-431.
- Arbeitsgruppe Curriculare Standards Bildungswissenschaften beim Ministerium für Wissenschaft, Weiterbildung, Forschung und Kultur in Rheinland-Pfalz (2004): Standards der Lehrerbildung (2004): Curriculare Standards des Faches Bildungswissenschaften und Standards der Systementwicklung. Abschlussbericht. Mainz: MWWFK.
- Arbeitsgruppe Internationale Vergleichsstudie (2003): Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Staaten. Hrsg. v. Bundesministerium für Bildung und Forschung. Bonn: BMBF.
- Arnold, A. (1999): Education Development Plans: Meetings Targets and Improving Schools. Slough: NFER, EMIE.
- Association of Teachers and Lecturers (ATL) (2006): Coming Full Circle: The Impact of New Labour's Education Policies in Primary School Teachers' Work. Report Summary. London: ATL.
- Avenarius, H./Ditton, H./Döbert, H./Klemm, K./Klieme, E./Rürup, M./Tenort, H.-E./Weishaupt, H./Weiss, M. (2003): Bildungsbericht für Deutschland. Erste Befunde. Opladen: Leske+Budrich.
- Bache, I. (2003): Governing Through Governance: Education Policy Control under New Labour. In: Political Studies 51, S. 300-314.
- Baker, E.L. (2004a): Aligning Curriculum, Standards, and Assessments: Fulfilling the Promise of School Reform. University of California, Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Baker, E.L. (2004b): Reforming Education Through Policy Metaphors. In: Fitzner, Th. (Hrsg.): Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform. Bad Boll: Evangelische Akademie, S. 150-163.
- Ball, S. (1993): Culture, Cost and Control: Self-Management and Entrepreneurial Schooling in England and Wales. In: Smyth, J. (Hrsg.): A Socially Critical View of the Self-Managing School. London: Falmer, S. 70-75.
- Ball, S. (1994): Education Reform. A Critical and Post-Structural Approach. Buckingham: OUP.
- Ballantyne, P.F. (2002): Psychology, Society, and Ability Testing (1859-2002): Transformative Alternatives to Mental Darwinism and Interactionism. URL: <http://www.comnet.ca/pballan/Index/html> (Stand 2. August 2006).

- Bandura, A. (1977): Self-Efficacy: Toward a Unifying Theory of Behavioral Change. In: *Psychological Review* 84, S. 191-215.
- Bartnitzky, H. (2006): Wie VERA und Verwandtes die Bildungsqualität beschädigen. Die potemkinschen Dörfer der gegenwärtigen Schulpolitik. In: *Die Deutsche Schule* 98, S. 201-213.
- Bastian, J. (2007): *Einführung in die Unterrichtsentwicklung*. Weinheim: Beltz.
- Batteson, Ch. (1997): A Review of Politics in the „Moment of 1976“. In: *British Journal of Educational Studies* 45, S. 363-377.
- Beck, E./Guldimann, M./Zutavern, M. (1991): Eigenständig lernende Schülerinnen und Schüler. In: *Zeitschrift für Pädagogik* 37, S. 735-768.
- Beck, K. (2006): Standards – ein Mittel zur Qualitätssicherung in der Lehrerbildung? In: *Seminar* 12, Nr. 2, S. 5-20.
- Bellmann, J. (2006): Bildungsforschung und Bildungspolitik im Zeitalter „Neuer Steuerung“. In: *Zeitschrift für Pädagogik* 52, S. 487-504.
- Bennett, N. (1976): *Teaching Styles and Pupil Progress*. London: Open Books.
- Berends, M. (2000): Teacher-Reported Effects of New American Schools Design: Exploring Relationships to Teacher Background and School Context. In: *Educational Evaluation and Policy Analysis* 22, S. 65-82.
- Berends, M./Chune, J./Schuyler, G./Stockly, S./Briggs, R.J. (2002): *Challenges of Conflicting School Reforms: Effects of the New American Schools in a High-Poverty District*. Santa Monica, CA: RAND.
- Berg, R.v.d. (2002): Teachers' Meaning Regarding Educational Practice. In: *Review of Educational Research* 72, S. 577-625.
- Bericht der Landesregierung an den Bildungsausschuss des Landtags zur Lehrkräftefortbildung (2004). Potsdam, 14. April 2004.
- Berliner, D.C. (1986): In Pursuit of the Expert Pedagogue. In: *Educational Researcher* 15, Nr. 7, S. 5-13.
- Berliner, D.C./Biddle, B.J. (1995): *The Manufactured Crisis. Myths, Fraud, and the Attack on America's Public Schools*. Reading, MA: Addison Wesley Publishing Company.
- Berman, P. (1980): Thinking About Programmed and Adaptive Implementation: Matching Strategies to Situations. In: Ingram, H./Mann, D. (Hrsg.): *Why Policies Succeed or Fail*. London: Sage, S. 205-227.
- Berman, P./McLaughlin, M.W./Bass, G.V./Pauly, E./Zellman, G.L. (1977): *Federal Programs Supporting Educational Change. Vol. VII: Factors Affecting Implementation and Continuation*. Santa Monica, CA and Washington, D.C: RAND Corporation and US Office of Education.
- Berner, E./Stolz, S. (2006): *Literaturanalyse zu Entwicklung, Anwendung und insbesondere Implementation von Standards in Schulsystemen: Nordamerika*. Pädagogisches Institut der Universität Zürich, Fachbereich Allgemeine Pädagogik. Im Auftrag der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren EDK.

- URL: http://www.edk.ch/PDF_Downloads/Hamos/Literaturanalyse_3.pdf
(Stand 22. März 06).
- Bertemes, J. [o.J.]: Eingliederung des Projektes „BiStaMath“ in das Gesamtkonzept „Bildungsstandards Luxemburg“ (BiSta beta version 0.83).
- Bertschy, B. (2004): Fachdidaktische Konzeption einer berufsfeldorientierten Pädagogik für die Lehrer/-innenbildung. Bern: Peter Lang.
- Bieber, G. (2007): Damit man auch die richtigen Schlüsse ziehen kann... Aufgabenkonstruktion und Strategien der Rückmeldung bei zentralen Prüfungen. In: Pädagogik 59, Nr. 3, S. 21-25.
- Bildungsplan 2004 (2004): Realschule. Stuttgart: Ministerium für Kultus, Jugend und Sport Baden-Württemberg.
- Bildungsstandards und bayerische Lehrpläne [o.J.]: URL:
<http://www.km.bayern.de/km/schule/qualitaetssicherung/standards/index/shtml>
(Stand 10. Februar 2006).
- Björklund, A./Clark, M.A./Edin, P.-A./Fredriksson, P./Krueger, A. (2005): The Market Comes to Education in Sweden: An Evaluation of Sweden's Surprising School Reform. London: Russell Sage Foundation.
- Björklund, A./Edin, P.-A./Fredriksson, P./Krueger, A. (2004): Education, Equality, and Efficiency – An Analysis of Swedish School Reform During the 1990s. Stockholm: SNS Welfare Policy Group.
- Blömeke, S. (2006): Fast Fish – Loose Fish: International-vergleichende Forschung zur Wirksamkeit der Lehrerausbildung. In: Hilligus, A.H./Rinkens, H.-D. (Hrsg.): Standards und Kompetenzen – neue Qualität in der Lehrerbildung? Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit Verlag, S. 189-213.
- Blum, R.E. (1997): Barriers to Scaling up School Change. Paper presented at the annual conference of the American Education Research Conference, Chicaco, IL, March.
- Blum, W./Drüke-Noe, C./Hartung, R./Köller, O. (2006): Bildungsstandards Mathematik konkret Sekundarstufe I: Aufgabenbeispiele, Unterrichtsanregungen, Fortbildungsideen. Berlin: Cornelsen.
- Board of Education (Great Britain) (1905): Handbook of suggestions for the consideration of teachers. London: Board of Education.
- Board of Education (Great Britain)/Hadow, W. H. (1931): Report of the Consultative Committee on the Primary School. London: His Majesty's Stationery Office.
- Bodilly, S.J. (1996): Lessons from New American Schools Development Corporation's Demonstration Phase. Santa Monica, CA: RAND.
- Bodilly, S.J./Berends, M. (1999): Necessary District Support for Comprehensive School Reform. In: Ortfield, G./DeBray, E.H. (Hrsg.): Hard Work for Good Schools: Facts, Not Fads, in Title I Reform. Cambridge, MA: Harvard University, The Civil Rights Project, S. 111-119.

- Boekaerts, Monique (1996): Self-Regulated Learning at the Junction of Cognition and Motivation. In: *European Psychologist* 1, S. 100-112.
- Bonsen, M. (2005): Professionelle Lerngemeinschaften in der Schule. In: Holtappels, H.G./Höhmann, K. (Hrsg.): *Schulentwicklung und Schulwirksamkeit*. Weinheim, München: Juventa, S. 180-195.
- Bonsen, M./Gathen, J.v.d./Iglhaut, C./Pfeiffer, H. (2002): Die Wirksamkeit von Schulleitung. Empirische Annäherungen an ein Gesamtmodell schulischen Leitungshandelns. Weinheim, München: Juventa.
- Bonsen, M./Gathen, J.v.d./Pfeiffer, H. (2002): Wie wirkt Schulleitung? Schulleitungshandeln als Faktor für Schulqualität. In: Rolff, H.-G./Holtappels, H.G./Klemm, K./Pfeiffer, H./Schulz-Zander, R. (Hrsg.): *Jahrbuch der Schulentwicklung*, Bd. 12. Weinheim, München: Juventa, S. 287-322.
- Bonsen, M./Rolff, H.-G. (2006): Professionelle Lerngemeinschaften von Lehrerinnen und Lehrern. In: *Zeitschrift für Pädagogik* 52, S. 167-184.
- Borko, H. (2004): Professional Development and Teacher Learning: Mapping the Terrain. In: *Educational Researcher* 33, Nr. 8, S. 3-15.
- Borman, G.D./Hewes, G.M./Overman, L.T./Brown, S. (2003): Comprehensive School Reform and Achievement: A Meta-Analysis. In: *Review of Educational Research* 73, S. 125-230.
- Böttcher, W. (2003) Bildung, Standards, Kerncurricula. In: *Die Deutsche Schule* 95, Nr. 2, S. 152-164.
- Boyd, W. (1987): Public Education's Last Hurray? Schizophrenia, Amnesia, and Ignorance in School Politics. In: *Educational Evaluation and Policy Analysis* 9, S. 85-100.
- Bransford, J.D./Brown, A.L./Cocking, R.R. (Hrsg.) (1999): *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.
- Brighouse, T. (1995): The History of Inspection. In: Brighouse, T./Moon, B. (Hrsg.): *School Inspection*. London: Pitman.
- Broadfoot, P. M. (1996): *Education, Assessment, and Society. A Sociological Analysis*. Buckingham: Open University Press.
- Brockmeyer, R. (1997): Länderbericht Deutschland. In: Österreichisches Bundesministerium für Unterricht und kulturelle Angelegenheiten (Hrsg.): *Schulleitung und Schulaufsicht. Neue Rollen und Aufgaben im Schulwesen einer dynamischen und offenen Gesellschaft*. Innsbruck, Wien: Studienverlag, S. 119-161.
- Brommer, R. (1995): Was ist ‚pedagogical content knowledge‘? Kritische Anmerkungen zu einem fruchtbaren Forschungsprogramm. In: Hopmann, S./Riquarts, K. (Hrsg.): *Didaktik und/oder Curriculum*. 33. Beiheft der *Zeitschrift für Pädagogik*. Weinheim: Beltz, S. 105-112.
- Brophy, J. (Hrsg.) (2004): *Using Video in Teacher Education*. Amsterdam: Elsevier.
- Büeler, X./Buholzer, A./Kummer, A./Roos, M. (2001): Zwischenbericht 3 der wissenschaftlichen Evaluation des Luzerner Projekts „Schulen mit Profil“. Universität Zürich: FS&S.

- Buholzer, A. (2000): Das Innovationsklima an Schulen. Aarau: Sauerländer.
- Bünder, W./Parchmann, I. (2004): Lehrerarbeit in Lehrgemeinschaften. In: Beiträge zur Lehrerbildung 22, Nr. 1, S. 29-40.
- Bundesministerium für Bildung, Wissenschaft und Kultur (BMBWK) (2004): Bildungsstandards in Österreich. In: Austrian Education News 38, S. 1-5. URL: <http://www.bmbwk.gv.at/medienpool/11664/aen38.pdf> (Stand 16. März 2006).
- Burgess, S. (2005): School Choice in England. URL: www.bris.ac.uk/Depts/CMPO (Stand 20. Juli 2006).
- Callaghan, J. (1976): Speech at a Foundation Stone-Laying Ceremony at Ruskin College, Oxford, on October 18, 1976. URL: <http://education.guardian.co.uk/thegreatdebate/story/0,9860,574645,00.html> (Stand 23. Juli 2006).
- Callahan, R. (1962): Education and the Cult of Efficiency. A Study of the Social Forces that Have Shaped the Administration of the Public Schools. Chicago und London: The University of Chicago Press.
- Carlson, R. (1965): The Adaption of Educational Innovation. Eugene: Oregon Press.
- Chatterji, M. (2002): Models and Methods for Examining Standards-Based Reforms and Accountability Initiatives: Have the Tools of Inquiry Answered Pressing Questions on Improving Schools? In: Review of Educational Research 72, S. 345-386.
- Chemie im Kontext (2007). URL: <http://www.chik.de/index2.htm> (Stand 26. April 2007).
- Chitty, C. (1998): The „Moment of 1976“ Revisited. In: British Journal of Educational Studies 46, S. 318-323.
- Chubb, J.E. (Hrsg.) (2005): Within Our Reach: How America Can Educate Every Child. Lanham, MD: Rowman & Littlefield Publishers.
- Clowes, G.A. (2003): Homeschooling Update. In: School Reform News, January, S. 13-16.
- Clune, W. (1998): Toward a Theory of Systemic School Reform: The Case of Nine NSF Statewide Systemic Initiatives. Madison, WI: National Institute for Science Education.
- Coburn, C.E. (2001): Collective Sensemaking About Reading: How Teachers Mediate Reading Policy in Their Professional Communities. Educational Evaluation and Policy Analysis 23, S. 145-170.
- Coburn, C.E. (2002): Beyond Decoupling: Rethinking the Relationship between the Institutional Environment and the Classroom. Paper presented at the American Sociological Association Conference, Chicago, IL, August.
- Coburn, C.E. (2003): Rethinking Scale: Moving Beyond Numbers to Deep and Lasting Change. Educational Researcher 32, Nr. 6, S. 3-12.
- Cochran-Smith, M. (2005): The New Teacher Education: For Better or for Worse? In: Educational Researcher 34, Nr. 7, S. 3-17.

- Coffman, W.E. (1993): A King Over Egypt, Which Knew Not Joseph. In: Educational Measurement: Issues and Practice 12, Nr. 2, S. 5-8.
- Cohen, D.K./Ball, D.L. (1996): Capacity and Building Capacity for Instruction. Paper presented at the annual meeting of the American Education Research Association, New York.
- Cohen, D.K./Ball, D.L./Raudenbusch, S. W. (2000): Resources, Instruction, and Research. Seattle, WA: University of Washington, Center for the Study of Teaching and Policy.
- Cohen, D.K./Hill, H.C. (1997): Instructional Policy and Classroom Performance: The Mathematics Reform in California. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL, April.
- Cohen, D.K./Hill, H.C. (2000): Instructional Policy and Classroom Performance: The Mathematics Reform in California. Teachers College Record 102, S. 296-345.
- Cohen, D.K./Hill, H.C. (2001): Learning policy: When State Education Reform Works. New Haven, CT: Yale University Press.
- Cohen, D.K./Raudenbush, S.W./Loewenberg Ball, D. (2002): Resources, Instruction, and Research. In: Mosteller, F./Boruch, R. (Hrsg.): Evidence Matters. Randomized Trials in Education Research. Washington, DC: Brookings Institution Press, S. 80-119.
- Connell, M.W./Sheridan, K./Gardner, H. (2003): On Abilities and Domains. In: Sternberg, R.J./Grigorenko, E.L. (Hrsg.): The Psychology of Abilities, Competencies, and Expertise. Cambridge: Cambridge University Press, S. 126-155.
- Cooper, R./Slavin, R.E./Madden, N.A. (1998): Success for All: Improving the Quality of Implementation of Whole-School Change through the Use of a National Reform Network. In: Education and Urban Society 30, S. 385-408.
- Coray, C./Graf, U. (2004): E-Assessments in der Volksschule. URL: <http://www.cybersystems.ch/kunden/Artikel%20Stellwerk%20E-Assessment.pdf> (Stand 20. März 2006).
- Corcoran, T./Christman, J.B. (2002): The Limits and Contradictions of Systemic Reform: The Philadelphia Story. Philadelphia, PA: Consortium for Policy Research in Education.
- Corcoran, T./Goertz, M.E. (1995): Instructional Capacity and High Performance Schools. Educational Researcher 24, Nr. 9, S. 27-31.
- Coulby, D. (2001): Beyond the National Curriculum. London: Routledge.
- Council of Chief State School Officers (CCSSO) (2003): Glossary of Assessment Terms and Acronyms. Used in Assessing Special Education Students. Washington, DC: CCSSO.
- Covington, M.V. (1996): The Myth of Intensification. Educational Researcher 25, Nr. 8, S. 24-27.
- Cox, C. B./Dyson, A.E. (Hrsg.) (1969a): Fight for Education: A Black Paper. London: Critical Quarterly Society.

- Cox, C. B./Dyson, A.E. (Hrsg.) (1969b): *Black Paper Two: The Crisis in Education*. London: Critical Quarterly Society.
- Cox, C. B./Dyson, A.E. (Hrsg.) (1971): *The Black Papers on Education*. London: Davis-Poynter Ltd.
- Criblez, L. (2006): *Lehrpläne und Bildungsstandards: Was Schülerinnen und Schüler lernen sollten*. Festschrift zum 65. Geburtstag von Rudolf Künzli. Bern: h.e.p.
- Cuban, L. (1984): *Transforming the Frog into a Prince: Effective School Research, Policy, and Practice at the District Level*. In: *Harvard Educational Review* 54, S. 129-151.
- Cuban, L. (2004): *Looking through the Review Mirror at School Accountability*. In: Sirotnik, K. (Hrsg.): *Holding Accountability Accountable*. New York: Teachers College Press, S. 18-34.
- Cunningham, G.K. (2004): *Learning from Kentucky's Failed Accountability System*. In: Evers, W.M./Walberg, H.J. (Hrsg.): *Testing Student Learning, Evaluating Teaching Effectiveness*. Stanford, CA: Hoover Press 2004, S. 245-301.
- Curriculare Standards des Fachs Deutsch [o.J.]: *Grundlegende Empfehlungen der Arbeitsgruppe für Leitbild, Kompetenzen und Inhalte*. Mainz: MWWFK.
- Dachverband Schweizer Lehrerinnen und Lehrer (LCH) (2005a): *Manifest für ein leistungsfähiges Bildungswesen*. Verabschiedet am 11. Juni. URL: http://www.lch.ch/docs/presse/050611_Manifest.pdf (Stand 14. März 2006).
- Dachverband Schweizer Lehrerinnen und Lehrer (LCH) (2005b): *LCH-Position zu Schulleitungen vom 7. Dezember 2005*. URL: http://www.lch.ch/docs/presse/051207_LCH_Position_Schulleitungen.pdf (Stand 14. September 2006).
- Daft, R.L./Weick, K.E. (1984): *Toward a Model of Organizations as Interpretation Systems*. In: *Academy of Management Review* 9, S. 284-295.
- Dann, H.-D. (1983): *Subjektive Theorien: Irrweg oder Forschungsprogramm? Zwischenbilanz eines kognitiven Konstrukts*. In: Montada, L./Reusser, K./Steiner, G. (Hrsg.): *Kognition und Handeln*. Stuttgart: Klett-Cotta, S. 77-92.
- Dann, H.-D. (1994): *Pädagogisches Verstehen: Subjektive Theorien und erfolgreiches Handeln von Lehrkräften*. In: Reusser, K./Reusser-Weyeneth, M. (Hrsg.): *Verstehen*. Bern: Hans Huber, S. 163-182.
- Darling-Hammond, L. (1990): *Teacher Evaluation in Transition: Emerging Roles and Evolving Methods*. In: Millmann, J./Darling-Hammond, L. (Hrsg.): *The New Handbook of Teacher Evaluation: Assessing Elementary and Secondary School Teachers*. Newbury Park, CA: Sage Publications, S. 17-32.
- Darling-Hammond, L. (1994): *Performance-Based Assessment and Educational Equity*. In: *Harvard Educational Review* 64, S. 5-30.
- Darling-Hammond, L. (1998): *Teachers and Teaching: Testing Policy Hypotheses from a National Commission Report*. In: *Educational Researcher* 27, Nr. 1, S. 5-15.

- Darling-Hammond, L. (1999): Educating Teachers for the Next Century: Rethinking Practice and Policy. In: Griffin, G.A. (Hrsg.): The Education of Teachers. Ninety-eighth Yearbook of the National Society for the Study of Education. Chicago, IL: The University of Chicago Press, S. 221-256.
- Darling-Hammond, L. (2000a): How Teacher Education Matters. In: Journal of Teacher Education 51, S. 166-173.
- Darling-Hammond, L. (2000b): Teacher Quality and Student Achievement: A Review of State Policy Evidence. In: Education Policy Analysis Archives 8, Nr. 1. URL: <http://epaa.asu.edu/epaa/v8n1/> (Stand 14. September 2006).
- Darling-Hammond, L./Bransford, J.D. (Hrsg.) (2006): Preparing Teachers for a Changing World: What Teachers Should Learn and be Able to do. San Francisco: Jossey-Bass.
- Darling-Hammond, L./Loewenberg Ball, D. (1998): Teaching for High Standards: What Policymakers Need to Know and Be Able to Do. Philadelphia, PA: Consortium for Policy Research in Education.
- Darling-Hammond, L./Youngs, P. (2002): Defining „High Qualified Teachers“: What Does „scientifically-based research“ Actually Tell Us? In: Educational Researcher 31, Nr. 9, S. 13-25.
- Daschner, P. (2006): Selbständige Schulen. Entwicklungen, Balancen, Arbeitskulturen. In: Pädagogik 58, Nr. 10, S. 6-11.
- Datnow, A./Hubbard, L./Mehan, H. (1998): Educational Reform Implementation: A Co-Constructive Process. Research Report 5. Santa Cruz CA und Washington DC: Center for Research on Education, Diversity & Excellence.
- Daun, H. (2004): What Makes Swedish Education What It Is? Some Facts and Reflections. In: Fitzner, Th. (Hrsg.): Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform. Bad Boll: Evangelische Akademie, S. 62-102.
- Davis, J. (2002): The Inner London Education Authority and the William Tyndale Junior School Affair, 1974-1976. In: Oxford Review of Education 28, S. 275-298.
- Deci, E.L./Ryan, R.M. (1985). Intrinsic Motivation and Self-Determination in Human Behavior. New York: Plenum Publishing Co.
- Demaine, J. (Hrsg.) (2002): Education Policy and Contemporary Politics. London: Palgrave Macmillan.
- Demmer, M./Schweitzer, J. (2005): Es fährt ein Zug nach nirgendwo... Zwischenbilanz einer unaufhaltsamen Entwicklung. In: Becker, G./Bremerich-Vos, A./Demmer, M./Maag Merki, K./Proebe, B./Schwippert, K./Stäudel, L./Tillmann, K.-J. (Hrsg.): Standards. Unterrichten zwischen Kompetenzen, zentralen Prüfungen und Vergleichsarbeiten. Friedrich Jahresheft XXIII, S. 68-69.
- Dempster, N. (2000): Guilty or Not: The Impact and Effects of Site-Based Management on Schools. In: Journal of Educational Administration 38, S. 47-63.
- Department for Education and Employment (DfEE) (1997): Excellence in Schools. London: DfEE.

- Department for Education and Employment (DfEE) (2000): *The Role of the Local Education Authority in School Education*. Nottingham: DfEE.
- Department for Education and Employment (DfEE) (2001): *Schools: Building on Success*. Norwich: The Stationery Office.
- Department for Education and Skills (DfES) (2004a): *National Standards for Headteachers*. Nottingham: DfES.
- Department for Education and Skills (DfES) (2004b): *Five Year Strategy for Children and Learners*. URL:
<http://www.dfes.gov.uk/publications/5yearstrategy/docs/DfES5Yearstrategy.pdf> (Stand 30. April 2007).
- Department for Education and Skills (DfES) (2005): *Higher Standards, Better Schools for All. More Choice for Parents and Pupils*. Norwich: The Stationery Office.
- Department for Education and Skills (DfES) (2006a): *A Short Guide to the Education and Inspections Bill 2006*. URL:
<http://www.dfes.gov.uk/publications/educationandinspectionsbill/docs/Guide%20to%20the%20Education%20and%20Inspections%20Bill.pdf> (Stand 4. September 2006).
- Department for Education and Skills (DfES) (2006b): *The Five Year Strategy for Children and Learners: Maintaining the Excellent Progress*. URL:
<http://www.google.ch/firefox?client=firefox-a&rls=org.mozilla:de:official> (Stand 30. April 2007).
- Department for Education and Skills (DfES)/Prime Minister's Strategy Unit (2006): *School Reform: A Survey of Recent International Experience*. June. London: DfES/Prime Minister's Strategy Unit.
- Desimone, L.M. (2002): *How Can Comprehensive School Reform Models Be Successfully Implemented?* In: *Review of Educational Research* 72, S. 433-479.
- Desimone, L.M./Garet, M./Birman, B./Porter, A./Yoon, K. (2003): *Improving Teachers' In-Service Professional Development in Mathematics and Science: The Role of Postsecondary Institutions*. In: *Educational Policy* 17, S. 613-649.
- Desimone, L.M./Porter, A./Garet, M./Yoon, K./Birman, B. (2002): *Effects of Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study*. In: *Educational Evaluation and Policy Analysis* 24, S. 81-112.
- Desimone, L.M./Smith, T.M./Hayes, S.A./Frisvold, D. (2005): *Beyond Accountability and Average Mathematics Scores: Relating State Education Policy Attributes to Cognitive Achievement Domains*. *Educational Measurement: Issues and Practice* 24, Nr. 4, S. 5-18.
- Deutsches PISA-Konsortium (2001a): *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Deutsches PISA-Konsortium (2001b): *PISA 2000. Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland*. Opladen: Leske + Budrich.

- Ditton, H. (2000): Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung. *Zeitschrift für Pädagogik* 41, S. 73-92.
- Ditton, H./Merz, D./Edelhäusser, T. (2002): Einstellungen von Lehrkräften und Schulleiter/innen zu zentralen Testuntersuchungen an Schulen. In: *Empirische Pädagogik* 16, S. 17-33.
- Döbert, H. (2002): Deutschland. In: Döbert, H./Hörner, W./Kopp, B.v./Mitter, W. (Hrsg.): *Die Schulsysteme Europas*. Hohengehren: Schneider Verlag, S. 92-114.
- Doolaard, S./Karstanje, P. (2001): Gebruik van publieke prestatie-indicatoren voor schoolverbetering. In: Dijkstra, A.B./Karsten, S./Veenstra, R./Visscher, A.J. (Hrsg.): *Het oog der natie. Scholen op rapport. Standarden voor de publicatie van schoolprestaties*. Assen: Van Gorcum 2001, S. 155-172.
- Dubs, R. (2005a): *Die Führung einer Schule. Leadership und Management*. Zürich: SKV Verlag.
- Dubs, R. (2005b): *Bildungsstandards: Das Problem der schulpraktischen Umsetzung*. In: *Seminar* 11, Nr. 4, S. 15-33.
- Dunford, J./Sharp, P. (1990): *The Education System in England and Wales*. London: Longman.
- Durham Directions (2003): *System Plan 2004-2009*. Durham, ON: Durham District School Board.
- Eder, F./Thonhauser, J. (2002): Österreich. In: Döbert, H./Hörner, W./Kopp, B.v./Mitter, W. (Hrsg.): *Die Schulsysteme Europas*. Hohengehren: Schneider Verlag, S. 366-384.
- Edmonds, E.L. (1999): *The School Inspector*. London: Routledge.
- Education Act 1993 (1993): London: HMSO.
- Education Week on the Web (2005): Standards. URL: <http://www.edweek.org/rc/issues/standards/> (Stand 16. November 2005).
- Eggenberger, K./Staub, F.C. (2001): Gesichtspunkte und Strategien zur Gestaltung von Unterrichtsnachbesprechungen: Eine Fallstudie. In: *Beiträge zur Lehrerbildung* 19, S. 199-216.
- Ehmke, T./Leiss, D./Blum, W./Prenzel, M. (2006): Entwicklung von Testverfahren für die Bildungsstandards Mathematik. In: *Unterrichtswissenschaft* 34, S. 220-238.
- Eikenbusch, G. (2007): Lehrer und Schulen in Zeiten der Zentralen Prüfungen. Gewinne – Verluste – Interessen. In: *Pädagogik* 59, Nr. 3, S. 6-11.
- Ekholm, M. (1997): Steuerungsmodelle in Europa. Schwedische Erfahrungen mit alternativen Ordnungsmodellen. In: *Zeitschrift für Pädagogik* 43, S. 597-608.
- Ekholm, M. (2005): Curriculum in Sweden: Use and Discussion. In: Letschert, J.: *Curriculum Development Re-Invented*. Enschede: SLO, S. 88-96.
- Elmore, R.F. (1993): The Role of Local School Districts in Instructional Improvement. In: Fuhrman, S.H. (Hrsg.): *Designing Coherent Education Policy*. San Francisco, CA: Jossey-Bass, S. 96-124.

- Elmore, R.F. (1996): Getting to Scale with Good Educational Practice. *Harvard Educational Review* 66, S. 1-26.
- Elmore, R.F. (2002): The Testing Trap. In: *Harvard Magazine* 105, Nr. 1, S. 35.
- Elmore, R.F./Burney, D. (1997): Investing in Staff Development and Instructional Improvement in Community School District #2, New York City. New York: National Commission on Teaching and America's Future. Teachers College, Columbia University. URL: <http://www.iadb.org/int/DRP/ing/Red4/Documents/ElmoreAbril4-5-2002ing.pdf> (Stand 29. Januar 2006).
- Elmore, R.F./Burney, D. (1999): Investing in Teacher Learning: Staff Development and Instructional Improvement. In: Darling-Hammond, L./Sykes, G. (Hrsg.): *Teaching as the Learning Profession: Handbook of Policy and Practice*. San Francisco, CA: Jossey-Bass, S. 263-291.
- Erickson, G./Lander, R. (2007): Der Kitt, der ein wachsendes System zusammenhält? Nationale Tests als Kern der Qualitätssicherung in Schweden. In: *Pädagogik* 59, Nr. 3, S. 32-35.
- Erpenbach, W.J./Forte, E. (2005): Statewide Educational Accountability under the No Child Left Behind Act. A Report on 2005 Amendments to State Plans. A Report Commissioned by the CCSSO Accountability Systems and Reporting State Collaborative – SCASS. Washington, DC: Council of Chief School Officers.
- Erziehungsdepartement Kanton Schwyz, Amt für Volksschulen (2004): *Geleitete Volksschulen im Kanton Schwyz. Gemeinsam Schule gestalten, gemeinsam Schule entwickeln*. Schwyz: Stelle für Schulentwicklung.
- Eulau, H. (1969): *Micro-Macro Political Analysis*. Chicago, IL: Aldine Publishing.
- Euler, D./Sloane, P.F.E. (1998): Implementation als Problem in der Modellversuchsforschung. In: *Unterrichtswissenschaft* 26, S. 312-326.
- Evers, W./Walberg, H. (Hrsg.) (2002): *School Accountability*. Stanford, CA: Hoover Press.
- Evers, W./Walberg, H. (Hrsg.) (2004): *Testing Student Learning, Evaluating Teaching Effectiveness*. Stanford, CA: Hoover Press.
- Farkas, F./Friedman, W./Biese, J./Shaw, G. (1994): *First Things First: What Americans Expect from Public Schools*. New York: Public Agenda.
- Farnsworth, C. (2002): *Data Collection and Use in Schools*. PREL Briefing Paper. Pacific Resources for Education and Learning. Honolulu und Department of Education, Washington, DC. URL: http://www.prel.org/products/fa_/data-collection.htm (Stand 12. Dezember 2005).
- Felber, F. (2005): *Evaluationsansatz der Niederlande. Qualitätsevaluation durch das niederländische Inspektorat „Inspectie van het Onderwijs“*. Luzern: Fachstelle für Evaluation des Kanton Luzern.
- Fend, H. (1987): „Gute Schulen – schlechte Schulen“. Die einzelne Schule als pädagogische Handlungseinheit. In: *Die Deutsche Schule* 78, S. 275-293.

- Fend, H. (1998): Qualität im Bildungswesen. Schulforschung zu Systembedingungen, Schulprofilen und Lehrerleistung. Weinheim: Juventa.
- Fend, H. (2000): Qualität und Qualitätssicherung im Bildungswesen: Wohlfahrtsstaatliche Modelle und Marktmodelle. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.): Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule. 41. Beiheft der Zeitschrift für Pädagogik. Weinheim: Beltz S. 55-72.
- Fend, H. (2003): Beste Bildungspolitik oder bester Kontext für das Lernen? Über die Verantwortung von Bildungspolitik für pädagogische Wirkungen. In: Larcher, S./Oelkers, J. (Hrsg.): Die besten Ausbildungssysteme. Thematischer Bericht der Erhebung PISA 2000. Neuchâtel: Bundesamt für Statistik, S. 29-38.
- Fend, H. (2005): Systemsteuerung im Bildungswesen – Anschlussfähigkeiten an die Schulwirklichkeit. In: Maag Merki, K./Sandmeier, A./Schuler, P./Fend, H. (Hrsg.): Berichte aus dem Forschungsbereich Schulqualität & Schulentwicklung. Zürich: FS & S, Pädagogisches Institut, Universität Zürich, S. 15-27.
- Fend, H. (2006): Neue Theorie der Schule. Einführung in das Verstehen von Bildungssystemen. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Ferguson, R.F. (1991): Paying for Public Education: New Evidence on How and Why Money Matters. In: Harvard Journal on Legislation 28, S. 465-498.
- Fernandez, C./Yoshida, M. (2004): Lesson Study: A Japanese Approach to Improving Mathematics Teaching and Learning. Mahwah: Lawrence Erlbaum Associates.
- Ferrini-Mundy, J./Martin, W. G. (2000): Principles and Standards for School Mathematics. Reston: National Council of Teachers of Mathematics (NCTM).
- Finn, Ch.E. (2002): Real Accountability in K-12 Education. The Marriage of Ted and Alice. In: Evers, M.W./Walberg, H.J. (Hrsg.): School Accountability. Stanford, CA: Hoover Press, S. 23-46.
- Finn, Ch.E./Petrilli, M.J. (Hrsg.) (2000): The State of the State Standards 2000. Washington, DC: Thomas B. Fordham Foundation.
- Firestone, W.A./Fitz, J./Broadfoot, P. (1999): Power, Learning, and Legitimation: Assessment Implementation Across Levels in the United States and the United Kingdom. In: American Educational Research Journal 36, S. 759-793.
- Firestone, W.A./Pennell, J.R. (1997): Designing State-Sponsored Teacher Networks: A Comparison of Two Cases. In: American Educational Research Journal 34, S. 237-266.
- Fishman, J.J./Marx, R.W./Best, S./Tal, R. (2003): Linking Teacher and Student Learning to Improve Professional Development in Systemic Reform. In: Teaching and Teacher Education 19, S. 643-658.
- Fitz, J. (2003): The Politics of Accountability: A Perspective From England and Wales. In: Peabody Journal of Education 78, S. 230-241.
- Fletcher-Campbell, F./Lee, B. (2003): A Study of the Changing Role of Local Education Authorities in Raising Standards of Achievement in Schools. Nottingham: DfES.

- Flude, M. (1989): *The Education Reform Act, 1988: Its Origins and Implications*. London: Falmer.
- Forum Bildung (2001): *Empfehlungen des Forum Bildung*. Bonn. URL: http://www.bmbf.de/pub/empfehlungen_des_forum_bildung.pdf (Stand 13. März 2006).
- Forum on Educational Accountability (2007): *Redefining Accountability: Improving Student Learning by Building Capacity*. URL: [http://www.edaccountability.org/capacity%20building%20FEAExecSum%20\(1\).pdf](http://www.edaccountability.org/capacity%20building%20FEAExecSum%20(1).pdf) (Stand 4. Mai 2007).
- Frankfurter Erklärung vom 10. Oktober 2005. In: Frost, U. (Hrsg.) (2006): *Unternehmen Bildung. Die Frankfurter Einsprüche und kontroverse Positionen zur aktuellen Bildungsreform*. In: *Vierteljahresschrift für Wissenschaftliche Pädagogik*, Sonderheft. Paderborn: Schöningh, S. 16-21.
- Freudenthaler, H.H./Specht, W. (2005): *Bildungsstandards aus Sicht der Anwender. Evaluation der Pilotphase I zur Umsetzung nationaler Bildungsstandards in der Sekundarstufe I*. ZSE Report Nr. 69. Graz: ZSE, Abt.: Evaluation und Schulforschung.
- Freudenthaler, H./Specht, W. (2006): *Bildungsstandards: Der Implementationsprozess aus der Sicht der Praxis. Ergebnisse einer Fragebogen-Studie nach dem ersten Jahr der Pilotphase II*. ZSE Report Nr. 71. Graz: ZSE, Abt.: Evaluation und Schulforschung.
- Friedman, W./Duffet, A. (1997): *Getting By: What American Teenagers Really Think About Their Schools*. New York: Public Agenda.
- Fries, A.-V. (1998): *Über die Bindungskraft heutiger Lehrpläne*. Referat gehalten am Kongress Arbeit und Bildung SGBF, ETH und Universität Zürich, September. URL: <http://www.lehrplan.ch/d/research/results/entwicklung/fries.pdf> (Stand 16. September 2006).
- Fuhrman, S. (1999): *The New Accountability*. Consortium for Policy Research in Education (CPRE), Policy Briefs. URL: <http://www.cpre.org/Publications/rb27.pdf> (Stand 9. Januar 2006).
- Fuhrman, S. (2004): *Introduction*. In: Fuhrman, S./Elmore, R.F. (Hrsg.): *Redesigning Accountability Systems for Education*. New York: Teachers College Press.
- Fuhrman, S./O'Day, J. (1996): *Rewards and Reform: Creating Educational Incentives that Work*. San Francisco, CA: Jossey-Bass.
- Fullan, M. (1982): *The Meaning of Educational Change*. New York: Teachers College Press.
- Fullan, M. (1994): *Implementation of Innovations*. In: Husen, T./Postlethwaite, T.N. (Hrsg.): *The International Encyclopedia of Education*. 2nd Ed. Oxford: Pergamon, S. 2839-2847.
- Fullan, M. (2000): *The Return of Large-Scale Reform*. In: *Journal of Educational Change* 2, S. 1-23.

- Fullan, M. (2001): *Leading in a Culture of Change*. San Francisco, CA: Jossey-Bass.
- Fullan, M. (1991): *The New Meaning of Educational Change*. 2nd Ed. New York: Teachers College Press.
- Fullan, M./Pomfret, A. (1977): Research on Curriculum and Instruction Implementation. In: *Review of Educational Research* 47, Nr. 1, S. 335-397.
- Fürst, D. (2004): Regional Governance. In: Benz, A. (Hrsg.): *Governance – Regieren in komplexen Regelsystemen*. Wiesbaden: VS-Verlag, S. 45-64.
- Gallucci, C. (2003): Communities of Practice and the Mediation of Teachers' Response to Standards-Based Reform. In: *Education Policy Analysis Archives* 11, Nr. 35. URL: <http://epaa.asu.edu/epaa/v11n35/> (Stand 23. Juni 2006).
- Gamson, D. (1998): The Challenge of Going to Scale: A Review of the Literature. Paper presented at the Pew Forum, Pittsburg, PA.
- Gandal, M./McGiffert, L. (2003): The Power of Testing. In: *Educational Leadership* 60, Nr. 5, S. 39-42.
- Gardner, H. (1983): *Frames of Mind. The Theory of Multiple Intelligences*. New York: Basic Books.
- Garet, M./Porter, A./Desimore, L./Birman, B./Yoon, K. (2001): What Makes Professional Development Effective? Results Form a National Sample of Teachers. In: *American Educational Research*, 38, Nr. 4, S. 915-945.
- Gärtner, H. (2007): *Unterrichtsmonitoring. Evaluation eines videobasierten Qualitätszirkels zur Unterrichtsentwicklung*. Münster: Waxmann.
- Gewerkschaft Erziehung und Wissenschaft (GEW) (2004a): „Zweite Chance für die KMK“ – Gute Bildungsstandards brauchen ein Konzept, Zeit, wissenschaftlichen Sachverstand und Akzeptanz. In: Schlömerkemper, J. (Hrsg.): *Bildung und Standards. Zur Kritik der „Instandardsetzung“ des deutschen Bildungswesens*. 8. Beiheft der Deutschen Schule. Weinheim: Juventa, S. 152-183.
- Gewerkschaft Erziehung und Wissenschaft (GEW) (2004b): „Chancen nicht weiter verspielen!“ – Nur „gute“ Bildungsstandards fördern ein chancengleiches und leistungsfähiges Schulsystem. In: Schlömerkemper, J. (Hrsg.): *Bildung und Standards. Zur Kritik der „Instandardsetzung“ des deutschen Bildungswesens*. 8. Beiheft der Deutschen Schule. Weinheim: Juventa, S. 197-218.
- Giaquinta, J.B. (1973): The Process of Organizational Change in Schools. In: Kerlinger, F.N. (Hrsg.): *Review of Research in Education*, Bd. 1. Itasca, IL: Peacock, S. 178- 208.
- Gill, B.P./Timpane, M.P./Ross, K.E./Brewer, D.J. (2001): *Rhetoric versus Reality: What we Know and What we Need to Know About Voucher and Charter Schools*. Santa Monica, CA: RAND.
- Gigerenzer, G. (2000): *Adaptive Thinking. Rationality in the Real World*. Oxford: Oxford University Press.
- Glaser, R. (1981): The Future of Testing. A Research Agenda for Cognitive Psychology and Psychometrics. In: *American Psychologist* 36, S. 923-936.

- Glaser, R. (1986): The Integration of Instruction and Testing. In: Freeman, E. (Hrsg.): The Redesign of Testing in the Twenty-First Century. Proceedings of the 1985 ETS Invitational Conference. Princeton, NJ: Educational Testing Service, S. 45-58.
- Glaser, R. (1990): Testing and Assessment: O Tempora! O Mores! Pittsburgh, PA: University of Pittsburgh, Learning, Research and Development Center.
- Glennan, T.K./Bodilly, S.J./Galegher, J./Kerr, K.A. (2004): Expanding the Reach of Education Reforms. Perspectives from Leaders in the Scale-up of Educational Interventions. Santa Monica, CA: RAND.
- Glickman, C. (2002): Leadership for Learning: How to Help Teachers Succeed. Alexandria, VA: Association for Supervision and Curriculum Development.
- Goddard, R./Hoy, W.K./Woolfolk Hoy, A. (2004): Collective Efficacy Beliefs: Theoretical Developments, Empirical Evidence, and Future Directions. In: Educational Researcher 33, Nr. 3, S. 3-13.
- Goertz, M.E./Floden, R.E./O'Day, J.A. (1995): Studies of Systemic Reform. New Brunswick, NJ: Rutgers University, Consortium for Policy Research in Education.
- Goldhaber, D.D./Brewer, D.J. (2000): Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement. In: Educational Evaluation and Policy Analysis 22, S. 129-146.
- Goldstein, H. (2001): Using Pupil Performance Data for Judging Schools and Teachers: Scope and Limitations. University of London, Institute of Education. URL: <http://www.mlwin.com/hgpersonal/Using%20pupil%20performance%20data%20for%20judging%20schools%20and%20teachers.pdf> (Stand 24. März 2006).
- Gonon, Ph. (1998): Das internationale Argument in der Bildungsreform. Die Rolle internationaler Bezüge in den bildungspolitischen Debatten zur schweizerischen Berufsbildung und zur englischen Reform der Sekundarstufe II. Bern: Peter Lang (Explorationen. Studien zur Erziehungswissenschaft, hrsg. v. J. Oelkers, Bd. 20).
- Goodson, I.F./Hopmann, St./Riquarts K. (Hrsg.) (1999): Das Schulfach als Handlungsrahmen. Vergleichende Untersuchungen zur Geschichte und Funktion der Schulfächer. 7. Beiheft Bildung und Erziehung. Köln, Weimar, Wien: Böhlau.
- Gorard, S. (2006): Value-added is of Little Value. In: Journal of Education Policy 21, Nr. 2, S. 235-243.
- Grant, S. (1995): Meeting the Uneven Challenge of Reform in Mathematics and Reading. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA, April.
- Grant, S./Peterson, P./Shojgreen-Downer, A. (1996): Learning to Teach Mathematics in the Context of System Reform. In: American Educational Research Journal 33, S. 500-541.
- Gräsel, C. (1997): Problemorientiertes Lernen. Göttingen: Hogrefe.
- Gräsel, C./Jäger, R./Willke, H. (2006): Konzeption einer übergreifenden Transferforschung und Einbeziehung des internationalen Forschungsstandes. In: Nickolaus,

- R./Gräsel, C. (Hrsg.): Innovation und Transfer. Expertisen zur Transferforschung. Hohengehren: Schneider, S. 447-556.
- Gräsel, C./Parchmann, I. (2004): Implementationsforschung – oder der steinige Weg, Unterricht zu verändern. In: Unterrichtswissenschaft 32, S. 196-214.
- Green, J./Winters, M./Forster, G. (2004): Testing High-Stakes Tests: Can We Believe the Results of Accountability Tests? In: Teachers College Record 106, S. 1124-1144.
- Green, Th. (1980): Predicting the Behavior of the Educational System. New York: University of Syracuse Press.
- Greenwald, R./Hedges, L./Laine, R. (1996): The Effect of School Resources on Student Achievement. In: Review of Educational Research 66, S. 361-396.
- Griffin, G.A. (Hrsg.) (2002): Rethinking Standards Through Teacher Preparation Partnerships. Albany, NY: State University of New York Press.
- Grob, U./Maag Merki, K. (2001): Überfachliche Kompetenzen. Theoretische Grundlegung und empirische Erprobung eines Indikatorensystems. Bern: Peter Lang (Explorationen. Studien zur Erziehungswissenschaft, hrsg. v. J. Oelkers, Bd. 31).
- Gross Ophoff, J./Koch, U./Hosenfeld, I./Helmke A. (2006): Ergebnisrückmeldungen und ihre Rezeption im Projekt VERA. In: Kuper, H./Schneewind, J. (Hrsg.): Rückmeldung und Rezeption von Forschungsergebnissen. Zur Verwendung wissenschaftlichen Wissens im Bildungsbereich. Münster: Waxmann, S. 19-40.
- Grossman, P. (1996): Of Regularities and Reform: Navigating the Subject-Specific Territory of High Schools. In: McLaughlin, M.W./Oberman, I. (Hrsg.): Teacher Learning: New Policies, New Practices. New York: Teachers College Press, S. 39-47.
- Grossman, P./Thompson, C./Valencia, S. (2001): District Policy and Beginning Teachers: Where the Trains Shall Meet. Seattle, WA: University of Washington, Center for the Study of Teaching and Policy.
- Gruber, K.H. (1987): Ignoring Plowden: On the Limited Impact of the Plowden Report in Germany and Austria. In: Oxford Review of Education 13, S. 57-65.
- Gruber, K.H. (2004): Bildungsstandards: „World class“, PISA-Durchschnitt und österreichische Mindest-Standards. In: Erziehung und Unterricht 154, S. 666-677.
- Gschwend, R./Claude, A. (Hrsg.) (2004): Unterrichtsentwicklung – zum Stand der Diskussion. Bern: EDK.
- Guskey, T.R./Passaro, P.D. (1994): Teacher Efficacy: A Study of Construct Dimensions. In: American Educational Research Journal 31, S. 627-643.
- Guthrie, J.W. (Hrsg.) (1990): Educational Evaluation and Policy Analysis 12, Nr. 3 (Special issue).
- Habermas, J. (1971): Vorbereitende Bemerkungen zu einer Theorie der kommunikativen Kompetenz. In: Habermas, J./Luhmann, N. (Hrsg.): Theorie der Gesellschaft oder Sozialtechnologie? Frankfurt a.M.: Suhrkamp, S. 101-141.
- Häcker, H./Leutner, D./Amelang, M. (Hrsg.) (1998): Standards für pädagogisches und psychologisches Testen. Supplementum 1 der Diagnostica und der Zeitschrift für Differentielle und Diagnostische Psychologie.

- Haenisch, H. (1994): Bedingungen für eine erfolgreiche Umsetzung curricularer Innovationen in der Schule. Soest: Landesinstitut für Schule und Weiterbildung.
- Haider, G./Eder, F./Specht, W./Spiel, C. (2003): Zukunft Schule. Strategien und Maßnahmen zur Qualitätsentwicklung. Reformkonzept der österreichischen Zukunftskommission. Erstfassung vom 17. Oktober 2003. Wien: BMBWK. URL: http://www.bmbwk.gv.at/medienpool/10473/Konzept_Zukunft.pdf (Stand 17. März 2006).
- Haider, G./Eder, F./Specht, W./Spiel, C. (2005): Das Reformkonzept der Zukunftskommission. Abschlussbericht. Wien: MBWK. URL: http://www.bmbwk.gv.at/medienpool/12421/zk_endbericht_neu.pdf (Stand 25. Mai 2006).
- Hall, G.E./Hord, S.M. (1987): Change in Schools. Albany, NY: State University of New York Press.
- Hall, G.E./Hord, S.M. (2001): Implementing Change. Patterns, Principles, and Pot-holes. Boston, MA: Allyn & Bacon.
- Hall, P.M./McGinty, J.J.W. (1997): Policy as the Transformation of Intentions: Producing Programs from Statutes. In: The Sociological Quarterly 38, S. 439-467.
- Hallinger, P./Bickman, L./Davis, K. (1996): School Context, Principal Leadership, and Student Reading Achievement. In: The Elementary School Journal 96, S. 527-549.
- Haney, W. (2000): The Myth of the Texas Miracle in Education. In: Education Policy Analysis Archives 8, Nr. 41. URL: <http://epaa.asu.edu/epaa/v8n41/> (Stand 15. Mai 2006).
- Hansche, L.N. (1998): Meeting the Requirements of Title I: Handbook for the Development of Performance Standards. Washington, DC: U.S. Department of Education.
- Hanushek, E. (2002): Teacher Quality. In: Izumi, L./Evers, W. (Hrsg.): Teacher Quality. Palo Alto, CA: Hoover Institution, S. 1-12.
- Hargreaves, D.H. (1995): Inspection and School Improvement. In: Cambridge Journal of Education 25, S. 117-125.
- Harvey, L./Green, D. (2000): Qualität definieren. Fünf unterschiedliche Ansätze. In: Helmke, A./Hornstein, W./Terhart, E. (Hrsg.): Qualität und Qualitätssicherung im Bildungsbereich: Schule, Sozialpädagogik, Hochschule. 41. Beiheft der Zeitschrift für Pädagogik. Weinheim: Beltz, S. 17-39.
- Hatch, T. (2001): It takes Capacity to Build Capacity. In: Education Week 20, Nr. 22, S. 44-47.
- Hawley, W.D./Valli, L. (1999): The Essentials of Effective Professional Development: A New Consensus. In: Darling-Hammond, L./Sykes, G. (Hrsg.): Teaching as the Learning Profession. San Francisco, CA: Jossey-Bass, S. 127-150.
- Haynes, N.M. (1998): Lessons Learned. In: Journal of Education for Students Placed at Risk 3, S. 87-99.

- Helmke, A. (2004a): Von der Evaluation zur Innovation. In: Das Seminar 10, Nr. 2, S. 90-112.
- Helmke, A./Hosenfeld, I. (2003): Vergleichsarbeiten (VERA): eine Standortbestimmung zur Sicherung schulischer Kompetenzen. In: Schulverwaltung Hessen, Rheinland-Pfalz, Saarland, S. 10-13, 41-43. URL: http://www.uni-landau.de/vera/downloads/schulverwaltung_hrs.pdf (Stand 16. März 2006).
- Helmke, A./Hosenfeld, I. (2004): Vergleichsarbeiten – Kompetenzmodelle – Standards. In: Jäger, R.S./Frey, A./Wosnitza, M. (Hrsg.): Lernprozesse, Lernumgebungen und Lerndiagnostik. Wissenschaftliche Beiträge zum Lernen im 21. Jahrhundert. Landau: Verlag Empirische Pädagogik, S. 56-75.
- Helmke, A./Hosenfeld, I. (2005): Standardbezogene Unterrichtsevaluation. In: Brägger, G./Bucher, B./Landwehr, N. (Hrsg.): Schlüsselfragen zur externen Schulevaluation. Bern: h.e.p., S. 127-151.
- Helmke, U. (2004b): Die Entwürfe für das Fach Deutsch weisen deutliche Mängel auf. In: Schlömerkemper, J. (Hrsg.): Bildung und Standards. Zur Kritik der „Instandardsetzung“ des deutschen Bildungswesens. 8. Beiheft der Deutschen Schule. Weinheim: Juventa, S. 101-120.
- Hendriks, M.A./Doolaard, S./Bosker, R.J. (2001): Using School Effectiveness as a Knowledge Base for Self-Evaluation in Dutch Schools: the ZEBO-project. Manuskript, Enschede.
- Herrmann, U. (1991): Sozialgeschichte des Bildungswesens als Regionalanalyse. Die höheren Schulen Westfalens im 19. Jahrhundert. Köln: Böhlau.
- Herrmann, U. (2003): Bildungsstandards – Erwartungen und Bedingungen, Grenzen und Chancen. In: Zeitschrift für Pädagogik 49, S. 625-639.
- Herzog, W. (2006): Bildungsstandards – Aussichten einer Selbstverständlichkeit. Referat zur Eröffnung der Studientage der Pädagogischen Hochschule Bern vom 18. Oktober 2006.
- Hess, F. (2001): Tear Down This Wall: The Case for a Radical Overhaul of Teacher Certification. Washington, DC: Progressive Policy Institute.
- Hessisches Kultusministerium (2007): Hessischer Referenzrahmen Schulqualität (HRS). URL: http://www.bak-online.de/files/hrs_2007-02-05neu_1.pdf (Stand 30. April 2007).
- Heubert, J.P./Hauser, R.M. (Hrsg.) (1999): High Stakes: Testing for Tracking, Promotion, and Graduation. Washington, DC: National Academy Press, National Research Council, Committee on Appropriate Test Use.
- Hiebert, J./Gallimore, R./Stigler, J.W. (2002): A Knowledge Base for the Teaching Profession: What Would It Look Like and How Can We Get One? In: Educational Researcher 31, Nr. 5, S. 3-15.
- Hiebert, J./Gallimore, R./Garnier, H./Givvin, K.B./Hollingsworth, H./Jacobs, J. (2003): Teaching Mathematics in Seven Countries. Results from the TIMSS 1999

- Video Study. Washington, DC: U.S. Department of Education, National Center for Education Studies.
- Hill, H.C. (2001): Policy is Not Enough: Language and Interpretation of State Standards. In: *American Educational Research Journal* 38, S. 289-318.
- Holmes, E. (2005): *School Inspection: A Teacher's Guide for Preparing, Surviving and Evaluating Ofsted Inspection*. New Ed. London: Routledge Falmer.
- Holtappels, H.G. (2004): *Schulprogramme – Instrumente der Schulentwicklung*. München: Juventa.
- Holtappels, H.G./Rolff, H.-G. (2005): *Datenrückmeldung der Erhebung 2005, Selbständige Schule NRW*. URL: http://www.5comma7.com/bess/materialien/download/Rueckmeldung_allgemein_2005.pdf (Stand 19. April 2007).
- Hord, S.M./Rutherford, W.L./Huling-Austin, L./Hall, G.E. (1987): *From Taking Charge of Change*. Association for Supervision and Curriculum Development. Alexandria, VA.
- Hosenfeld, I./Schrader, F.-W./Helmke, T. (2006): Von der Rezeption zur Ergebnissrückmeldung: Leistungsevaluation im Spannungsfeld von System-Monitoring und Schulentwicklung. In: Hosenfeld, I./Schrader, F.-W. (Hrsg.): *Schulische Leistung. Grundlagen, Bedingungen, Perspektiven*. Münster: Waxmann, S. 289-313.
- Hovestadt, G./Kessler, N. (2005): 16 Bundesländer. Eine Übersicht zu Bildungsstandards und Evaluationen. In: Becker, G./Bremerich-Vos, A./Demmer, M./Maag Merki, K./Proebe, B./Schwippert, K./Stäudel, L./Tillmann, K.-J. (Hrsg.): *Standards. Unterrichten zwischen Kompetenzen, zentralen Prüfungen und Vergleichsarbeiten*. Friedrich Jahresheft XXIII, S. 8-10.
- Hoxby, C. (2002): *School Choice and School Productivity (or Could School Choice be a Tide that Lifts All Boats?)*. NBER Working Paper Nr. 8873. New York und Stanford: National Bureau of Economic Research.
- Hoxby, C. (2003): *School Choice and School Competition: Evidence from the United States*. In: *Swedish Economic Policy Review* 10, S. 11-66.
- Huber, A. (2005): Verbesserung der Unterrichtsqualität durch ‚Wechselseitiges Lehren und Lernen‘ (WELL). In: Schnebel, S. (Hrsg.): *Schulentwicklung im Spannungsfeld von Bildungssystem und Unterricht*. Hohengehren: Schneider.
- Huber, Ch./Späni, M./Schmellentin, C./Criblez, L. (2006): *Bildungsstandards in Deutschland, Österreich, England, Australien, Neuseeland und Südostasien. Literaturbericht zu Entwicklung, Implementation und Gebrauch von Standards in nationalen Schulsystemen*. Fachhochschule Nordwestschweiz, Pädagogische Hochschule Aarau. Im Auftrag der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren EDK. URL: http://www.edk.ch/PDF_Downloads/Hamos/Literaturanalyse_1.pdf (Stand 16. März 2006).

- Huberman, A.M./Miles, M.B. (1984): *Innovation Up Close. How School Improvement Works*. New York: Plenum.
- Humpert, W./Hauser B./Nagl, W. (2006): Was (zukünftige) Lehrpersonen über wissenschaftliche Methoden und Statistik wissen sollen und wollen. In: *Beiträge zur Lehrerbildung* 24, S. 231-245.
- Independent School Council (ISC) (2006): *ISC Census 2006*. London: Independent School Council.
- Ingersoll, R.M. (2000): *Teacher Turnover, Teacher Shortages, and the Organization of Schools*. Seattle, WA: University of Washington, Center for the Study of Teaching and Policy.
- Ingram, D./Colby, J. (1998): *Taking Standards Beyond the Classroom Door: A Process for Implementation*. Paper presented at the annual meeting of the American Education Research Association, San Diego, CA, April.
- Inspectie van het Onderwijs (2005): *Supervisory Framework for Primary Education. Content and Working Method of the Inspection Supervision*. Utrecht: Inspectie van het Onderwijs. URL: <http://194.13.31.214/Documents/pdf/EngelsToezichtkaderpo05.pdf> (Stand 12. August 2006).
- Institut für Qualitätsentwicklung an Schulen Schleswig-Holstein (IQSH) (2005): *Bilanz 2004*. Kronshagen: IQSH. URL: <http://www.infokumi.lernnetz.de/materials/PBab4PnM.pdf> (Stand 5. März 2006).
- Institut für Qualitätsentwicklung im Bildungswesen (IQB). URL: <http://www.iqb.hu-berlin.de/arbberiche/Bildungsmonitor> (Stand 27. April 2007).
- Jäger, M. (2006): *Ist-Analyse der schulischen Unterstützungssysteme. Zwei Studien zur Verbreitung von SINUS-Transfer in den am Programm beteiligten Ländern*. IPN: Kiel.
- Jäger, R.S./Mengelkamp, C. (2006): *Das Bildungsbarometer: Begründung, Implementation und erste Ergebnisse*. In: Hosenfeld, I./Schrader, F.-W. (Hrsg.): *Schulische Leistung. Grundlagen, Bedingungen, Perspektiven*. Münster: Waxmann, S. 41-63.
- James, E. (1984): *Benefits and Costs of Privatized Public Services: Lessons from the Dutch Educational System*. In: *Comparative Education Review* 28, S. 605-624.
- Joftus, S./Berman, I. (1998): *Great Expectations? Defining and Assessing Rigor in State Standards for Mathematics and English Language Arts*. Washington, DC: Council for Basic Education.
- Jones, K./Sinkinson, A. (2000): *A Critical Analysis of Ofsted Judgements of the Quality of Mathematical Initial Teacher Education Courses*. In: *Evaluation and Research in Education* 14, S. 79-93.
- Jones, M.G./Jones, B.D./Hargrove, T. (Hrsg.) (2003): *The Unintended Consequences of High-Stakes Testing*. Lanham, MD: Rowman & Littlefield Publishers.
- Kahl, R. (2004): *Treibhäuser der Zukunft. Wie in Deutschland Schulen gelingen*. Buch und 3 DVDs. Hamburg und Weinheim: Archiv der Zukunft/Beltz.

- Kallós, D. (2003): Teachers and Teacher Education in Sweden. Recent Developments. Paper presented at the international meeting „La formazione iniziale degli insegnanti in Europa. Percorsi attuali e futuri“, organized by C.I.R.E. (Centro Interdipartimentale di Ricerche Educative), Università di Bologna, January.
- Kallós, D./Lundahl-Kallós, L. (1994): Recent Changes in Teachers' Work in Sweden: Professionalization or What? In: Kallós, D./Lindblad, S. (Hrsg.): New Policy Contexts for Education: Sweden and United Kingdom. Umeå University: Educational Reports, Nr. 42, S. 140-168.
- Kanton Aargau, Departement Bildung, Kultur (2006): 80 Prozent aller 5. Klassen beteiligten sich am Check 5. URL: http://www.ag.ch/leistungstest/de/pub/check5/zweite_durchfuehrung.php (Stand 20. März 2006).
- Kanton Zürich (2005): Gesetzessammlung: Volksschulgesetz vom 7. Februar 2005. URL: http://www.vsa.zh.ch/file_uploads/bibliothek/k_453_GesetzeVerordnungenR/k_454_NeuesVSG/2128_0_2128_0_volksschulgesetz_7_2_05.pdf (Stand 21. März 2006).
- Keating, D.P. (1978): A Search for Social Intelligence. In: Journal of Educational Psychology 70, S. 218-223.
- Kelchtermans, G. (2006): Teacher Collaboration and Collegiality as Workplace Conditions. A Review. In: Zeitschrift für Pädagogik 52, S. 220-237.
- Kellaghan, Th. (Hrsg.) (2002): International Handbook of Educational Evaluation. Part 1. Dordrecht, NL: Kluwer Academic Publisher.
- Kennedy, M. (1998): Form and Substance in Inservice Teacher Education. Research Monograph No. 13, National Institute for Science Education, S. 1-22. (<http://www.msu.edu/~mkennedy/publications/valuePD.html>).
- Kent, K./Lingman, C. (2000): California's Course: Collaboration Crafts a System for Professional Learning. In: Journal of Staff Development 21, Nr. 3, S. 31-36. URL: <http://www.nsd.org/library/publications/jsd/kent213.cfm> (Stand 26. Januar 2006).
- Kieler Thesen, vorgestellt auf dem Kieler Schulleitersymposium 24./25. September 2004. URL: <http://www.infokumi.lernnetz.de/materials/jxRMTfIH.pdf> (Stand 5. März 2006).
- Klassenscockpit [o.J.]: Ein Instrument zur Qualitätssicherung in der Volksschule. St. Gallen: Kantonaler Lehrmittelverlag St. Gallen. URL: <http://www.klassenscockpit.ch/informationen/unterstufe.aspx> (Stand 20. März 2006).
- Klieme, E. (2004a): Begründung, Implementation und Wirkung von Bildungsstandards: Aktuelle Diskussionslinien und empirische Befunde. In: Zeitschrift für Pädagogik 50, S. 625-634.

- Klieme, E. (2004b): Der Beitrag von Bildungsstandards zur Qualitätssicherung und Qualitätsentwicklung in Schulen. KMK-Fachtagung „Implementation von Bildungsstandards“, Berlin, April. URL: http://www.dipf.de/publikationen/volltexte/klieme_kmk_042004.pdf (Stand 14. Mai 2006).
- Klieme, E. (2005a): Zur Bedeutung von Evaluation für die Schulentwicklung. In: Maag Merki, K./Sandmeier, A./Schuler, P./Fend, H. (Hrsg.): Berichte aus dem Forschungsbereich Schulqualität und Schulentwicklung. Zürich: FS & S, Pädagogisches Institut, Universität Zürich, S. 40-61.
- Klieme, E. (2005b): Bildungsqualität und Standards. Anmerkungen zu einem umstrittenen Begriffspaar. Friedrich Jahresheft XXIII, S. 6-7.
- Klieme, E./Avenarius, H./Blum, W./Döbrich, P./Gruber, H./Prenzel, M./Reiss, K./Riquarts, K./Rost, J./Ternorth, H.-E./Vollmer, H.J. (2003): Zur Entwicklung nationaler Bildungsstandards. Eine Expertise. Hrsg. v. Bundesministerium für Bildung und Forschung. Bonn: BMBF.
- Klieme, E./Avenarius, H./Döbert, H./van Ackeren, I./Bos, W./Klemm, K./Lehmann, R.H./von Kopp, B./Schwippert, K./Stroka, W./Weiss, M. (2003): Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten. Bonn: BMBF.
- Klieme, E./Hartig, J. (im Druck): Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. Zeitschrift für Pädagogik.
- Klieme, E./Leutner, D. (2006): Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. URL: <http://www.kompetenzdiagnostik.de/> (Stand 30. April 2007).
- Klieme, E./Reusser, K./Pauli, Ch. (2003): Unterrichtsqualität und mathematisches Verständnis im internationalen Vergleich. Ein Forschungsprojekt und erste Schritte zur Realisierung. In: Unterrichtswissenschaft 31, S. 194-205.
- Klippert, H. (2004): Lehrerbildung. Unterrichtsentwicklung und der Aufbau neuer Routinen. Weinheim: Beltz.
- Knapp, M.S. (1995): Education Policy and the Improvement of Teaching: Two Accounts of the Early Implementation of the California Mathematics Framework. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA, April.
- Knapp, M.S. (1997): Between Systemic Reform and the Mathematics and Science Classroom: The Dynamics of Innovation, Implementation, and Professional Learning. In: Review of Educational Research 67, S. 227-266.
- Knapp, M.S. (2002): Understanding How Policy Meets Practice: Two Takes on Local Responses to a State Reform Initiative. University of Washington, Center for the Study of Teaching and Policy. URL: <http://depts.washington.edu/ctpmail/PDFs/PolicyPractice-MSK-06-2002.pdf> (Stand 12. Januar 2006).

- Knapp, M.S./Bamburg, J.D./Ferguson, M.C./Hill, P. (1998): Converging Reforms and the Working Lives of Frontline Professional in Schools. In: *Educational Policy* 12, S. 397-418.
- Kobarg, M./Schwindt, K./Seidel, T. (2007): Lehrpersonen analysieren Unterricht – Die Bedeutung der persönlichen Relevanz des Videomaterials für die Analyse von Unterricht. 4. Tagung der Sektion „Empirische Bildungsforschung“ der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE). Wuppertal: 19.-21. März 2007.
- Koerner, J. (1968): *Who Controls American Education?* Boston: Beacon Press.
- Kogan, M./Maden, M. (1999): An Evaluation of Evaluators: The Ofsted System of School Inspection. In: Cullingford, C. (Hrsg.): *An Inspector Calls. Ofsted and its Effect on School Standards.* London: Kogan Page, S. 9-31.
- Köller, O. (2005): Neue Besen kehren gut: Das Institut zur Qualitätsentwicklung im Bildungswesen soll die nationalen Bildungsstandards in Deutschland überprüfen. In: *Psychologie in Erziehung und Unterricht* 52, S. 281-286.
- Köller, O./Köller, M. (im Druck): Bedeutung von Schulleiterinnen und Schulleitern für die Qualität von Unterricht. In: Riecke-Baulecke, T. (Hrsg.): *SchulleitungPlus: Grundsätze und Verfahren wirksamer Führung.* München: Oldenbourg.
- Komisar, B.P. (1966): The Paradox of Equality in Schooling. In: *Teachers College Record*, S. 251-254.
- Komisar, B.P./McClellan, J.E. (1961): The Logic of Slogans. In: Smith, B.O./Ennis, R.H. (Hrsg.): *Language and Concepts in Education.* Chicago: Rand McNally 1968, S. 194-214.
- Kornhaber, M.L. (2004): Appropriate and Inappropriate Forms of Testing, Assessment, and Accountability. In: *Educational Policy* 18, S. 45-70.
- Korthagen, F.A.J./Kessels, J.P.A.M. (1999): Linking Theory and Practice: Changing the Pedagogy of Teacher Education. In: *Educational Researcher* 28, Nr. 4, S. 4-17.
- Kos, O./Lehmann, R./Brenstein, E./Holtsch, D. (2005): *Bildungsportale – Wegweiser im Netz: Gestaltung – Nutzung – Evaluation.* Frankfurt a.M.: Peter Lang.
- Kotthoff, H.-G. (2003): *Bessere Schule durch Evaluation? Internationale Erfahrungen. Studien zur International und Interkulturell Vergleichenden Erziehungswissenschaft, Bd. 1.* Münster: Waxmann.
- Krammer, K./Hugener, I. (2005): Netzbasierte Reflexion von Unterrichtsvideos in der Ausbildung von Lehrpersonen. Eine Explorationsstudie. In: *Beiträge zur Lehrerbildung* 23, S. 51-61.
- Krammer, K./Reusser, K. (2004): Unterrichtsvideos als Medium der Lehrerinnen- und Lehrerbildung. In: *Das Seminar* 10, Nr. 4, S. 80-101.
- Krammer, K./Reusser, K. (2005): Unterrichtsvideos als Medium der Aus- und Weiterbildung von Lehrpersonen. In: *Beiträge zur Lehrerbildung* 23, S. 35-50.
- Krammer, K./Schnetzer, C.L./Ratzka, N./Pauli, C./Reusser, K. (2007): Veränderungen im Wissen von Lehrpersonen durch videobasierte Unterrichtsanalyse – Ergebnisse aus der Evaluation einer netzbasierten Weiterbildung mit Unterrichtsvideos. 4.

- Tagung der Sektion „Empirische Bildungsforschung“ der Deutschen Gesellschaft für Erziehungswissenschaft (DGfE). Wuppertal: 19.-21. März 2007.
- Krapp, M. (2004): Nach der Beschlussfassung in der Kultusministerkonferenz: Wie werden Standards wirksam? In: Fitzner, Th. (Hrsg.): Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform. Bad Boll: Evangelische Akademie, S. 310-321.
- Kubinger, K.D./Frebort, M./Holoher-Ertl, S./Pletschko, T. (2006): Standards-Tests zu den Bildungsstandards in Österreich. Wissenschaftlicher Hintergrund und Hinweise zur Interpretation der Ergebnisse der Standard-Tests. Wien: BMBWK. URL: <http://www.gemeinsamlernen.at/site/Verwaltung/mOBibliothek/Bibliothek/Standards%20Handbuch+Kubinger.pdf> (Stand 21. April 2007).
- Kuiper, W./Akker, J. v.d./Hooghoff, H./Letschert, J. (2005): Curriculum Policy and School Practice in a European Comparative Perspective. In: Letschert, J.: Curriculum Development Re-Invented. Enschede: SLO, S. 56-77.
- Kultusministerkonferenz (KMK) (2001): Erste Konsequenzen aus den Ergebnissen der PISA-Studie. Beschluss vom 5./6. Dezember 2001.
- Kultusministerkonferenz (KMK) (2002): Bildungsstandards zur Sicherung von Qualität und Innovation im föderalen Wettbewerb der Länder. Beschluss vom 23./24. Mai 2002.
- Kultusministerkonferenz (KMK) (2003): Vereinbarung über Bildungsstandards für den Mittleren Schulabschluss (Jahrgangsstufe 10). Beschluss vom 4. Dezember 2003.
- Kultusministerkonferenz (KMK) (2004a): Vereinbarung über Bildungsstandards für den Primarbereich (Jahrgangsstufe 4). Beschluss vom 15. Oktober 2004.
- Kultusministerkonferenz (KMK) (2004b): Standards für die Lehrerbildung: Bildungswissenschaften. Beschluss vom 16. Dezember 2004.
- Kultusministerkonferenz (KMK) (2005): Bildungsstandards der Kultusministerkonferenz. Erläuterungen zur Konzeption und Entwicklung. München und Neuwied: Luchterhand.
- Kultusministerkonferenz (KMK) (2006a): Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring (Entwurf).
- Kultusministerkonferenz (KMK) (2006b): Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. Beschluss vom 2. Juni 2006.
- Kunz, A./Halbheer, U./Maag Merki, K. (2005): „Problemlösekompetenz“ in Zürcher Gymnasien. Kooperations-Level als Indikator für eine Lernende Schule. Referat gehalten am Kongress der Schweizerischen Gesellschaft für Bildungsforschung – SSRE, September.
- Kuntze, S. (2004): Das binationale und videobasierte Lehrerinnen- und Lehrerfortbildungsprojekt „MuBiL“. In: GDM-Mitteilungen 79, 112-118.
- Kuntze, S. (2006): Video Technology in the Assessment of an In-Service Teacher Learning Program – Differences in Mathematics Teachers’ Judgements on Instructional quality. In: Zentralblatt für Didaktik der Mathematik 38, Nr. 5, S. 413-421.

- Künzli, R. (1986): Topik des Lehrplandenkens I. Architektur des Lehrplanes: Ordnung und Wandel. Kiel: Bärbel Mende Verlag Wissenschaft+Bildung (Kieler Beiträge zu Unterricht und Erziehung, hrsg. v. J. Petersen, Bd. 2).
- Kuper, H./Schneewind, J. (Hrsg.): Rückmeldung und Rezeption von Forschungsergebnissen. Zur Verwendung wissenschaftlichen Wissens im Bildungsbereich. Münster, New York, München, Berlin: Waxmann.
- La Marca, P.M./Redfield, D./Winter, P.C./Bailey, A./Despriet, L. (2000): State Standards and State Assessment Systems: A Guide to Alignment. Washington, DC: Council of Chief State School Officers.
- Labudde, P. (2004): Unser Unterricht im Spiegel von TIMSS und PISA – Resultate und Entwicklungsperspektiven. In R. Gschwend & A. Claude (Hrsg.): Unterrichtsentwicklung – zum Stand der Diskussion. Bern: EDK.
- Ladd, H.F. (2003): Comment on Caroline M. Hoxby: School Choice and School Competition: Evidence from the United States. In: Swedish Economic Policy Review 10, S. 67-76.
- Lam, T.C.M. (2004): Issues and Strategies in Standards-Based School Reform: The Canadian Experience. In: Fitzner, Th. (Hrsg.): Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform. Bad Boll: Evangelische Akademie, S. 103-149.
- Lampert, M. (1999): Knowing Teaching from the Inside Out: Implications of Inquiry in Practice for Teacher Education. In: Griffin, G.A. (Hrsg.): The Education of Teachers. Ninety-eighth Yearbook of the National Society for the Study of Education. Chicago, IL: The University of Chicago Press, S. 167-184.
- Landert, Ch./Stamm, M./Trachsler, E. (1998): Die Erprobungsfassung des Lehrplans für die Volksschule des Kantons Zürich. Bericht über die externe wissenschaftliche Evaluation im Auftrag der Bildungsdirektion des Kantons Zürich. Teil I: Synthese. Zürich: Bildungsdirektion des Kantons Zürich.
- Langen, B./Driscoll, P. (2004): Arbeiten mit Standards in England. In: Grundschule 36, Nr. 4 (Beilage), S. 17-20.
- Larcher Klee, S. (2005): Einstieg in den Lehrberuf. Untersuchungen der Identitätsentwicklung von Lehrerinnen und Lehrern im ersten Berufsjahr. Bern, Stuttgart, Wien: Haupt Verlag.
- Lauer, P.A./Snow, D./Martin-Glenn, M./Van Buhler, R.J./Stoutemyer, K./Snow-Renner, R. (2005): The Influence of Standards on K-12 Teaching and Student Learning. A Research Synthesis. Aurora, CO: Mid-continent Research for Education and Learning (McREL). URL: [http://www.edaccountability.org/capacity%20building%20FEAExecSum%20\(1\).pdf](http://www.edaccountability.org/capacity%20building%20FEAExecSum%20(1).pdf) (Stand 4. Mai 2007).
- Lee, J./Fitz, J. (1997): HMI and OFSTED: Evolution or Revolution in School Inspection? In: British Journal of Educational Studies 45, S. 39-52.
- Leithwood, K. (2000): Organizational Learning and School Improvement. Greenwich, CT: JAI.

- Leithwood, K./Menzies, T. (1998): Forms and Effects of School-Based Management: A Review. In: *Educational Policy* 12, S. 325-346.
- Leutner, D. (2006): Instruktionspsychologie. In: Rost, D.H. (Hrsg.): *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz, S. 261-270.
- Levin, H.M. (1974): A Conceptual Framework for Accountability in Education. In: *School Review* 82, S. 363-391.
- Lewis, L./Parsad, B./Carey, N./Bartfai, N./Farris, E./Smerdon, B./Green, B. (1999): *Teacher Quality: A Report on the Preparation and Qualification of Public School Teachers*. U.S. Department of Education, National Center for Education Statistics, Washington, DC: U.S. Government Printing Office.
- Lewis, C./Perry, R./Murata, A. (2006): How Should Research Contribute to Instructional Improvement? The Case of Lesson Study. In: *Educational Researcher* 35, Nr. 3, 3-14.
- Lieberman, A. (1995): Practices that Support Teacher Development: Transforming Conceptions of Professional Learning. In: *Phi Delta Kappan* 76, S. 591-596.
- Lieberman, A./McLaughlin, M.W. (1996): Networks for Educational Change: Powerful and Problematic. In: McLaughlin, M.W./Oberman, I. (Hrsg.): *Teacher Learning: New Policies, New Practices*. New York: Teachers College Press, S. 63-72.
- Liket, Th. (1995): *Freiheit und Verantwortung: Das niederländische Modell des Bildungswesens*. Gütersloh: Bertelsmann Stiftung.
- Liket, Th. (1997): Niederlande. In: Döbert, H./Geissler, G. (Hrsg.): *Schulautonomie in Europa*. Baden-Baden: Nomos, S. 247-264.
- Liket, Th. (1998): Niederlande. In: Bertelsmann Stiftung (Hrsg.): *Innovative Schulsysteme im internationalen Vergleich. Dokumentation zum Symposium und Festakt, Bd. 2*. Gütersloh: Verlag Bertelsmann Stiftung.
- Linn, R.L. (2000): Assessments and Accountability. In: *Educational Researcher* 29, Nr. 2, S. 4-16.
- Linn, R.L. (2003): Accountability: Responsibility and Reasonable Expectations. In: *Educational Researcher* 32, Nr. 7, S. 3-13.
- Linn, R.L./Baker, E.L./Betebenner, D.W. (2002): Accountability Systems: Implications of Requirements of the No Child Left Behind Act of 2001. In: *Educational Researcher* 31, Nr. 6, S. 3-16.
- Lipowsky, F. (2004): Was macht Fortbildungen für Lehrkräfte erfolgreich? In: *Die deutsche Schule* 96, Nr. 4, S. 462-479.
- Loeb, S./Strunk, K. (1995): *Accountability and Local Control: Response to Incentives With and Without Authority over Resource Generation and Allocation*. Unpubl. paper: Stanford University.
- Lorenz, J.H. (2005): Standards im Mathematikunterricht der Grundschule. In: *Grundschule* 36, Nr. 4 (Beilage), S. 11-14.
- Lortie, D.C. (1975): *Schoolteacher*. Chicago: University of Chicago Press.

- Loucks-Horsley, S. (1996): Professional Development for Science Education: A Critical and Immediate Challenge. In: Bybee, R. (Hrsg.): National Standards and the Science Curriculum of the Biological Sciences Curriculum Study. Dubuque, IA: Kendall/Hunt Publishing, S. 83-95.
- Loucks-Horsley, S./Hall, G. (1979): Implementing Innovations in Schools: A Concerns-Based Approach. Austin, TX: Research and Development Center for Teacher Education, University of Texas.
- Loucks-Horsley, S./Stiegelbauer, S. (1991): Using Knowledge of Change to Guide Staff Development. In: Miller, L./Lieberman, A. (Hrsg.): Staff Development for Education in the '90s: New Demands, New Realities, New Perspectives. New York: Teachers College Press, S. 15-36.
- Lucyshyn, J. (2006a): Implementation von Bildungsstandards in Österreich. Salzburg: Projektmanagement Bildungsstandards, BMBWK. URL: http://www.pi.salzburg.at/standards/LUCYSHYN_Implementation%20von%20Bildungsstandards%20in%20Oesterreich.pdf (Stand 22. Januar 2006).
- Lucyshyn, J. (2006b): Implementation von Bildungsstandards in Österreich. Arbeitsbericht. Salzburg: Bundesinstitut für Bildungsforschung, Innovation und Entwicklung des Bildungswesens (BIFIE). URL: <http://www.pi-wien.at/interplus/bildungsstandards/arbeitsbericht.pdf> (Stand 22. Januar 2006).
- Lundahl, L. (2002): From Centralisation to Decentralisation: Governance in Sweden. In: European Educational Research Journal 4, S. 625-636.
- Lusi, S. (1997): The Role of State Departments of Education in Complex School Reform. New York: Teachers College Press.
- Maag Merki, K./Büeler, X. (2002): Schule im Spannungsfeld von Demokratie und Markt. In: Holtappels, H.G./Klemm, K./Pfeiffer, H./Rolf, H.-G./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung, Bd. 12. Daten, Beispiele und Perspektiven. Weinheim: Juventa, S. 131-161.
- MacDonald, A. (1897/1898): Experimental Study of Children, Including Anthropometrical and Psycho-Physical Measurements of Washington School Children, and a Bibliography. In: Report of the Commissioner of Education. Washington, DC: National Printing Office, chapters XXI/XXII.
- Madaus, G.F. (1985): Public Policy and the Testing Profession – You've Never Had It so Good? In: Educational Measurement: Issues and Practices 4, Nr. 4, S. 5-11.
- Maehr, M.L./Maehr, J.M. (1996): Schools Aren't As Good As They Used to Be; They Never Were. In: Educational Researcher 25, Nr. 8, S. 21-24.
- Mannsfeld, K. (2004): Nationale Standards in einem föderalen System: Was tun die Länder? In: Fitzner, Th. (Hrsg.): Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform. Bad Boll: Evangelische Akademie, S. 296-309.
- Marlette, S.M./Goldston, J.M. (2003): Implementation of the Kansas Education Standards: A Principal/Teacher Perspective. Paper presented at the annual meeting of

- the Association for the Education of Teachers of Science. St. Louis, MO, January/February.
- Marsh, D.D./Odden, A.R. (1991): Implementation of the California Mathematics and Science Frameworks. In: Odden, A.R. (Hrsg.): Educational Policy Implementation. Albany: State University of New York Press, S. 219-240.
- Marsh, J.A. (2000): Connecting Districts to the Policy Dialogue: A Review of Literature on the Relationship of Districts with States, Schools, and Communities. Working paper. Seattle, WA: Center for the Study of Teaching and Policy.
- Marzano, R.J./Mayer, F./Dean, C.B. (2000): Research into Practice Series: Implementing Standards in the Classroom. Aurora, CO: Mid-continent Research for Education and Learning (McREL).
- Massell, D. (1998): State Strategies for Building Capacity in Education: Progress and Continuing Challenges. University of Pennsylvania, Graduate School of Education: Consortium for Policy Research in Education. URL: <http://www.cpre.org/Publications/rr41.pdf> (Stand 23. November 2005).
- Massell, D./Kirst, M./Hoppe, M. (1997): Persistence and Change: Standards-Based Reform in Nine States. University of Pennsylvania, Graduate School of Education: Consortium for Policy Research in Education.
- Mayer, D./Mullens, J./Moore, M. (2001): Monitoring School Quality. U.S. Department of Education, National Center for Education Statistics, Washington, DC: U.S. Government Printing Office.
- McClelland, D.C. (1973): Testing for Competence Rather Than for „Intelligence“. In: American Psychologist 28, S. 1-14.
- McDermott, K.A. (2000): Barriers to Large-Scale Success of Models for Urban School Reform. Educational Evaluation and Policy Analysis 22, S. 83-93.
- McDonnell, L.M./Elmore, R.F. (1987): Getting the Job Done: Alternative Policy Instruments. In: Educational Evaluation and Policy Analysis 9, S. 133-152.
- McLaughlin, M.W./Mitra, D. (2001): Theory-Based Change and Change-Based Theory: Going Deeper and Going Broader. In: Journal of Educational Change 2, S. 301-323.
- McLaughlin, M.W./Shepard, L.A. (1995): Improving Education through Standards-Based Reform: A Report of the National Academy of Education Panel on Standards-Based Educational Reform. Stanford, CA: National Academy of Education.
- McLaughlin, M.W./Talbert, J.E. (2001): Professional Communities and the Work of High School Teaching. Chicago: University of Chicago Press.
- McNeil, L. (2000): Contradictions of School Reform. Educational Costs of Standardized Testing. New York und London: Routledge.
- Medina, K./St. John, M. (1997): The Nature of Teacher Leadership: Lessons Learned from the California Subject Matter Projects. Inverness, CA: Inverness Research Associates.
- Mehrens, W.A./Kaminiski, J. (1989): Methods for Improving Standardized Test Scores: Fruitful, Fruitless, or Fraudulent? In: Educational Measurement: Issues and Practice 8, Nr. 1, S. 14-22.

- Meier, D.W./Wood, G.H. (Hrsg.) (2004): *Many Children Left Behind: How the No Child Left Behind Act is Damaging our Children and Our Schools*. Boston: Beacon Press.
- Messner, R. (2004): Was Bildung von Produktion unterscheidet. In: Schlömerkemper, J. (Hrsg.): *Bildung und Standards. Zur Kritik der „Instandardsetzung“ des deutschen Bildungswesens*. 8. Beiheft der Deutschen Schule. Weinheim: Juventa, S. 26-47.
- Messner, R. (2005): Die Neugestaltung der Schulinspektion im Spannungsfeld von Evaluation und Schulentwicklung. Referat gehalten anlässlich der Tagung ‚Schulinspektion – Schulrat mit neuer Gewalt?‘, 9.-11.12.2005, Schmitten. URL: http://www.evangelische-akademie.de/_old/materialien/054959/messner.pdf (Stand 22. April 2007).
- Messner, R. (2006): Bildungsstandards und Schulentwicklung – ein vernachlässigter Zusammenhang. In: *Seminar 12*, Nr 2, S. 21-36.
- Messner, H./Reusser, K. (2000): Berufliches Lernen als lebenslanger Prozess. In: *Beiträge zur Lehrerbildung* 18, Nr. 3, S. 277-294.
- Meyer, H. (2003). Zehn Merkmale guten Unterrichts. In: *Pädagogik* 10, 36-43.
- Michlin, M./Seppanen, P./Sheldon, T. (2001): *Charting a New Course: A Study of the Adoption and Implementation of Standards-Based Mathematics Curricula in Eight Minnesota School Districts*. Final Report to SciMath. Minnesota University, Minneapolis, Center for Applied Research and Educational Improvement, Fall.
- Mid-continent Research for Education and Learning (McREL) (2000): *Noteworthy Perspectives on Implementing Standards-Based Education*. November. URL: <http://www.mcrel.org/topics/productDetail.asp?productID=99> (Stand 4. Januar 2006).
- Mid-continent Research for Education and Learning (McREL) (2006): *Content Knowledge. A Compilation of Content Standards for K-12 Curriculum in Both Searchable and Browsable Formats*. 4th Ed.: *The process of This Work*. URL: <http://www.mcrel.org/standards-benchmarks/docs/process.asp> (Stand 13. Januar 2006).
- Miles, K.H./Darling-Hammond, L. (1998): Rethinking the Allocation of Teacher Resources: Some Lessons from High-Performing Schools. In: *Educational Evaluation and Policy Analysis* 20, S. 9-29.
- Miller, R. (Hrsg.) (1995): *Educational Freedom for a Democratic Society. A Critique of National Educational Goals, Standards, and Curriculum*. Brandon, VT: Psychology Press/Holistic Education Press.
- Ministère de l'Éducation Nationale, de la Formation Professionnelle et des Sports (MENFP) (2004): *PISA 2000. Kompetenzen von Schülern im internationalen Vergleich*. Nationaler Bericht Luxemburg. Luxembourg: MENFP.
- Ministerium für Bildung und Frauen des Landes Schleswig-Holstein (Hrsg.) (2005): *EVIT. Evaluation im Team*. Handbuch. URL: http://evit.lernnetz2.de/allgemein/materials/evit_handbuch18_10_2005.pdf (Stand 21. März 2006).

- Miron, G. (1996): Choice and the Quasi-Market in Swedish Education. In: Walford, G. (Hrsg.): Oxford Studies in Comparative Education 6, Nr. 1, S. 33-47.
- Moelands, H.A./Ouberg, M.J. (1998): School Self Evaluation in Primary Schools in the Netherlands. Paper presented at the European Conference of Educational research (ECER98), Ljubljana, Slovenia, September.
- Monk, D.H. (1994): Subject Area Preparation of Secondary Mathematics and Science Teachers and Student Achievement. In: Economics of Education Review 13, Nr. 2, S. 1-21.
- Morita, E. (2005): Lesson Study: Kooperative Lehrerweiterbildung in Japan. In: Beiträge zur Lehrerbildung 23, S. 398-409.
- Moser, U. (2003a): Leistungsbeurteilung der Schülerinnen und Schüler als Instrument zur Sicherung der Bildungsqualität. Expertise zuhanden der Stabsabteilung Entwicklung und Qualitätssicherung des Departements Bildung, Kultur und Sport des Kantons Aargau. Aarau: Departement Bildung, Kultur und Sport.
- Moser, U. (2003b): Klassenscockpit im Kanton Zürich. Ergebnisse einer Befragung von Lehrerinnen und Lehrern der 6. Klassen über ihre Erfahrungen im Rahmen der Erprobung von Klassenscockpit im Schuljahr 2002/2003. Bericht zuhanden der Bildungsdirektion des Kantons Zürich. Ms. Zürich.
- Moser, U. (2004): ‚Stellwerk‘ soll Jugendliche auf die richtige Schiene weisen. In: Moser, U./Keller, F. (2005): Erste Ergebnisse Check 5. Zwischenbericht zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau. Zürich: Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich.
- Moser, U./Tresch, S. (2005): Leistungen messen und beurteilen. Handreichung zum Umgang mit Ergebnissen von Leistungstests. Aarau: Departement Bildung, Kultur und Sport des Kantons Aargau.
- Müller, Ch. (2006): Volksschulentwicklung im Kanton Graubünden. Diss. phil. Universität Zürich, Pädagogisches Institut (Fachbereich Allgemeine Pädagogik). Ms. Zürich.
- Müller, K./Silver, R. (2006): Analyse de la littérature critique sur le développement, l'usage et l'implémentation de standards dans un système éducative: USA. Réalisé dans le cadre du projet Harnos de la CDIP sous la direction de M. Behrens. Im Auftrag der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren EDK. URL: http://www.edk.ch/PDF_Downloads/Harnos/Literaturanalyse_2.pdf (Stand 22. März 2006).
- Müller-Jung, A. (1928): Schulmäßiges Wissen und Können bei Schulentlassenen. Berlin: Union Deutsche Verlagsanstalt.
- Munby, H./Russell, T./Martin, A.K. (2001): Teachers' Knowledge and How It Develops. In: Richardson, V. (Hrsg.): Handbook of Research on Teaching. Washington, DC: American Educational Research Association, S. 877-904.

- Münder, N. (2007): Die Wirksamkeit von Parallelarbeiten als Instrument zur Qualitätssicherung und Qualitätsentwicklung aus der Sicht von Lehrerinnen und Lehrern. Eine empirische Studie an Münsteraner Grundschulen, Hauptschulen und Gymnasien. Diss. Universität Münster.
URL:http://miami.uni-muenster.de/servlets/DerivateServlet/Derivate-3751/diss_muender.pdf (Stand 23. April 2007).
- Murnane, R.J./Levy, F. (1996): *Teaching the New Basic Skills: Principles for Educating Children to Thrive in a Changing Economy*. New York: Free Press.
- Nachtigall, C./Kröhne, U. (2006): Methodische Anforderungen an schulische Leistungsmessung – auf dem Weg zu fairen Vergleichen. In: Kuper, H./Schneewind, J. (Hrsg.): *Rückmeldung und Rezeption von Forschungsergebnissen. Zur Verwendung wissenschaftlichen Wissens im Bildungsbereich*. Münster, New York, München, Berlin: Waxmann, S. 59-74.
- National Commission on Excellence in Education (1983): *A Nation at Risk: The Imperative for Educational Reform*. Washington, DC: U.S. Government Printing Office.
- National Commission on Teaching and America's Future (1996): *What Matters Most: Teaching for America's Future*. New York: NCTAF.
- National Council of Teachers of Mathematics (NCTM) (1989): *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (NCTM) (1991): *Professional Standards for Teaching Mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (NCTM) (1995): *Assessment Standards for School Mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (NCTM) (2000): *Principles and Standards for School Mathematics*. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (NCTM) (2006): *High-Stakes Tests. A Position of the National Council of the Teachers of Mathematics*. URL: http://www.nctm.org/about/position_statements/highstakes.htm (Stand 3. August 2006).
- National Council on Education Standards and Testing (NCEST) (1992): *Raising Standards for American Education*. Washington, DC: NCEST.
- National Curriculum (2004): *Foundation Stage, Years 1 and 2: Assessment and Reporting Arrangements*. London: DfES.
- National Education Association (NEA) (2004): *Teacher Quality: Moving Forward*. Washington, DC: NEA.
- National Education Goals Panel (NEGP) (1991): *The National Education Goals Report: Building a Nation of Learners*. Washington, DC: NEGP.
- National Education Goals Panel (NEGP) (1993): *Promises to Keep: Creating High Standards for American Students. Report on the Review of Educational Standards*

- from the Goals 3 and 4 Technical Planning Group to the National Education Goals Panel. Washington, DC: NEGP.
- National Governors' Association (NGA) (1989): *From Rhetoric to Action*. Washington, DC: NGA.
- National Research Council (NRC) (1999): *Testing, Teaching, and Learning. A Guide for States and School Districts*. Washington, DC: NRC. URL: <http://www.nap.edu/books/0309065348/html/R1.html> (16. November 2005).
- National Research Council (NCR) (2001b): *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press.
- National Union of Teachers (NUT) (2005): *A Good School for Every Child. The Response of the National Union of Teachers to the Government's White Paper: High Standards, Better Schools for All*. London: NUT.
- Natriello, G. (1996): *Diverting Attention from Conditions in American Schools*. In: *Educational Researcher* 25, Nr. 8, S. 7-9.
- Nelson, B.S./Hammerman, J.K. (1996): *Reconceptualizing Teaching: Moving toward the Creation of Intellectual Communities of Students, Teachers, and Teacher Educators*. In: McLaughlin, M.W./Oberman, I. (Hrsg.) (1996): *Teacher Learning: New Policies, New Practices*. New York: Teachers College Press, S. 3-21.
- Nessel, I. (2005): *Zwei Wege. Brandenburg und Rheinland-Pfalz*. In: Becker, G./Bremmerich-Vos, A./Demmer, M./Maag Merki, K./Proebe, B./Schwippert, K./Stäudel, L./Tillmann, K.-J. (Hrsg.): *Standards. Unterrichten zwischen Kompetenzen, zentralen Prüfungen und Vergleichsarbeiten. Friedrich Jahresheft XXIII*, S. 22-25.
- Neuweg, G.H. (2004): *Bildungsstandards in Österreich. Über die gute Absicht, die Vereinbarkeit von Einsicht und Aufsicht und die gebotene Vorsicht*. In: *Pädaktuell*, Nr. 2, S. 4-13. URL: http://www.wipaed.jku.at/mitarb/Neuweg/Neuweg_Standards_Paedaktuell.pdf (Stand 16. März 2006).
- Neveling, A. (2005): *Kollegiale Evaluation (KoEv) – Ein Verfahren zum Lernen am eigenen Modell*. In: *Seminar* 11, Nr. 4, S. 78-99.
- Newmann, F.M./King, M.B./Rigdon, M. (1997): *Accountability and School Performance: Implications from Restructuring Schools*. In: *Harvard Educational Review* 67, S. 41-74.
- Nicaise, I./Esping-Andersen, G./Pont, B./Turnstall, P. (2005): *Equity in Education. Thematic Review. Sweden Country Note*. Paris: OECD.
- Nickolaus, R./Gräsel, C. (2006) (Hrsg.): *Innovation und Transfer. Expertisen zur Transferforschung*. Hohengehren: Schneider.
- Nickolaus, R./Ziegler, B./Abel, M./Eccard, C./Aheimer, R. (2006): *Transferkonzepte, Transferprozesse und Transfereffekte ausgewählter Modell- und Schulversuchsprogramme*. In: Nickolaus, R./Gräsel, C. (Hrsg.): *Innovation und Transfer. Expertisen zur Transferforschung*. Hohengehren: Schneider, S. 24-71.

- Niggli, A. (2000): Lernarrangements erfolgreich planen. Didaktische Anregungen zur Gestaltung offener Arbeitsformen. Aarau: Sauerländer.
- Nisbet, J. (1990): Rapporteur's Report. In: Council of Europe/Scottish Council for Research in Education (Hrsg.): *The Evaluation of Educational Programmes: Methods, Uses, and Benefits*. Amsterdam: Swets & Zeitlinger, S. 1-9.
- O'Day, J.A. (1995): Systemic Reform in California. In: Goertz, M.E./Floden, R.E./O'Day, J.A. (Hrsg.): *Studies of Education Reform: Systemic Reform*. Bd. II: Case Studies. Newark, NJ: Rutgers, the State University of New Jersey, Center for Policy Research in Education, S. 1-38.
- O'Day, J.A. (2002): Complexity, Accountability, and School Improvement. In: *Harvard Educational Review* 72, S. 293-329. <http://gseweb.harvard.edu/hepg/oday.html> (Stand 28. Dezember 2005).
- O'Neil, H.F./Abedi, J./Miyoshi, J./Mastergeorge, A. (2004): Monetary Incentives for Low-Stakes Tests. In: Fitzner, Th. (Hrsg.): *Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform*. Bad Boll: Evangelische Akademie, S. 164-201.
- O'Reilly, F.E. (1996): *Educational Accountability: Current Practices and Theories in Use*. Cambridge, MA: Harvard University, Consortium for Policy Research in Education.
- O'Shea, M.R. (2003): *Implementing State Academic Standards in the Classroom*. Paper presented at the annual meeting of the American Association of Colleges for Teacher Education. New Orleans, LA, January.
- Odden, A.R. (Hrsg.) (1991): *Education Policy Implementation*. Albany, NY: The State University of New York Press.
- Odden, A.R./Marsh, D. (1989): State Education Reform Implementation: A Framework for Analysis. In: Hannaway, J./Crowson, R. (Hrsg.): *The Politics of Reforming School Administration*. Philadelphia, PA: Falmer, S. 41-59.
- OECD (1991): *Schulen und Qualität. Ein internationaler OECD-Bericht*. Frankfurt a.M.: Lang (Englisches Original 1989).
- OECD (1995): *Schools under Scrutiny*. Paris: OECD.
- OECD (2001): *Lernen für das Leben. Erste Ergebnisse der internationalen Schulleistungsstudie PISA 2000*. Paris: OECD.
- OECD (2005a): *Formative Assessment: Improving Learning in Secondary Classroom*. Paris: OECD.
- OECD (2005b): *Education at a Glance. OECD Indicators 2005*. Paris: OECD.
- OECD (2005c): *PISA 2003 Technical Report*. Paris: OECD.
- Oelkers, J. (2004): Zum Problem von Standards aus historischer Sicht. In: *Neue Sammlung* 44, S. 179-200.
- Ofsted (2004): *Strategic Plan 2005 to 2008*. London: Ofsted.
- Ofsted (2005): *William Tyndale Community Primary School Inspection Report July 2005*. London: Ofsted. URL: <http://www.ofsted.gov.uk/reports/> (Stand 31. Juli 2006).

- Olson, L. (1999): Shining a Spotlight on Results (Quality counts 1999). In: Education Week 18, Nr. 17. URL: <http://counts.edweek.org/sreports/qc99/ac/mc/mc-intro.htm> (Stand 15. Juli 2006).
- Onderwijsraad (2002): De Kern van het doel. Reactie op het advies van de commissie Wijnen over de kerndoelen basisonderwijs. Den Haag: Onderwijsraad.
- Orfield, G./Kornhaber, M.L. (Hrsg.) (2001): Raising Standards or Raising Barriers? Inequality and High-Stakes Testing in Public Education. New York: The Century Foundation Press.
- Ostermeier, Ch. (2004): Kooperative Qualitätsentwicklung in Schulnetzwerken. Münster: Waxmann.
- Ostermeier, Ch./Carstensen, C.H./Prenzel, M./Geiser, H. (2004): Kooperative unterrichtsbezogene Qualitätsentwicklung in Netzwerken. Ausgangsbedingungen für die Implementation im BLK-Modellversuchsprogramm SINUS. In: Unterrichtswissenschaft 32, S. 215-235.
- Pädagogische Hochschule Zürich (2004): Strategie 2005-2008. Zürich: PHZH. Vielfältiges Ms.
- Page, R. (1995): Who Systemizes the Systemizers? Policy and Practice in Interactions in a Case of State-Level Systemic Reform. In: Theory into Practice 34, S. 21-29.
- Pauli, C./Reusser, K. (2006): Von international vergleichenden Video Surveys zur videobasierten Unterrichtsforschung und -entwicklung. In: Zeitschrift für Pädagogik 52, S. 774-797.
- Peek, R./Dobbelstein, P. (2006): Benchmarks als Input für die Schulentwicklung. Das Beispiel der Lernstandserhebungen in Nordrhein-Westfalen. In: Kuper, H./Schneewind, J. (Hrsg.): Rückmeldung und Rezeption von Forschungsergebnissen. Münster: Waxmann, S. 41-58.
- Pennell, J.R./Firstone, W.A. (1996): Changing Classroom Practice through Teacher Networks: Matching Program Features with Teacher Characteristics and Circumstances. In: Teachers College Record 98, S. 46-76.
- Penuel, W.R./Kim, D.T./Michalchik, V./Lewis, S./Means, B./Murphy, R./Korbak, C./Whaley, A./Allen, J. E. (2002): Using Technology to Enhance Connections Between Home and School. A Research Synthesis. SRI International. URL: <http://ctl.sri.com/publications/> (Stand 1. Juni 2005).
- Peterson, G.J. (1999): Demonstrated Actions of Instructional Leaders: An Examination of Five California Superintendents. In: Education Policy Analysis Archives 7, Nr. 18. URL: <http://epaa.asu.edu/epaa/v7n18.html> (25. Juni 2006).
- Peterson, P. (Hrsg.) (2003): Our Schools and Our Future: Are We Still at Risk? Stanford, CA: Hoover Institution Press.
- Peterson, P./Rabe, B./Wong, K. (1986): When Federalism Works. Washington DC: The Brookings Institution.
- Petko, D. (2005): Das virtuelle Klassenbuch. Offener Unterricht und Elternmitarbeit mit Schülerinformationssystemen. In: Die Neue Schulpraxis 9, S. 56-59.

- Pfeiffer, H. (2004): Schule im Spannungsfeld von Demokratie und Markt. In: Holtapfels, H.G./Klemm, K./Pfeiffer, H./Rolff, H.-G./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung, Bd. 13. Daten, Beispiele und Perspektiven. Weinheim: Juventa, S. 51-81.
- PHZ Schwyz/PHZ Zug (2004): Professionelle Standards für die Aus- und Weiterbildung von Lehrpersonen der Vorschule und Primarstufe. Vervielfältigtes Ms.
- Pintrich, P.R. (2000): The Role of Goal Orientation in Self-Regulated Learning. In: Boekaerts, M./Pintrich, P.R./Zeidner, M.H. (Hrsg.): Handbook of Self-Regulation. San Diego: Academic Press, S. 451-502.
- PISA 2000 – Zentrale Handlungsfelder. Zusammenfassende Darstellung der laufenden und geplanten Massnahmen in den Ländern. Stand: 7. Oktober 2002. Beschluss der 299. Kultusministerkonferenz vom 17./18. Oktober 2002. URL: <http://www.kmk.org/schul/pisa/massnahmen.pdf> (Stand 13. März 2006).
- Plewis, I./Goldstein, H. (1998): The 1997 Education White Paper – A Failure of Standards. In: British Journal of Curriculum and Assessment 8, S. 17-20.
- Plowden, B. (1987): „Plowden“ Twenty Years On. In: Oxford Review of Education 13, S. 119-124.
- Polanyi, M. (1985): Implizites Wissen. Frankfurt a.M.: Suhrkamp.
- Porter, A.C. (1993): School Delivery Standards. In: Educational Researcher 22, Nr. 5, S. 24-30.
- Porter, A.C. (1994): National Standards and School Improvement in the 1990s: Issues and Promise. In: American Journal of Education 102, S. 421-449.
- Porter, A.C. (1995): The Uses and Misuses of Opportunity-to-Learn Standards. In: Educational Researcher 24, Nr. 1, S. 21-27.
- Porter, A.C./Floden, R./Freeman, D./Schmidt, W./Schwille, J. (1988): Content Determinants in Elementary School Mathematics. In: Grouws, D./Cooney, T. (Hrsg.): Perspectives on Research on Effective Mathematics Teaching. Reston, VA: NCTM, S. 96-113.
- Power, M. (1997): The Audit Society: Rituals of Verification. Oxford: Oxford University Press.
- Prenzel, M. (2000): Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts. Ein Modellversuchsprogramm von Bund und Ländern. In: Unterrichtswissenschaft 28, S. 103-126.
- Prenzel, M. (2005): Bildungsstandards und die Weiterentwicklung des Mathematikunterrichts. In: Henn, H.-W./Kaiser, G. (Hrsg.): Mathematikunterricht im Spannungsfeld von Evolution und Evaluation. Festschrift für Werner Blum. Hildesheim: Franzbecker, S. 261-271.
- Prenzel, M./Carstensen, C.H./Senkbeil, M./Ostermeier, Ch./Seidel, T. (2005): Wie schneiden SINUS-Schulen bei PISA ab? In: Zeitschrift für Erziehungswissenschaft 8, S. 540-562.

- Price, J./Ball, D./Luks, S. (1994): *Marshalling Resources for Reform: The Role of District Administrators*. Paper presented at the annual meeting of the American Educational Research Association. New Orleans, LA, April.
- Pull, J. (2002): *Luxemburg*. In: Döbert, H./Hörner, W./Kopp, B.v. (Hrsg.): *Die Schulsysteme Europas*. Hohengehren: Schneider, S. 282-295.
- Putnam, R./Borko, H. (1997): *Teacher Learning: Implications of New Views of Cognition*. In: Biddle, B.J./Good, T.L./Goodson, I.F. (Hrsg.): *The International Handbook of Teachers and Teaching*. Dordrecht: Kluwer, S. 1223-1296.
- Quinlan, M./Scharaschkin, A. (1999): *National Curriculum Testing: Problems and Practicalities*. Paper presented at 1999 BERA Conference, Brighton.
- Ramseier, E. (2007): *Ansprache des Preisträgers anlässlich der Übergabe des CORECHED-Preises durch Bundesrätin Doris Leuthard am 18. April 2007 in Bern*.
- Ratke, F.-O. (2006): *Das neue Erziehungsregime. Steuerungserwartungen, Kontrollphantasien und Rationalitätsmythen*. In: Frost, U. (Hrsg.) (2006): *Unternehmen Bildung. Die Frankfurter Einsprüche und kontroverse Positionen zur aktuellen Bildungsreform*. Vierteljahresschrift für Wissenschaftliche Pädagogik, Sonderheft. Paderborn: Schöningh, S. 45-49.
- Ratzka, N./Lipowsky, F./Krammer, K./Pauli, Ch. (2005): *Lernen mit Unterrichtsvideos. Ein Fortbildungskonzept zur Entwicklung von Unterrichtsqualität*. In: *Pädagogik* 57, Nr. 5, S. 30-33.
- Ravitch, D. (1995): *National Standards in American Education: A Citizen's Guide*. Washington, DC: Brookings Institution.
- Ravitch, D. (2002): *Testing and Accountability, Historically Considered*. In: Evers, M.W./Walberg, H.J. (Hrsg.): *School Accountability*. Stanford, CA: Hoover Press 2002, S. 9-21.
- Redding, S. (2000): *Parents and Learning*. UNESCO Educational Practices Series, 2. UNESCO. URL: <http://www.ibe.unesco.org/International/Publications/EducationalPractices/prachome.htm> (Stand 1. Juni 2005).
- Regeringens Skrivelse (1996/97) *Utvecklingsplan för Förskola, Skola och Vuxenutbildning: Kvalitet och likvärdighet*. Stockholm: Fritzes.
- Regeringens Skrivelse (1998/99): *Samverkan, ansvar och utveckling: Utvecklingsplan för Förskola, Skola och Vuxenutbildning*. Stockholm: Riksdagens tryckeriexpedition.
- Reinmann-Rothmeier, G./Mandl, H. (1998): *Wenn kreative Ansätze versanden: Implementation als verkannte Aufgabe*. In: *Unterrichtswissenschaft* 26, S. 292-311.
- Resnick, D.P. (1982): *History of Educational Testing*. In: Wigdor, A.K./Garner, W.R. (Hrsg.): *Ability Testing: Uses, Consequences, and Controversies. Part II*. Washington DC: National Academy Press, S. 173-194.
- Resnick, L.B./Hall, M.W. (2001): *The Principles of Learning: Study tools for educators [CD-ROM, version 2.0]*. University of Pittsburgh, Learning Research and Development Center, Institute for Learning.

- Resnick, L.B./Nolan, K. (1995): Standards for Education. In: Ravitch, D. (Hrsg.): *Debating the Future of American Education: Do We Need National Standards and Assessments?* Washington, DC: Brookings Institution, S. 94-119.
- Resnick, D.P./Resnick, L.B. (1983): Improving Educational Standards in American Schools. In: *Phi Delta Kappan* 65, S. 178-180.
- Resnick, D.P./Resnick, L.B. (1985): Standards, Curriculum, and Performance. A Historical and Comparative Perspective. In: *Educational Researcher* 14, Nr. 4, S. 5-20.
- Reusser, K. (2001): Unterricht zwischen Wissensvermittlung und Lernen lernen. Alte Sackgassen und neue Wege in der Bearbeitung eines pädagogischen Jahrhundertproblems. In: Finkbeiner, C./Schnaitmann, G.W. (Hrsg.): *Lehren und Lernen im Kontext empirischer Forschung und Fachdidaktik*. Donauwörth: Auer, S. 106-140.
- Reusser, K. (2005a): Problemorientiertes Lernen – Tiefenstruktur, Gestaltungsformen, Wirkung. In: *Beiträge zur Lehrerbildung* 23, S. 159-182.
- Reusser, K. (2005b): Situiertes Lernen mit Unterrichtsvideos in der Lehrerinnen- und Lehrerbildung. In: *Journal für Lehrerinnen- und Lehrerbildung* 5, Nr. 2, S. 8-18.
- Reusser, K./Pauli, C. (Hrsg.) (2003): *Mathematikunterricht in der Schweiz und in weiteren sechs Ländern. Bericht über die Ergebnisse einer internationalen und schweizerischen Video-Unterrichtsstudie. Doppel-CD-ROM (Schlussbericht mit Videodokumentation)*. Universität Zürich: Pädagogisches Institut.
- Rhyn, H. (1997): Länderbericht Schweiz. In: Österreichisches Bundesministerium für Unterricht und kulturelle Angelegenheiten (Hrsg.): *Schulleitung und Schulaufsicht. Neue Rollen und Aufgaben im Schulwesen einer dynamischen und offenen Gesellschaft*. Innsbruck, Wien: Studienverlag, S. 163-187.
- Richards, C. (2001): *School Inspection in England: A Re-Appraisal (Impact Paper No. 9)*. London: Institute of Education.
- Richardson, V./Placier, P. (2001): Teacher Change. In: Richardson, V. (Hrsg.): *Handbook of Research on Teaching*. Washington, DC: American Educational Research Association.
- Riley, K. (1998): *Whose School is It Anyway?* London: Falmer.
- Rogers, E.M. (2003): *Diffusion of Innovations*. New York: The Free Press.
- Rolff, H.-G. (1993): *Wandel durch Selbstorganisation: Theoretische und praktische Hinweise für eine bessere Schule*. Weinheim: Juventa.
- Rolff, H.-G. (1995): Autonomie als Gestaltungs-Aufgabe. Organisationspädagogische Perspektiven. In: Daschner, P./Rolff, H.-G./Stryck, T. (Hrsg.): *Schulautonomie – Chancen und Grenzen*. Weinheim: Juventa, S. 31-54.
- Rolff, H.-G. (1998): Schulentwicklung als Entwicklung von Einzelschulen: viel Praxis, wenig Theorie und kaum Forschung. Ein Versuch Schulentwicklung zu systematisieren. In: Rolff, H.-G./Bauer, K.-O./Klemm, K./Pfeiffer, H. (Hrsg.): *Jahrbuch der Schulentwicklung*, Bd. 10. Weinheim und München: Beltz, S. 295-326.

- Rolff, H.-G. (2001): Professionelle Lerngemeinschaften. Eine wirkungsvolle Synthese von Unterrichts- und Personalentwicklung. In: Buchen, H./Horster, L./Rolff, H.-G. (Hrsg.): Schulleitung und Schulentwicklung. Berlin: Raabe, S. 1-14.
- Rolff, H.-G. (2002): Rückmeldung und Nutzung der Ergebnisse in großflächigen Leistungsuntersuchungen. Grenzen und Chancen. In: Rolff, H.-G./Holtappels, H.G./Klemm, K./Pfeiffer, H./Schulz-Zander, R. (Hrsg.): Jahrbuch der Schulentwicklung, Bd. 12. Weinheim, München: Juventa, S. 75-98.
- Rolff, H.-G. (2006): Was wissen wir über die Entwicklung von Schule? In: Pädagogik 58, Nr. 6, S. 42-47.
- Rolff, H.-G. (2007): Kontrovers. In: Pädagogik 1, S. 49.
- Rolff, H.-G./Bastian, J. (2002): Abschlussevaluation des Projektes „Schule & Co.“ Gütersloh: Bertelsmann Stiftung.
- Rosenshine, B. (2003): High-Stakes Testing: Another Analysis. In: Education Policy Analysis Archives 11, Nr. 24. URL: <http://epaa.asu.edu/epaa/v11n24/> (Stand 23. Mai 2006).
- Rost, D. (2006): Kompetenzstrukturen und Kompetenzmessung. In: Pädagogik der Naturwissenschaften Chemie 55, Nr. 8, S. 5-8.
- Rüdel, E. (2005): Standards at Work. Beauchamp College in Leicestershire. In: Becker, G./Bremerich-Vos, A./Demmer, M./Maag Merki, K./Proebe, B./Schwippert, K./Stäudel, L./Tillmann, K.-J. (Hrsg.): Standards. Unterrichten zwischen Kompetenzen, zentralen Prüfungen und Vergleichsarbeiten. Friedrich Jahresheft XXIII, S. 54-56.
- Rudolph, J.L. (2002): Scientists in the Classroom. The Cold War Reconstruction of American Science Education. New York, Houndmills, Basingstoke: Palgrave.
- Ruf, U./Gallin, P. (1991): Aufbau von Sprach- und Fachkompetenz beim Lernen mit Kernideen und Reisetagebüchern. In: Schweizer Schule 78, Nr. 9, S. 18-29.
- Ruiz-Primo, M.A. (2006): A Multi-Method and Multi-Source Approach for Studying Fidelity of Implementation. National Center for Research on Evaluation, Standards, and Student Testing (CRESST): Center for the Study of Evaluation (CSE), Graduate School of Education & Information Studies. University of California, Los Angeles.
- Rumpf, H. (2005): Bildungsstandards? Einwände gegen die absehbare Verödung des Lernens. In: Diskussion Musikpädagogik 27, S. 4-7.
- Rupp, A.A./Leucht, M./Hartung, R. (2006): „Die Kompetenzbrille aufsetzen“. Verfahren zur multiplen Klassifikation für Kompetenzdiagnostik in Unterricht und Testung. In: Unterrichtswissenschaft 34, S. 195-219.
- Rutter, M. (1983): School Effects on Pupil Progress: Research Findings and Policy Implications. In: Shulman, L./Sykes, G. (Hrsg.): Handbook of Teaching and Policy. New York: Longbean, S. 3-41.

- Rutter, M./Maughan, B./Mortimore, P./Ouston, J. (1979): *Fifteen Thousand Hours. Secondary Schools and their Effects on Children*. Cambridge: Harvard University Press.
- Rychen, D.D./Salganik, L.H. (Hrsg.) (2003): *Key Competencies for a Successful Life and a Well-Functioning Society*. Göttingen: Hogrefe & Huber.
- Sabatier, P.A. (1986): Top-Down and Bottom-Up Approaches to Implementation. Research. In: *Journal of Public Policy* 6, S. 21-48.
- Sabers, D.L./Sabers, D.S. (1996): Conceptualizing, Measuring, and Implementing Higher (High or Hire) Standards. In: *Educational Researcher* 25, Nr. 8, S. 19-21.
- Sachs Adams, G./Torgerson, Th.L. (1956): *Measurement and Evaluation for the Secondary-School Teachers: With Implications for Corrective Procedures*. New York: Dryden Press.
- Sacks, P. (1999): *Standardized Mind. The High Prize of America's Testing Culture and What Can We Do to Change It*. Cambridge, MA: Perseus Publishing.
- Saldern, M.v./Paulsen, A. (2004): Sind Bildungsstandards die richtige Antwort auf PISA? In: Schlömerkemper, J. (Hrsg.): *Bildung und Standards. Zur Kritik der „Instandardsetzung“ des deutschen Bildungswesens*. 8. Beiheft der Deutschen Schule. Weinheim: Juventa, S. 66-100.
- Sammons, P./Elliot, K./Welcomme, W./Taggart, B./Levacic, R. (2004): England. In: Döbert, H./Klieme, E./Sroka, W. (Hrsg.): *Conditions of School Performance in Seven Countries. A Quest for Understanding the International Variation of PISA Results*. Münster: Waxmann, S. 65-149.
- Sanders, W.L./Horn, S.P. (1995): Educational Assessment Reassessed: The Usefulness of Standardized and Alternative Measures of Student Achievement as Indicators for the Assessment of Educational Outcomes. In: *Education Policy Analysis Archives* 3, Nr. 6. URL: <http://epaa.asu.edu/epaa/v3n6/> (Stand 12. Juli 2006).
- Sanders, W.L./Horn, S.P. (1998): Research Findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implications for Educational Evaluation and Research. In: *Journal of Personnel Evaluation in Education* 12, S. 247-256.
- Sanders, W.L./Rivers, J.C. (1996): *Cumulative and Residual Effects of Teachers on Future Student Achievement. Research Progress Report*. Knoxville, TN: University of Tennessee, Value-Added Research and Assessment Center.
- Sanders, W.L./Topping, K.J. (1999): *Teacher Effectiveness and Computer Assessment of Reading: Relating Value Added and Learning Information System Data*. Knoxville, TN: University of Tennessee Value Added Research and Assessment Center.
- Sarason, S. (1990): *The Predictable Failure of Educational Reform. Can We Change Course Before It's Too Late?* San Francisco, CA: Jossey-Bass.
- Scardamalia, M./Bereiter, C. (1994): Computer Support for Knowledge-Building Communities. In: *The Journal of the Learning Sciences* 3, S. 265-283.
- Scheerens, J./Bosker, R. (1997): *The Foundations of Educational Effectiveness*. New York: Pergamon.

- Scheerens, J./Glas, C./Letschert, J. (2005): Experteninterview mit Jaap Scheerens, Cees Glas und Jos Letschert. Durchgeführt am 27. Juni 2005 an der Universität Twente.
- Scheerens, J./Glas, C./Thomas, S.M. (2003): Educational Evaluation, Assessment, and Monitoring. A Systematic Approach. Lisse: Swets & Zeitlinger.
- Schley, W./Schratz, M. (2005): Leadership Academy – Chancen und Perspektiven für die Entwicklung von Schule und Lernen. In: *Erziehung und Unterricht* 155, S. 950-956.
- Schmoker, M. (2004): Tipping Point: From Feckless Reform to Substantive Instructional Improvement. In: *Phi Delta Kappan* 85, S. 424-432.
- Schmoker, M. (2006): Results Now: How We Can Achieve Unprecedented Improvements in Teaching and Learning? Alexandria, VA: Association for Supervision & Curriculum Development.
- Schneewind, J. (2006): Gutachten zu Form und Einsatz von Ergebnisrückmeldungen für die Unterrichtsentwicklung. Die Instrumente Klassenscockpit, Orientierungsarbeiten, Check 5. Im Auftrag der Bildungsdirektion des Kantons Zürich. Berlin: Freie Universität Berlin.
- Schön, D.A. (1987): Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Profession. San Francisco: Jossey-Bass.
- Schrader, F./Helmke, A. (2002): Alltägliche Leistungsbeurteilung durch Lehrer. In: Weinert, F.E. (Hrsg.): *Leistungsmessungen in Schulen*. 2. Aufl. Weinheim: Beltz, S. 45-58.
- Schrader, F.-W./Helmke, A. (2003): Evaluation – und was danach? Ergebnisse der Schulleiterbefragung im Rahmen der Rezeptionsstudie WALZER. In: *Schweizerische Zeitschrift für Bildungswissenschaften* 25, S. 79-110.
- Schratz, M. (1997): Länderbericht Österreich. In: Österreichisches Bundesministerium für Unterricht und kulturelle Angelegenheiten (Hrsg.): *Schulleitung und Schulaufsicht. Neue Rollen und Aufgaben im Schulwesen einer dynamischen und offenen Gesellschaft*. Innsbruck, Wien: Studienverlag, S. 189-235.
- Schratz, M. (1996): Die neue Qualität von Schulleitung. Schule als lernende Organisation. In: Specht, W./Thonhauser, J. (Hrsg.): *Handbuch zur Schulentwicklung*. Innsbruck: Studien-Verlag, S. 160-189.
- Schratz, M. (1996): Die Rolle der Schulaufsicht in der autonomen Schulentwicklung. Innsbruck, Wien: Studienverlag.
- Schratz, M./Bauk-van Vugt, A. (2000): Qualitätsstandards für Schulprogramme. URL: <http://bebis.cidsnet.de/fortbildung/unterstuetzung/forum/> (Stand 12. März 2006).
- Schwarz, T. (1997): Schulsystem und Lehrerausbildung in den Niederlanden. In: Kreienbaum, M.A./Meyer, K./S. Rathmann/Späth, A./Streb, B./Susenberg, P./Eckert, E. (Hrsg.): *Bildungslandschaft Europa*. Bielefeld: Kleine, S. 25-31.
- Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) (2004): HARMOS: Zielsetzungen und Konzeption. Juni 2004. Bern: EDK. URL: http://www.edk.ch/PDF_Downloads/Harmos/Harmos_Weissbuch_d.pdf#search

- =%22edk%20harmos%20zielsetzung%20konzeption%20juni%202004%22 (Stand 2. Februar 2006).
- Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) (2006): Interkantonale Vereinbarung über die Harmonisierung der obligatorischen Schule. HarmoS-Konkordat. Bericht zur Vernehmlassung 16.2.2006-30.11.2006. URL: http://www.edk.ch/PDF_Downloads/Vernehmlassungen/Harmos/Bericht_d.pdf (Stand 18. März 2006).
- Scruton, R. (1987): Expressionist Education. In: *Oxford Review of Education* 13, S. 39-44.
- Seago, N. (2004): Using Video as an Object of Inquiry for Mathematics Teaching and Learning. In: Brophy, J. (Hrsg.): *Using video in teacher education*. Amsterdam: Elsevier, S. 259-286.
- Seidel, T./Prenzel, M./Rimmele, R./Meyer, L./Dalehefte, I.M. (2004): *Lernprogramm LUV. Lernen aus Unterrichtsvideos für Physiklehrkräfte*. Kiel: IPN.
- Selbständige Schule (2007). URL: http://www.selbststaendige-schule.nrw.de/S__FaAie8GX (Stand 19 April 2007).
- Senge, P. (1996): *Die fünfte Disziplin*. Stuttgart: Klett-Cotta.
- Senkbeil, M./Drechsel, B./Rolff, H.-G./Bonsen, M./Zimmer, K./Lehmann, R.H./Neumann, A. (2004): Merkmale und Wahrnehmungen von Schule und Unterricht. In: Deutsches PISA-Konsortium (Hrsg.): *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs*. Münster: Waxmann, S. 296-314.
- Sexton, St. (1987): *Our Schools: A Radical Policy*. Warlington: Institute for Economic Affairs Education Unit.
- Shavelson, R./Baxter, G./Pine, J. (1992): Performance Assessments. Political Rhetoric and Measurement Reality. In: *Educational Researcher* 21, Nr. 4, S. 22-27.
- Shavit, Y./Blossfeld, H.P. (1993): *Persistent Inequality*. Boulder, CO: Westview Press.
- Shepard, L.A. (1990): Inflated Test Score Gains: Is the Problem Old Norms or Teaching to the Test? In: *Educational Measurement: Issues and Practice* 9, Nr. 3, S. 15-22.
- Shepard, L.A. (2000): The Role of Assessment in a Learning Culture. In: *Educational Researcher* 29, Nr. 7, S. 4-14.
- Sherin, M.G./Han, S.Y. (2004): Teacher Learning in the Context of a Video Club. In: *Teaching and Teacher Education* 20, S. 163-183.
- Sherin, M. G./Linsenmeier, K.A./van Es, E.A. (2006): *Selecting Video Clips to Promote Mathematics Teachers' Discussion of Student Thinking*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. Verfügbar unter: <http://www.professional-vision.org/> (Stand 15. Mai 2007).

- Shields, P.M./Marsh, J.A./Marder, C./Wilson, Ch.L./SRI International (1998): A Case Study of California's SSI (CAMS), 1992-1997. National Science Foundation. URL: <http://www.sri.com/policy/cep/pubs/ssi/casip.pdf> (Stand 24. November 2005).
- Shulman, L. (1986): Those Who Understand: Knowledge Growth in Teaching. In: *Educational Researcher* 15, Nr. 2, S. 4-14.
- Shulman, L. (1987): Knowledge and teaching: Foundations of the New Reform. In: *Harvard Educational Review* 57, Nr. 1, S. 1-21.
- Sill, H.-D. (2004): Bemerkungen zu den aktuellen Bildungsstandards. In: *Mitteilungen der GDM*, Nr. 78, S. 72-75.
- Simons, H. (1981): *The Sciences of the Artificial*. Cambridge, MA: The MIT Press.
- Sipple, J.W./Killeen, K. (2004): Context, Capacity, and Concern: A District-Level Analysis of the Implementation of Standards-Based Reform in New York State. In: *Education Policy* 18, S. 456-490.
- Sirotnik, K. (2004): *Holding Accountability Accountable*. New York: Teachers College Press.
- Skolverket (1998a): *Vem tror på skolan? Attityder till skolan 1997 (Rapport Nr. 144)*. Stockholm: Skolverket.
- Skolverket (1998b): *A Summary: Who Believes in our Schools? Attitudes to the Swedish School 1997*. Stockholm: Skolverket.
- Skolverket (1998c): *Tillsynen av skolan: en skrift framtagen som stöd för kommuneras egentillsyn (Allmänna råd 1998:1)*. Stockholm: Skolverket.
- Slenning, K. (2004): *Managing Educational Standards in a Decentralised Organisation: Critical Points of Transition*. In: Fitzner, Th. (Hrsg.): *Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform*. Bad Boll: Evangelische Akademie, S. 54-61.
- Smith, M.L./Rottenberg, C. (1991): Unintended Consequences of External Testing in Elementary Schools. In: *Educational Measurement: Issues and Practice* 10, Nr. 4, S. 7-11.
- Smith, M.S./O'Day, J.A. (1991): Systemic School Reform. In: Fuhrman, S.H./Malen, B. (Hrsg.): *The Politics of Curriculum and Testing*. Philadelphia, PA: Falmer, S. 233-268.
- Smith, M.S./O'Day, J.A. (1993): Systemic Reform and Educational Opportunity. In: Fuhrman, S.H. (Hrsg.): *Designing Coherent Education Policy*. San Francisco, CA: Jossey-Bass, S. 250-312.
- Smitt, M. (2002): Continuous Professional Development for Teachers. In: Gassner, O. (Hrsg.): *Strategies of Change in Teacher Education. European Views*. Feldkirch: European Network on Teacher Education Policies (ENTEP), S. 175-179.
- Smylie, M.A. (1995): Teacher Learning in the Workplace: Implications for School Reform. In: Guskey, T.R./Huberman, M. (Hrsg.): *Professional Development in Education*. New York: Teachers College Press, S. 92-113.

- Smylie, M.A. (1996): From Bureaucratic Control to Building Human Capital: The Importance of Teacher Learning in Education Reform. In: *Educational Researcher* 25, Nr. 9, S. 9-11.
- Söderström, M./Uusitalo, R. (2005): Vad innebar införandet av fritt skolval i Stockholm för segregeringen i skolan? Uppsala: Institutet för Arbetsmarknadspolitisk Utvärdering (IFAU).
- Söderström, M./Uusitalo, R. (2005a): School Choice and Segregation: Evidence from an Admission reform. Uppsala: Institutet för Arbetsmarknadspolitisk Utvärdering (IFAU).
- Solzbacher, C./Minderop, D. (Hrsg.) (2007): Bildungsnetzwerke und Regionale Bildungslandschaften. Ziele und Konzepte, Aufgaben und Prozesse. München: LinkLuchterhand.
- Sophian, S. (1997): Beyond Competence: The Significance of Performance for Conceptual Development. In: *Cognitive Development* 12, S. 281-303.
- Sosniak, L.A. (1999): Professional and Subject Matter Knowledge for Teacher Education. In: Griffin, G.A. (Hrsg.): *The Education of Teachers. Ninety-eighth Yearbook of the National Society for the Study of Education*. Chicago, IL: The University of Chicago Press, S. 185-204.
- Sparks, D./Hirsh, S. (1997): *A New Vision for Staff Development*. Alexandria, VA: ASCD/National Staff Development Council.
- Specht, W. (2003): Zum Stand der Entwicklung nationaler Bildungsstandards: Prinzipien – Bewertungen – Perspektiven. Ms. Graz.
- Specht, W. (2006): Von den Mühen der Ebene. Entwicklung und Implementation von Bildungsstandards in Österreich. Ms.
- Specht, W./Freudenthaler, H. (2004): Bildungsstandards. Bedingungen ihrer Wirksamkeit. In: *Erziehung und Unterricht* 154, S. 618-629.
- Spencer, D.A. (1996): Teachers and Educational Reform. In: *Educational Researcher* 25, Nr. 9, S. 15-17, 40.
- Spillane, J.P. (1996): School Districts Matter: Local Educational Authorities and State Instructional Policy. In: *Educational Policy* 10, S. 63-87.
- Spillane, J.P. (1999): External Reform Initiatives and Teachers' Efforts to Reconstruct Practice: The Mediating Role of Teachers' Zones of Enactment. In: *Journal of Curriculum Studies* 31, S. 143-175.
- Spillane, J.P. (2002a): Local Theories of Teacher Change: The Pedagogy of District Policies and Programs. In: *Teachers College Records* 104, S. 377-420.
- Spillane, J.P. (2002b): District Policymaking and State Standards: A Cognitive Perspective on Implementation. In: Hightower, A./Knapp, M.S./Marsh, J.A./McLaughlin, M. (Hrsg.): *School Districts and Instructional Renewal*. New York: Teachers College Press.
- Spillane, J.P. (2004): *Standards Deviation. How Schools Misunderstand Education Policy*. Cambridge, MA: Harvard University Press.

- Spillane, J.P./Reiser, B.J./Reimer, T. (2002): Policy Implementation and Cognition: Reframing and Refocusing Implementation Research. In: *Review of Education Research* 72, S. 387-431.
- Spillane, J.P./Thompson, C. (1997): Reconstructing Conceptions of Local Capacity: The Local Education Agency's Capacity for Ambitious Instructional Reform. In: *Education Evaluation and Policy Analysis* 19, S. 185-203.
- Spillane, J.P./Thompson, C.L./Lubienski, C./Jita, L./Reimann, C. (1995): The Local Government Policy System Affecting Mathematics and Science Education in Michigan: Lessons from Nine School Districts. East Lansing: Michigan State University College of Education/MSSI Policy Program Review Component.
- Spillane, J.P./Zeuli, J.S. (1999): Reform and Teaching: Exploring Patterns of Practice in the Context of National and State Mathematics Reforms. In: *Educational Evaluation and Policy Analysis* 21, S. 1-27.
- Spinner, K. (2005): Der standardisierte Schüler. In: *Didaktik Deutsch* 18, Nr. 4, S. 4-13.
- Spring, J. (1989): *The Sorting Machine Revisited: American Educational Policy Since 1945*. New York: Longman.
- St. John, M./Dickey, K./Heenan, B./Hirabayashi, J./Medina, K./Ramage, K. (1995): *Evaluating a Statewide Professional Development System*. Inverness, CA: Inverness Research Associates.
- Staatsinstitut für Schulqualität und Bildungsforschung (ISB) (2005): *KMK-Bildungsstandards. Konsequenzen für die Arbeit an bayerischen Schulen*. München: ISB.
- Stander, K. (2001): *Grant-Maintained Schools: A Study of Two Schools' Finance Management and Spending Patterns*. Diss. University of Oklahoma. URL: <http://e-archiv.library.okstate.edu/dissertations/AAI3021649/> (Stand 20. August 2006).
- Stanfield, J./Tooley, J. (2005): *Market Solutions for British Education*. In: Roy, S./Clarke, J. (Hrsg.): *Margaret Thatcher's Revolution: How It Happened and What It Meant*. London: Institute of Economic Affairs.
- Statistics of Education (2004): *Schools in England 2004 Edition*. London: TSO.
- Staub, F.C. (2001): *Fachspezifisch-Pädagogisches Coaching: Förderung von Unterrichtsexpertise durch Unterrichtsentwicklung*. In: *Beiträge zur Lehrerbildung* 19, S. 175-198.
- Staub, F.C. (2004): *Fachspezifisch-Pädagogisches Coaching: Ein Beispiel zur Entwicklung von Lehrerfortbildung und Unterrichtskompetenz als Kooperation*. In: Lenzen, D./Baumert, J./Watermann, R./Trautwein, U. (Hrsg.): *PISA und die Konsequenzen für die erziehungswissenschaftliche Forschung*. 3. Beiheft der Zeitschrift für Erziehungswissenschaft. Wiesbaden: VS Verlag Sozialwissenschaften, S. 113-142.
- Staub, F.C. (2005): *Videos im Fachspezifisch-Pädagogischen Coaching*. In: *Journal für Lehrerinnen- und Lehrerbildung* 5, Nr. 2, S. 26-30.
- Staub, F.C. (2006): *Wenn der Coach kommt... Diagnose und Unterrichtskompetenz stärken durch neue Beratungsformen*. In: Becker, G./Horstkemper, M./Risse,

- E./Stäudel, L./Werning, R./Winter, F. (Hrsg.): Diagnostizieren und Fördern. Stärken Entdecken – Können entwickeln. Friedrich Jahresheft XXIV, S. 138-140.
- Staub, F.C./Stern, E. (2002): The Nature of Teachers' Pedagogical Content Beliefs Matters for Students' Achievement Gains: Quasi-Experimental Evidence from Elementary Mathematics. In: *Journal of Educational Psychology* 94, S. 344-355.
- Stäudel, L./Blum, W. (2005): Prozessqualität entwickeln. Impulse für Fachkollegien. In: Becker, G./Bremerich-Vos, A./Demmer, M./Maag Merki, K./Proebe, B./Schwipert, K./Stäudel, L./Tillmann, K.-J. (Hrsg.): Standards. Unterricht zwischen Kompetenzen, zentralen Prüfungen und Vergleichsarbeiten. Friedrich Jahresheft XXIII, S. 115-117.
- Stecher, B.M./Hamilton, L.S. (2002): Putting Theory to the Test: Systems of „Educational Accountability Should Be Held Accountable“ In: *Rand Review* 26, Nr. 1, S. 16-23.
- Steffens, B./Bargel, T. (1993): Erkundungen zur Qualität von Schule. Neuwied: Luchterhand.
- Steinert, B./Klieme, E. (2003): Levels of Teacher Cooperation as Levels of School Development: A Criterion Referenced-Approach to School Evaluation. Paper presented at the European Conference on Educational Research, Hamburg, September.
- Steinert, B./Klieme, E./Maag Merki, K./Döbrich, P./Halbheer, U./Kunz, A. (2006): Lehrerkooperation in der Schule: Konzeption, Erfassung, Ergebnisse. In: *Zeitschrift für Pädagogik* 52, S. 185-204.
- Stokes, L.M./Sato, N.E./McLaughlin, M.W./Talbert, J.E. (1997): Theory-Based Reform and Problems of Change: Contexts that Matter for Teachers' Learning and Community. Stanford, CA: Stanford University.
- Strittmatter, A. (2000): Worauf bei der Selbstevaluation zu achten ist. 10 Lehren aus 5 Jahren FQS. Ein Forschungsbericht. In: Buchen, H./Horster, L./Rolff, H.-G. (Hrsg.): Schulleitung und Schulentwicklung. Berlin: Raabe, E 4.2, S. 1-24.
- Strittmatter, A. (2004): Die Standarddiskussion in der Schweiz. In: *Journal für Schulentwicklung* 8, Nr. 4, S. 39-46.
- Sunderman, G.L. (1999): Designing Title I Schoolwide Programs: A Comparison of Three Urban Districts. Johns Hopkins University, Baltimore, MD, Center for Social Organization of Schools.
- Swedish Ministry of Education and Science (1996): Information on the 1994 Curriculum for Compulsory School System (Lpo 94). Stockholm: Swedish Ministry of Education and Science.
- Swedish Ministry of Education and Science (1997): Summary of the National Development Plan for Pre-School, School, and Adult Education: Quality and Equivalence. Stockholm: Swedish Ministry of Education and Science.
- Tate, W.F. (2004): What is a Standard. In: Lester, F.K./Ferrini-Mundy, J. (Hrsg.): Proceedings on the NCTM Research Catalyst Conference. Reston, VA: NCTM, S. 15-

23. URL: <http://www.nctm.org/highered/sirg/sirg2.pdf>
(Stand 14. November 2005).
- Taylor, G./Shepard, L./Kinner, F./Rosenthal J. (2001): A Survey of Teachers' Perspectives on High-Stakes Testing in Colorado: What Gets Taught, What Gets Lost. Boulder, CO: University of Colorado, School of Education and the Public Interest Center.
- Teddlie, C./Reynolds, D. (2000): The International Handbook of School Effectiveness Research. London: Falmer.
- Terhart, E. (Hrsg.) (2000): Perspektiven der Lehrerbildung in Deutschland. Abschlussbericht der von der Kultusministerkonferenz eingesetzten Kommission. Weinheim, Basel: Beltz.
- Terhart, E. (2003): Wirkungen von Lehrerbildung: Perspektiven einer an Standards orientierten Evaluation. In: Journal für Lehrerinnen- und Lehrerbildung 3, S. 8-19.
- Terhart, E. (2006): Standards und Kompetenzen in der Lehrerbildung. In: Hilligus, A.H./Rinkens, H.-D. (Hrsg.): Standards und Kompetenzen – neue Qualität in der Lehrerbildung? Neue Ansätze und Erfahrungen in nationaler und internationaler Perspektive. Berlin: Lit Verlag, S. 29-42.
- Terhart, E. (2007): Standards in der Lehrerbildung – eine Einführung. In: Unterrichtswissenschaft 35, S. 2-14.
- The Plowden Report (1967): Children and their Primary Schools. A Report of the Central Advisory Council for Education. London: Her Majesty's Stationery Office.
URL: <http://www.dg.dial.pipex.com/documents/plowden00.shtml> (Stand 16. August 2006).
- The Swedish National Agency for Education (2003): School Choice and its Effects in Sweden. Offprint of Report 230. Stockholm: Fritzes.
- The Teaching Commission (2004): Teaching at Risk: A Call to Action. New York: The Teaching Commission.
- Thomas, R.M. (2005): High Stakes Testing. Coping with Collateral Damage. Mahwah, NJ und London: Lawrence Erlbaum Associates.
- Thomas B. Fordham Foundation (1999): The Teachers We Need and How to Get More of Them: A Manifesto. In: Kanstoroom, M./Finn, Ch.E. (Hrsg.): Better Teachers, Better Schools. Washington, DC: Thomas B. Fordham Foundation, S. 1-18.
- Tillmann, K.-J. (2001): Autonomie der Schule. Die bildungspolitische Diskussion in Deutschland und die Erfahrungen der Bielefelder Laborschule. In: Melzer, W./Sandfuchs, U. (Hrsg.): Was Schule leistet. Funktionen und Aufgaben von Schule. München: Juventa, S. 225-240.
- Tomlinson, S. (2001): Education in a Post-Welfare Society. Buckingham: Open University Press.
- Tooley, J./Dixon, P./Stanfield, J. (2003): Delivering Better Education. Market Solutions for Educational Improvement. London: Adam Smith Institute.

- Treasury (Great Britain) (2003): Every Child Matters. Presented to Parliament by the Chief Secretary to the Treasury by Command of Her Majesty. London: The Stationery Office (TSO).
- Tresch, S. (2007): Potenzial Leistungstest. Wie Lehrerinnen und Lehrer Ergebnisrückmeldungen zur Sicherung und Steigerung ihrer Unterrichtsqualität nutzen. Bern: h.e.p. verlag.
- Tresch, S./Moser, U. (2005): Check 5. Schlussbericht 2005 zuhanden des Departements Bildung, Kultur und Sport des Kantons Aarau. Zürich: Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich.
- Tyack, D./Tobin, W. (1994): The „Grammar“ of Schooling. Why Has It Been So Hard to Change? In: American Educational Research Journal 31, S. 453-479.
- Uhl, S. (2006): Die Bildungsstandards, die Outputsteuerung und ihre Kritiker. Hessisches Kultusministerium: Institut für Qualitätsentwicklung (auch unter: www.iq.hessen.de).
- U.S. Department of Education (Office of Postsecondary Education) (2002): Meeting the Highly Qualified Teachers Challenge. The Secretary's Annual Report on Teacher Quality. Washington, DC: U.S. Department of Education. URL: <http://www.title2.org/ADATitleIIReport2002.pdf> (Stand 28. Dezember 2005).
- U.S. Department of Education (Office of Postsecondary Education) (2003): Meeting the Highly Qualified Teachers Challenge. The Secretary's Second Annual Report on Teacher Quality. Washington, DC: U.S. Department of Education. URL: <http://www.ed.gov/about/reports/annual/teachprep/2003title-ii-report.pdf> (Stand 28. Dezember 2005).
- U.S. Department of Education (Office of Postsecondary Education) (2004): Meeting the Highly Qualified Teachers Challenge. The Secretary's Third Annual Report on Teacher Quality. Washington, DC: U.S. Department of Education. URL: <http://www.title2.org/TitleIIReport04.pdf> (Stand 28. Dezember 2005).
- Utbildningsdepartementet (2000): En förnyad lärarutbildning, Regeringens proposition 1999/2000: 135, Faktblad U00.016, Mai 2000.
- Ven, B. v.d. (2002): Niederlande. In: Döbert, H./Hörner, W./Kopp, B.v./Mitter, W. (Hrsg.): Die Schulsysteme Europas. Hohengehren: Schneider Verlag, S. 329-346.
- Victor, D.G. (2006): Recovering Sustainable Development. In: Foreign Affairs 85, Nr. 1, S. 91-103.
- Vroom, V.V./Jago, A.G. (2007): The Role of the Situation in Leadership. In: American Psychologist 62, Nr 1, S. 17-24.
- Wagenschein, M. (1968): Verstehen lehren. Weinheim: Beltz.
- Wahl, D. (2001): Nachhaltige Wege vom Wissen zum Handeln. In: Beiträge zur Lehrerbildung 19, 157-174.
- Wainer, H. (2000): Computerized Adaptive Testing. A Primer. Mahwah, NJ und London: Lawrence Erlbaum Associates.

- Walsh, K./Hale, C. (2004): *Increasing the Odds: How Good Policies Can Yield Better Teachers*. Washington, DC: National Council on Teacher Quality.
- Wang, J./Odell, S.J. (2002): *Mentored Learning to Teach and Standards-Based Teaching Reform: A Critical Review*. In: *Review of Educational Research* 7, S. 481-546.
- Washburne, C./Stearns, M.M. (1928): *Better Schools: A Survey of Progressive Education in American Public Schools*. New York: John Day Company.
- Watermann, R./Stanat, P./Kunter, M./Klieme, E./Baumert, J. (2003): *Schulrückmeldungen im Rahmen von Schulleistungsuntersuchungen: Das Disseminationskonzept von PISA 2000*. In: *Zeitschrift für Pädagogik* 49, S. 92-111.
- Webb, R./Vulliamy, G./Hakkinen, K./Hamalainen, S. (1998): *External Inspection or School Self-Evaluation? A Comparative Analysis of Policy and Practice in Primary Schools in England and Finland*. In: *British Educational Research Journal* 24, S. 539-556.
- Wehner, T./Legler, A./Siegrist, M. (2003): *Wissenschaftliche Evaluation der Mitarbeiterbeurteilung für Lehrkräfte der Zürcher Volksschule (EvaMAB)*. Bericht im Auftrag der Bildungsdirektion des Kantons Zürich. Zürich: ETHZ, Institut für Arbeitspsychologie.
- Weick, K.E. (1995): *Sensemaking in Organizations*. Thousand Oaks, CA: Sage Publications.
- Weih-Dengg, G. (2004): *Die Wiener Bildungsstandards*. In: *Erziehung und Unterricht* 154, S. 726-734.
- Weiner, L. (2003): *Research or „Cheerleading“? Scholarship on Community School District 2, New York City*. In: *Education Policy Analysis Archives* 11, Nr. 27. URL: <http://epaa.asu.edu/epaa/v11n27/> (13. Januar 2006).
- Weinert, F.E. (1999): *Concepts of Competence. Contribution within the OCED-Project Definition and Section of Competencies: Theoretical and Conceptual Foundations (DeSeCo)*. Neuchâtel: Bundesamt für Statistik.
- Weinert, F.E. (2001a): *Concepts of Competence: A Conceptual Clarification*. In: Rychen, D.S./Salganik, L.H. (Hrsg.): *Defining and Selecting Key Competencies*. Göttingen: Hogrefe & Huber, S. 45-65.
- Weinert, F.E. (2001b): *Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit*. In: Weinert, F.E. (Hrsg.): *Leistungsmessungen in Schulen*. Weinheim und Basel: Beltz, S. 17-31.
- Weinstein, R.S. (1996): *High Standards in a Tracked System of Schooling: For Which Students and With What Educational Supports?* In: *Educational Researcher* 25, Nr. 8, S. 16-18.
- Weishaupt, H. (2002): *Bildung und Region*. In: Tippelt, R. (Hrsg.): *Handbuch der Bildungsforschung*. Opladen: Leske + Budrich, S. 185-200.
- Welzel, M./Stadler, H. (Hrsg.) (2005): *Nimm doch mal die Kamera! Zur Nutzung von Videos in der Lehrerbildung. Beispiele und Empfehlungen aus den Naturwissenschaften*. Münster: Waxmann.

- Wendt, W. (2002): Orientierungsleitfaden zur Schulprogrammentwicklung. In: Schul-Verwaltung MO 12, S. 90-93. URL: <http://bebis.cidsnet.de/fortbildung/unterstuetzung/forum/> (Stand 12. März 2006).
- West, L./Staub, F.C. (2003): Content-Focused Coaching: Transforming Mathematics Lessons. Portsmouth, NH: Heinemann und Pittsburgh, PA: University of Pittsburgh.
- Whetton, C./Twist, E./Sainsbury, M. (2000): National Tests and Target Setting: Maintaining Consistent Standards. Paper presented at American Educational Research Association Annual Meeting, New Orleans. URL: http://www.nfer.ac.uk/publications/other-publications/conference-papers/pdf_docs/aera2000.PDF (Stand 7. September 2006).
- White, R.W. (1959): Motivation Reconsidered: The Concept of Competence. In: Psychological Review 66, S. 297-333.
- Whitehurst, G.J. (2002): Research on Teacher Preparation and Professional Development. Remarks Delivered at the White House Conference of Preparing Tomorrow's Teachers. URL: <http://www.ed.gov/admins/tchrqual/learn/preparingteachersconference/whitehurst.html> (Stand 12. Januar 2006).
- Whitford, B.L./Metcalf-Turner, P.M. (1998): Of promises and Unresolved Puzzles: Reforming Teacher Education with Professional Development Schools. National Society for the Study of Education (NSSE) Yearbook. Chicago, IL: University of Chicago Press.
- Whitty, G./Power, S./Halpin, D. (1998): Devolution and Choice in Education. The School, the State, and the Market. Buckingham: Open University Press.
- Wiburg, K./Brown, S. (2007): Lesson Study Communities: Increasing Achievement With Diverse Students. Thousand Oaks: Corwin Press.
- Wiggenhorn, G./Vorndran, O. (2002): Integrating IT into Schools. An International Study on Regional Implementation Strategies. Bertelsmann Foundation/Heinz Nixdorf Foundation. URL: www.netzwerk-medienschulen.de (Stand 6. Januar 2006).
- Wikström, C. (2005): Grade Stability in a Criterion-Referenced Grading System: The Swedish Example. In: Assessment in Education 12, S. 125-144.
- Wilson, D. (2003): Unravelling Control Freakery: Redefining Central-Local Government Relations. In: The British Journal of Politics & International Relations 5, S. 317-346.
- Wilson, S.M./Berne, J. (1999): Teacher Learning and the Acquisition of Professional Knowledge. An Examination of Research on Contemporary Professional Development. In: Iran-Nejad, A./Pearson, P.D. (Hrsg.): Review of Research in Education 24. Washington, DC: AERA, S. 173-209.
- Winch, Ch. (1987): Who Controls Education? URL: <http://members.aol.com/BevinSoc/l2edu.htm> (Stand 20. August 2006).

- Winick, D./Kress, S. (2004): Accountability Works in Texas. In: Evers, W.M./Walberg, H.J. (Hrsg.): *Testing Student Learning, Evaluation Teacher Effectiveness*. Stanford, CA: Hoover Press 2004, S. 303-321.
- Winter, F. (2002): Ein Instrument mit vielen Möglichkeiten. Leistungsbewertung anhand von Portfolios. In: Winter, F./Groeben, A.v.d./Lenzen, K.D. (Hrsg.): *Leistung sehen, fördern, bewerten. Neue Wege für die Schule*. Bad Heilbrunn: Klinkhardt, S. 175-183.
- Winter, F. (2004): Leistungsbewertung. Eine neue Lernkultur braucht einen anderen Umgang mit den Schülerleistungen. Schneider: Baltmannsweiler.
- Winter, F. (2005): Standards auch von unten? Warum die Schule die Leistungsbewertung nicht delegieren darf. In: Becker, G./Bremerich-Vos, A./Demmer, M./Maag Merki, K./Proebe, B./Schwippert, K./Stäudel, L./Tillmann, K.-J. (Hrsg.): *Standards. Unterrichten zwischen Kompetenzen, zentralen Prüfungen und Vergleichsarbeiten*. Friedrich Jahresheft XXIII, S. 76-77.
- Withford, B.L./Jones, K. (Hrsg.) (2000): *Accountability, Assessment, Teacher Commitment: Lessons from Kentucky's Reform Efforts*. Albany, NY: Suny Press.
- Wittmann, E. C. (2004): Was ist Mathematik und was bedeutet dies für die Entwicklung von Lehrmitteln für den Mathematikunterricht? In: Aeberli, C. (Hrsg.): *Lehrmittel neu diskutiert. Ergebnisse des 1. Schweizerischen Lehrmittelsymposiums vom 29. und 30. Januar 2004*. Zürich: Lehrmittelverlag.
- Wolf, A. (2000): A Comparative Perspective on Educational Standards. In: Goldstein, H./Heath, A. (Hrsg.): *Educational Standards*. Oxford: Oxford University Press, S. 9-37 (= Proceedings of the British Academy, Bd. 102).
- Wolf, P.J./Peterson, P.E./West, M.R. (2001): Results of a School Voucher Experiment: The Case of Washington D.C. After Two Years. Paper prepared for the annual meeting of the American Political Science Association, San Francisco, CA, August/September.
- Wong, K.K./Anagnostopoulos, D./Rutledge, S./Edwards, C. (2001): *The Challenge of Improving Instruction in Urban High Schools: Case Studies of the Implementation of the Chicago Academic Standards*. Chicago, IL: University of Chicago.
- Woodhead, Ch. (2004): The British Experience. In: *Education Next*, Summer, S. 52-57.
- Wright, S.P./Horn, S.P./Sanders, W.L. (1997): Teachers and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation. In: *Journal of Personnel Evaluation in Education* 11, S. 57-67.
- Yanow, D. (1996): *How Does Policy Mean? Interpreting Policy and Organizational Actions*. Washington, DC: Georgetown University Press.
- Youngs, P. (2001): District and State Policy Influences on Professional Development and School Capacity. In: *Educational Policy* 15, S. 278-301.
- Zahner Rossier, C./Berweger, S./Brühwiler, C./Holzer, T./Mariotta, M./Moser, U./Nicoli, M. (2004): *PISA 2003: Kompetenzen für die Zukunft. Erster nationaler Bericht*. Neuchâtel und Bern: BFS/EDK.

- Zeller, A. (2004): „Stellwerk“ soll Jugendliche auf die richtige Schiene weisen. In: Bildung Schweiz 149, Nr. 9, S. 8-9. URL: www.lch.ch/bildungschweiz/pdfs/2004/artikel/09/Stellwerk.pdf (Stand 20. März 2006).
- Ziegler, B. (2006): Transferstrategien und Transfermaßnahmen im Modellversuch SINUS-Transfer. In: Nickolaus, R./Gräsel, C. (Hrsg.): Innovation und Transfer – Expertisen zur Transferforschung. Schneider: Hohengehren, S. 117-142.
- Zucker, A.A./Shields, P.M. (1995): Evaluation of the National Science Foundation's Statewide Systemic Initiatives Program: Second Year Report – Second Year Case Studies: Connecticut, Delaware, Montana. Washington, DC: NSF, Directorate for Education and Human Resources.
- Zymek, B. (2007): Die Aktualität der regionalen Schulentwicklung als Gegenstand der empirischen Bildungsforschung. In: Zeitschrift für Pädagogik 53, Nr. 3, S. 279-283.

Diese Druckschrift wird im Rahmen der Öffentlichkeitsarbeit vom Bundesministerium für Bildung und Forschung unentgeltlich abgegeben. Sie ist nicht zum gewerblichen Vertrieb bestimmt. Sie darf weder von Parteien noch von Wahlbewerberinnen/Wahlbewerbern oder Wahlhelferinnen/Wahlhelfern während eines Wahlkampfes zum Zwecke der Wahlwerbung verwendet werden. Dies gilt für Bundestags-, Landtags- und Kommunalwahlen sowie für Wahlen zum Europäischen Parlament. Missbräuchlich ist insbesondere die Verteilung auf Wahlveranstaltungen und an Informationsständen der Parteien sowie das Einlegen, Aufdrucken oder Aufkleben parteipolitischer Informationen oder Werbemittel. Untersagt ist gleichfalls die Weitergabe an Dritte zum Zwecke der Wahlwerbung.

Unabhängig davon, wann, auf welchem Weg und in welcher Anzahl diese Schrift der Empfängerin/dem Empfänger zugegangen ist, darf sie auch ohne zeitlichen Bezug zu einer bevorstehenden Wahl nicht in einer Weise verwendet werden, die als Parteinahme der Bundesregierung zugunsten einzelner politischer Gruppen verstanden werden könnte.



Bundesministerium
für Bildung
und Forschung

