

# **Zur Entwicklung nationaler Bildungsstandards**

**- Eine Expertise -**

**Vorgestellt von**

***Edelgard Bulmahn***

**Bundesministerin für Bildung und Forschung**

***Karin Wolff***

**Präsidentin der Kultusministerkonferenz und Staatsministerin für Kultus (Hessen)**

***Prof. Dr. Eckhard Klieme***

**Deutsches Institut für Internationale Pädagogische Forschung (DIPF)**

**am 18. Februar 2003 in Berlin**

## **Autoren:**

*Prof. Dr. Eckhard Klieme* (Koordination), Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt a.M., Arbeitseinheit Bildungsqualität und Evaluation

*Prof. Dr. Hermann Avenarius*; Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt a.M., Arbeitseinheit Finanzierung und Steuerung des Bildungswesens

*Prof. Dr. Werner Blum*, Universität Kassel, Fachbereich Mathematik/ Informatik

*Dr. Peter Döbrich*, Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt a.M., Arbeitseinheit Bildungsqualität und Evaluation

*Prof. Dr. Hans Gruber*, Universität Regensburg, Lehrstuhl für Pädagogik III, Lehr-Lern-Forschung und Medienpädagogik

*Prof. Dr. Manfred Prenzel*, Leibniz-Institut für die Pädagogik der Naturwissenschaften, Kiel, Abt. Erziehungswissenschaft

*Prof. Dr. Kristina Reiss*, Universität Augsburg, Lehrstuhl für Didaktik der Mathematik

*Dr. Kurt Riquarts*, Leibniz-Institut für die Pädagogik der Naturwissenschaften, Kiel, Abt. Erziehungswissenschaft

*Prof. Dr. Jürgen Rost*, Leibniz-Institut für die Pädagogik der Naturwissenschaften, Kiel, Abt. Pädagogisch-psychologische Methodenlehre

*Prof. Dr. Heinz-Elmar Tenorth*, Humboldt-Universität Berlin, Vizepräsident für Lehre und Studium, Historische Erziehungswissenschaft

*Prof. Dr. Helmut J. Vollmer*, Universität Osnabrück, Fachbereich Sprach- und Literaturwissenschaft

Wissenschaftliche Mitarbeit: Jutta Laukart, DIPF; Dr. Andreas Paetz, DIPF

Koordination: Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt a.M.



Gefördert durch das Bundesministerium für Bildung und Forschung (BMBF)

# Inhaltsverzeichnis

|                                                                                                    |     |
|----------------------------------------------------------------------------------------------------|-----|
| Zusammenfassung                                                                                    | 4   |
| 1. Anlass, Ziel und Aufbau der Expertise                                                           | 6   |
| <i>Konzeption und Funktion von Bildungsstandards</i>                                               |     |
| 2. Konzeption von Bildungsstandards                                                                | 13  |
| 3. Beispiele: Standards und Curricula aus dem Bereich der Mathematik                               | 28  |
| 4. Zur Bedeutung von Standards für die Weiterentwicklung von Schule und Unterricht                 | 38  |
| <i>Grundlagen für die Entwicklung von Bildungsstandards</i>                                        |     |
| 5. Bildungsstandards, Kompetenzmodelle und Bildungsziele: Konstruktions- und Legitimationsprobleme | 45  |
| 6. Kompetenzmodelle aus pädagogisch-psychologischer Sicht                                          | 58  |
| 7. Testentwicklung: Methodische und diagnostische Grundlagen                                       | 67  |
| <i>Konsequenzen der Einführung von Standards für das Bildungssystem</i>                            |     |
| 8. Folgen für Lehrplanentwicklung und schulische Lehrpläne: Standards und Kerncurricula            | 74  |
| 9. Bildungsmonitoring, Evaluation und Individualdiagnostik auf der Basis von Standards             | 81  |
| 10. Aufgaben der Unterstützungssysteme                                                             | 90  |
| <i>Entwicklung und Implementation von Bildungsstandards in Deutschland</i>                         |     |
| 11. Die Infrastruktur für Entwicklung und Implementation von Bildungsstandards                     | 98  |
| 12. Stand und Perspektiven der Entwicklung von Bildungsstandards in Deutschland                    | 109 |
| Anhang: Beispiele für die weitere Entwicklung von Kompetenzmodellen                                |     |
| (a) Der Gemeinsame Europäische Referenzrahmen für Sprachen                                         | 122 |
| (b) Der <i>Common Framework for Science Outcomes</i> (Kanada)                                      | 131 |
| Literaturverzeichnis                                                                               | 141 |

# Zusammenfassung

Nationale Bildungsstandards formulieren verbindliche **Anforderungen an das Lehren und Lernen in der Schule**. Sie stellen damit innerhalb der Gesamtheit der Anstrengungen zur **Sicherung und Steigerung der Qualität schulischer Arbeit** ein zentrales Gelenkstück dar. Bildungsstandards benennen präzise, verständlich und fokussiert die wesentlichen Ziele der pädagogischen Arbeit, ausgedrückt als erwünschte Lernergebnisse der Schülerinnen und Schüler. Damit konkretisieren sie den Bildungsauftrag, den Schulen zu erfüllen haben.

Für die **Gestaltung von Bildungsstandards** werden in dieser Expertise folgende Vorschläge gemacht:

Bildungsstandards greifen allgemeine *Bildungsziele* auf. **Die Bildungsstandards legen fest, welche Kompetenzen die Kinder oder Jugendlichen bis zu einer bestimmten Jahrgangsstufe mindestens erworben haben sollen.** Die Kompetenzen werden so konkret beschrieben, dass sie in Aufgabenstellungen umgesetzt und prinzipiell mit Hilfe von *Testverfahren* erfasst werden können. Der Darstellung von Kompetenzen, die innerhalb eines Lernbereiches oder Faches aufgebaut werden, ihrer Teildimensionen und Niveaustufen, kommt in diesem Konzept ein entscheidender Platz zu. **Kompetenzmodelle** konkretisieren Inhalte und Stufen der allgemeinen Bildung. Sie formulieren damit eine **pragmatische Antwort auf die Konstruktions- und Legitimationsprobleme traditioneller Bildungs- und Lehrplandebatten**. Die Expertise benennt mehrere Beispiele für Kompetenzmodelle aus der Mathematik (etwa bei PISA), aus dem Fremdsprachenlernen und der Naturwissenschaftsdidaktik. Diese Modelle stützen sich auf fachdidaktisches und pädagogisch-psychologisches Wissen.

Die **erste Funktion** der Bildungsstandards besteht in der **Orientierung** der Schulen auf verbindliche Ziele. Lehrkräfte, aber auch Lernende und deren Eltern können sich darauf bei der pädagogischen Weiterentwicklung von Schule und Unterricht beziehen.

Kompetenzmodelle bieten den Lehrerinnen und Lehrern ein Referenzsystem für ihr professionelles Handeln. Bildungsstandards lassen den Schulen einen starken **Freiraum für die inner-schulische Lernplanung**, zumal wenn die Lehrpläne und Rahmenrichtlinien der Länder auf Kerncurricula begrenzt werden. Diesen Freiraum zu nutzen, kann die Schule voranbringen, erfordert aber auch **Unterstützung** durch Einrichtungen der Lehrerbildung, Schulaufsicht und Landesinstitute.

Eine **zweite Funktion** der Bildungsstandards besteht darin, dass auf ihrer Grundlage **Lernergebnisse erfasst und bewertet** werden. Mit Bezug auf die Bildungsstandards kann man überprüfen, ob die angestrebten Kompetenzen tatsächlich erworben wurden. So lässt sich feststellen, inwieweit das Bildungssystem seinen Auftrag erfüllt hat (**Bildungsmonitoring**), und die Schulen

erhalten eine Rückmeldung über die Ergebnisse ihrer Arbeit (**Schulevaluation**). Die Standards können auch Hinweise geben für die individuelle Diagnostik und Förderung. Allerdings legt die Expertise Wert darauf, dass Tests, die im Bildungsmonitoring und für die Schulevaluation eingesetzt werden, solche Individualdiagnostik aus methodischen Gründen meist nicht erlauben. Von einer Verwendung der Standards bzw. standard-bezogener Tests für Notengebung und Zertifizierung wird abgeraten.

Die Expertise beschreibt in den beiden letzten Kapiteln die **Infrastruktur**, die in Deutschland für Entwicklung und Implementation der Bildungsstandards sowie für Bildungsmonitoring und Schulevaluation entstehen sollte, und die **Arbeitsschritte** der nächsten Jahre.

Die bisherigen Arbeiten in den Ländern und in der KMK bilden einen guten Ausgangspunkt. Aufgaben der näheren Zukunft sind vor allem die **fachdidaktische Vertiefung in Kompetenzmodellen**, die Festlegung von **Mindeststandards**, die Entwicklung von **Aufgabenpools und Testverfahren** sowie die **Implementation** an den Schulen. Es wird betont, dass die verbindliche Festsetzung der Standards und die standard-bezogene Evaluation von Schulen **Angelegenheit der einzelnen Länder** sind. Für die Entwicklung von Standards und Tests sowie das (nationale) Bildungsmonitoring – das an PISA und andere internationale Untersuchungen gekoppelt werden könnte – schlägt die Expertise jedoch den **Aufbau einer nationalen Agentur** vor.

Wenn es gelingt, Bildungsstandards so zu gestalten, dass sich in ihnen eine Vision von Bildungsprozessen abzeichnet, eine moderne „Philosophie“ der Schulfächer, eine Entwicklungsperspektive für die Fähigkeiten von Schülern, dann können die Standards zu einem Motor der pädagogischen Entwicklung unserer Schulen werden.

# 1. Anlass, Ziel und Aufbau der Expertise

## 1.1 Die Ausgangslage:

### **Bildungsstandards als Reaktion auf Probleme des Bildungssystems**

Seit Veröffentlichung der TIMSS-Ergebnisse (Baumert, Lehmann et al. 1997, Baumert, Bos & Lehmann 2000 a, b) wird in Deutschland verstärkt über den Zustand und die Entwicklungsperspektiven des Bildungssystems debattiert. Das von Bund und Ländern gemeinsam getragene *Forum Bildung* und die große Resonanz auf die PISA-Studie (OECD 2001, Baumert et al. 2001, Baumert et al. 2002 a) haben diese Debatte in den vergangenen Monaten zu einer seit langem nicht mehr erlebten Intensität und Breite geführt. Im Vordergrund der Diskussion steht die Qualität von Schule und Unterricht im Bereich der Sekundarstufen I und II (5. bis 13. Schuljahr), aber zumindest mittelbar sind auch die Elementarerziehung in den Kindergärten und die Grundschulen herausgefordert. Über Landes- und Parteigrenzen hinweg sind Bildungspolitiker zu einschneidenden Reformen bereit.

Unabweisbar haben die empirischen Studien, die nach fast 20 Jahren erstmals die Realität der Schulen analysiert und im internationalen Kontext verglichen haben, gravierende Mängel offen gelegt. Der Vergleich von Lernergebnissen zwischen den Ländern der Bundesrepublik Deutschland und mit anderen Staaten ist vielfach Anlass für Enttäuschung und Kritik gewesen. Viel wichtiger als die „Rankings“ sind aber strukturelle Erkenntnisse, die Hinweise darauf geben können, an welchen Stellen Reformen notwendig sind. Die Studien verweisen auf Disparitäten, die unser Schulsystem nach wie vor in besonderem Maße kennzeichnen: zwischen Regionen, zwischen Kindern aus unterschiedlichen sozialen Schichten, zwischen in Deutschland Aufgewachsenen und Zugewanderten. Die Studien belegen die Bedeutung des sozialen und familiären Kontextes für den Schulerfolg, aber sie geben auch Hinweise zur Gestaltung von Lehr- und Lernprozessen, beispielsweise zur Leseförderung als fächerübergreifender Aufgabe oder zu einem anspruchsvollen, verständnisorientierten Unterricht in Mathematik und Naturwissenschaften.

In der Bildungspolitik und der Bildungsverwaltung haben TIMSS und PISA eine grundsätzliche Wende eingeleitet. Wurde unser Bildungssystem bislang ausschließlich durch den „Input“ gesteuert, d.h. durch Haushaltspläne, Lehrpläne und Rahmenrichtlinien, Ausbildungsbestimmungen für Lehrpersonen, Prüfungsrichtlinien usw., so ist nun immer häufiger davon die Rede, die Bildungspolitik und die Schulentwicklung sollten sich am „Output“ orientieren, d.h. an den Leistungen der Schule, vor allem an den Lernergebnissen der Schülerinnen und Schüler. Der Output von Bildungssystemen umfasst neben der Vergabe von Zertifikaten im Wesentlichen den Aufbau von

Kompetenzen, Qualifikationen, Wissensstrukturen, Einstellungen, Überzeugungen, Werthaltungen – also von Persönlichkeitsmerkmalen bei den Schülerinnen und Schülern, mit denen die Basis für ein lebenslanges Lernen zur persönlichen Weiterentwicklung und gesellschaftlichen Beteiligung gelegt ist. Die Schulen und die Bildungsadministration sollen – ungeachtet der Rolle, die die Schüler selbst und die Eltern spielen – Verantwortung dafür übernehmen, dass diese Ergebnisse tatsächlich erreicht werden. Der Output wird somit zum entscheidenden Bezugspunkt für die Beurteilung des Schulsystems und für Maßnahmen zur Verbesserung und Weiterentwicklung. Diese „Output-Orientierung“ lässt sich zugleich zwanglos mit einem gewandelten Verständnis staatlicher Steuerung verbinden, wie es auch in anderen Bereichen der Gesellschaft zu beobachten ist. Nicht mehr durch detaillierte Richtlinien und Regelungen, sondern durch Definition von Zielen, deren Einhaltung auch tatsächlich überprüft wird, sorgt der Staat für Qualität.

Die Kriterien, nach denen der Output des Bildungssystems beurteilt wird, und die zugehörigen Indikatoren (d.h. die Kennzahlen, die zur Qualitätsmessung herangezogen werden) können und müssen vielschichtig sein. Sowohl das Schulsystem insgesamt als auch einzelne Schulen sollten sicherlich daran gemessen werden, welcher Teil der Schülerinnen und Schüler die angestrebten Kompetenzen, Überzeugungen usw. tatsächlich erworben hat. Weder das „Systemmonitoring“ noch die Evaluation von Einzelschulen dürfen aber bei der Feststellung von Mittelwerten in Schulleistungstests und Einstellungsskalen stehen bleiben – so breit der abgedeckte Bereich an Kompetenzen und Einstellungen auch sein mag. Systemziele wie etwa der Abbau von Disparitäten, die Förderung von Kindern und Jugendlichen mit Benachteiligungen, die Integration von Migranten und die Begabtenförderung oder die Flexibilität und Offenheit von Bildungsverläufen müssen berücksichtigt werden. Will man – etwa im Rahmen einer nationalen Bildungsberichterstattung – das Bildungssystem bewerten oder einzelne Schulen evaluieren, darf man sich daher nicht nur auf das Durchschnittsniveau der Lernergebnisse stützen. Schule hat nicht nur die Funktion, individuelle Leistungsfähigkeit sicherzustellen, sondern sie dient auch der kulturellen Selbstverständigung und dem sozialen Zusammenhalt der Gesellschaft.

Gleichwohl bilden die Bildungsziele und die tatsächlich erreichten Lernergebnisse den Kern der Qualitätsdebatte. Diese und nur diese Thematik ist Gegenstand der vorliegenden Expertise. Damit rückt die Frage ins Zentrum, welche Erwartungen wir an die Lernergebnisse in den Schulen stellen und inwieweit diese erfüllt werden. Hierzu haben eine Reihe von Schulleistungsstudien, die in den vergangenen Jahren durchgeführt wurden, wichtige Erkenntnisse erbracht: LAU in Hamburg (Lehmann, Gänsfuß & Peek 1999), QUASUM in Brandenburg (Lehmann & Peek 1997), MARKUS in Rheinland-Pfalz (Helmke & Jäger 2002), landesweite Vergleichsarbeiten in anderen Ländern. Zuletzt hat PISA (Baumert et al. 2001, Baumert et al. 2002a) gezeigt:

- Die selbst gesteckten, in den Lehrplänen festgehaltenen Ziele werden häufig nicht erreicht. In der Lesekompetenz und in der Mathematik erreicht beispielsweise mindestens ein Viertel der Fünfzehnjährigen nicht das Niveau, das nach Meinung von Lehrplanexperten für einen erfolgreichen Abschluss der Sekundarstufe I unabdingbar wäre.
- Die Leistungen der Schüler variieren innerhalb Deutschlands so stark wie in keinem anderem PISA-Teilnehmerstaat. Im oberen Leistungsbereich kann Deutschland mit den meisten OECD-Staaten mithalten, aber im unteren Leistungsbereich fallen die deutschen Schüler erheblich gegen die Schüler anderer Staaten ab. Dies wurde als Hinweis darauf interpretiert, dass es in unserem Land an Mindeststandards fehlt, die in der Ausbildung beispielsweise von Lesekompetenz oder mathematischer Kompetenz erreicht werden müssen.
- Die Zusatzstudie PISA-E hat darüber hinaus gezeigt, dass es innerhalb Deutschlands erhebliche Unterschiede zwischen den Ländern gibt. Nicht nur das Leistungsniveau, sondern auch die Benotungsmaßstäbe unterscheiden sich deutlich von Land zu Land. Im Interesse des Gleichheitsgebotes (gleiche Bildungschancen für alle Jugendlichen, unabhängig vom Wohnort) sollten solche Diskrepanzen verringert werden.
- Die internationalen Ergebnisse der PISA-Studie, insbesondere die sehr guten Ergebnisse der skandinavischen Staaten und einiger anglo-amerikanischer Staaten, lassen vermuten, dass Länder, die systematische Qualitätssicherung betreiben – sei es durch regelmäßige Schulleistungsstudien oder durch ein dichtes Netz von Schulevaluationen – insgesamt höhere Leistungen erreichen.

Als Antwort auf diese Erkenntnisse haben Bildungspolitiker aller Parteien wiederholt vorgeschlagen, zukünftig verbindlich festzulegen, welche Kompetenzen Schülerinnen und Schüler in Deutschland erreichen sollen, und darüber hinaus die Erfüllung dieser „Standards“ regelmäßig zu prüfen. Die Kultusministerkonferenz (KMK) hat daher am 25. Juni 2002 beschlossen, nationale Bildungsstandards in Kernfächern für bestimmte Jahrgangsstufen und Abschlussklassen zu erarbeiten. Zur Überprüfung sind landesweite Orientierungs- und Vergleichsarbeiten vorgesehen, auch bundesweite Vergleichsuntersuchungen. Dazu sollen z.B. PISA sowie der gegenwärtig vorbereitete Deutsch-Englisch-Test DESI (DESI-Konsortium 2001, Beck & Klieme 2003) genutzt werden. Die hier gewonnenen Daten sollen zudem in eine zukünftige nationale Bildungsberichtserstattung einfließen. Ähnliche Vorschläge wurden im *Forum Bildung* von Bundes- und Landesvertretern gemeinsam formuliert.

## 1.2 Auftrag und Ziel der Expertengruppe

Wer die Diskussion der letzten Monate in Deutschland verfolgt hat, muss allerdings feststellen: Es fehlt noch an Wissen darüber, was genau Standards bedeuten, wie sie zu formulieren sind, welche Rolle sie bei der Qualitätsentwicklung im Bildungssystem spielen können und wie ihre Einhaltung überprüft werden kann. Das vorliegende Gutachten hat daher zum Ziel, das Konzept der Bildungsstandards fachlich zu klären und eine Rahmenkonzeption dafür zu vorzustellen, wie Bildungsstandards für das deutsche Schulsystem angelegt sein sollten und wie sie entwickelt und genutzt werden könnten. Dabei soll die internationale bildungspolitische und bildungswissenschaftliche Diskussion berücksichtigt werden. Die Konzeption von Bildungsstandards sollte zudem mit anderen Kriterien von System- und Schulqualität wie beispielsweise dem Abbau von Disparitäten kompatibel sein bzw. sogar zu deren Einlösung beitragen.

Die Anregung zur Erstellung des Gutachtens kam vom Bundesministerium für Bildung und Forschung. Es beauftragte das Deutsche Institut für Internationale Pädagogische Forschung (DIPF) in Frankfurt a.M., eine interdisziplinäre Expertengruppe zusammenzustellen und eine Expertise anzufertigen, die alle Akteure der Bildungspolitik darin unterstützen könnte, verbindliche nationale Bildungsstandards zu entwickeln, zu implementieren und für die Qualitätsentwicklung zu nutzen. Die Ständige Konferenz der Kultusminister und -senatoren der Länder (KMK) wurde von dem Vorhaben unterrichtet. Sie hat ihr Interesse erklärt, Ergebnisse der Expertise in der laufenden Arbeit an Bildungsstandards zu berücksichtigen, und stellt nun gemeinsam mit dem Bundesministerium für Bildung und Wissenschaft (BMBWF) das Gutachten in der Öffentlichkeit vor.

Der Arbeitsgruppe, die das nun vorliegende Gutachten binnen fünf Monaten erarbeitet hat, gehörten Experten aus dem DIPF, aus dem Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN) in Kiel und aus verschiedenen Hochschulen an. Sie repräsentieren sehr unterschiedliche Arbeitsgebiete, die für das Thema Bedeutung haben: Allgemeine Erziehungswissenschaft und Empirische Bildungsforschung, Lehr-Lern-Forschung, Bildungsrecht, Historisch-Systematische Erziehungswissenschaft, Pädagogisch-Psychologische Methodenlehre sowie die Fachdidaktiken der Mathematik und der Fremdsprachen.

Die Expertinnen und Experten waren sich schnell darin einig, das Thema „Bildungsstandards“ nicht isoliert und technisch zu behandeln, sondern in den Kontext von Bildungsreformen zu stellen. Das Gutachten fragt daher nicht nur nach dem Begriff des Standards und seiner „technischen“ Ausführung, sondern vor allem nach den Funktionen von Standards für die Qualitätsentwicklung im Bildungswesen. Wenn es gelingt, Standards so zu gestalten, dass sich in ihnen eine

Vision von Bildungsprozessen abzeichnet, eine moderne „Philosophie“ der Schulfächer, eine Entwicklungsperspektive für die Fähigkeiten von Schülern, dann könnten Standards selbst zum Motor der pädagogischen Entwicklung unserer Schulen werden. Bildungsstandards, wie sie hier verstanden werden, sorgen dafür, dass pädagogische Schulentwicklung sich an klaren, verbindlichen Zielen orientiert und aus Ergebnissen systematisch lernt. Sie stehen dabei keineswegs im Gegensatz zur Eigenverantwortlichkeit der Schulen und zur professionellen Verantwortung jedes einzelnen Lehrers. Sie werden den Unterricht gerade nicht in ein enges Korsett zwingen, sondern den Schulen inhaltliche und pädagogische Freiräume geben.

Selbstverständlich bedarf es längerer Entwicklungsarbeit, bis Bildungsstandards vorliegen, die diese Funktionen in vollem Umfang erfüllen können. Sollen die Standards an den Schulen wirksam werden, sind zudem vielfältige Unterstützungsmaßnahmen in der Lehrplanarbeit, in der Lehrerbildung, in der Schulaufsicht und -beratung, im Aufbau von Evaluationssystemen usw. nötig. Unser Bildungssystem kann aber nicht lange auf den Beginn dieser Reformen warten. Die Kultusministerkonferenz hat mit dieser Arbeit auch schon begonnen und dieses Gutachten soll dazu anregen, eine breite, aber auch wissenschaftlich fundierte und international anschlussfähige Diskussion über die weitere Strategie im Umgang mit nationalen Bildungsstandards in Deutschland zu führen.

### **1.3 Aufbau der Expertise**

Im Anschluss an diese Einleitung gliedert sich die Expertise in vier große Abschnitte:

#### ***1) Konzeption und Funktion von Bildungsstandards: Kapitel 2 bis 4***

Die drei Kapitel dieses Teils bilden den Kern des Gutachtens.

Kapitel 2 führt das Konzept der Bildungsstandards ein: Sie konkretisieren den Bildungsauftrag der Schule durch allgemein verbindliche Kompetenzanforderungen. Zentrale Merkmale guter Standards werden genannt und das hier ausgeführte Verständnis von Standards wird in die internationale Diskussion eingeordnet.

Zur Illustration und als Beleg für die Möglichkeit, Bildungsstandards auf differenzierten Kompetenzmodellen aufzubauen, werden in Kapitel 3 Beispiele aus der Mathematik dargestellt. (Ergänzende Beispiele aus dem Bereich des Fremdsprachenlernens und der Naturwissenschaften finden sich im Anhang.)

Was die Einführung von Bildungsstandards schließlich konkret für die Weiterentwicklung von Schule und Unterricht bedeutet, wird in Kapitel 4 erläutert. Es beschreibt die Funktion der Standards aus der Sicht von Lehrern, Schülern und Eltern, für die einzelne Schule und das Bildungssystem insgesamt. Es wird deutlich, dass Standards nicht nur ein neues Konzept der Steuerung im Bildungssystem einführen, sondern vor allem Anstöße zur pädagogischen Schulentwicklung geben. Es wird Wert darauf gelegt, dass Bildungsstandards der Qualitätsentwicklung im Bildungswesen dienen, aber keine Zentralisierung von Prüfungen und Zertifizierungen beinhalten.

## **II) Grundlagen für die Entwicklung von Bildungsstandards: Kapitel 5 bis 7**

Dieser Teil der Expertise erläutert die in Kapitel 2 vorgestellten Grundkomponenten ausführlicher aus wissenschaftlicher Sicht:

Gesellschaftlich vorgegebene *Bildungsziele* (Kapitel 5) bilden den Orientierungsrahmen für Bildungsstandards. Die Verankerung der Standards und Kompetenzmodelle in allgemeinen Bildungszielen gibt ihnen ein bildungstheoretisches Fundament und bietet zugleich eine pragmatische Lösung für einige klassische Konstruktions- und Legitimationsprobleme der Curriculumentwicklung.

*Kompetenzmodelle* (Kapitel 6) beschreiben auf der Basis fachdidaktischer Konzepte die Komponenten und Stufen der Kompetenzen und stützen sich dabei auf pädagogisch-psychologische Forschungen zum Aufbau von Wissen und Können.

*Testverfahren* (Kapitel 7) schließlich operationalisieren die Kompetenzanforderungen durch konkrete Aufgaben. Sie dienen unterschiedlichen diagnostischen und evaluativen Zielen und ermöglichen zugleich die empirische Überprüfung der Kompetenzmodelle.

## **III) Konsequenzen der Einführung von Bildungsstandards für das Bildungssystem: Kapitel 8 bis 10**

Welche Bedeutung die Einführung von Standards für die Weiterentwicklung der Schulen hat, wurde in Kapitel 4 bereits zusammenfassend dargestellt. In den Kapiteln 8 bis 10 werden einige zentrale, für das Bildungssystem als Ganzes relevante Faktoren detaillierter behandelt:

Die zentrale *Lehrplanentwicklung* wird durch Bildungsstandards prinzipiell entlastet, weil curriculare Regelungen unterhalb der Ebene der Standards tendenziell in die Schulen verlagert werden können. Kapitel 8 diskutiert diese neue Form von Lehrplanentwicklung. Es wird vorgeschlagen, Bildungsstandards vorerst Hand in Hand mit zentralen Kerncurricula zu entwickeln, die zusätzliche Hinweise auf Sequenzierung und unterrichtliche Umsetzung von Themen und Inhalten geben.

Welche Rolle standardbezogene Testverfahren für das *Bildungsmonitoring* auf der Ebene des Gesamtsystems und die *Evaluation* auf der Ebene von Einzelschulen haben, wird in Kapitel 9 erläutert. Hier wird auch diskutiert, unter welchen Bedingungen standardbezogene Testverfahren für die Individualdiagnostik, d.h. als Basis von Förderentscheidungen in Einzelfällen verwendet werden können.

Die Umsetzung der Bildungsstandards in den Schulen erfordert vermehrte Unterstützungstätigkeiten durch Lehrerbildung, Schulaufsicht und andere Akteure, die in Kapitel 10 behandelt werden.

#### **IV) *Entwicklung und Implementation von Bildungsstandards in Deutschland:* *Kapitel 11 bis 12***

In den beiden Schlusskapiteln werden die schulpraktischen sowie die bildungspolitischen und -rechtlichen Fragen behandelt, die mit der Einführung von Bildungsstandards verbunden sind, und es werden Empfehlungen für die weitere Arbeit gegeben.

Kapitel 11 unterscheidet verschiedene Aufgabenfelder im Rahmen der Standardentwicklung und -nutzung und schlägt vor, wie die erforderliche Infrastruktur aussehen kann.

Kapitel 12 schließlich berichtet über bisherige und laufende Arbeiten an Bildungsstandards in Deutschland und benennt Desiderata für die Arbeit der nächsten Jahre.

## 2. Konzeption von Bildungsstandards

### 2.1 Was in dieser Expertise unter „Bildungsstandards“ verstanden wird

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.

Bildungsstandards, wie sie in dieser Expertise konzipiert werden, greifen allgemeine *Bildungsziele* auf. Sie benennen die *Kompetenzen*, welche die Schule ihren Schülerinnen und Schülern vermitteln muss, damit bestimmte zentrale Bildungsziele erreicht werden. 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.

Bildungsstandards stellen damit innerhalb der Gesamtheit der Anstrengungen zur Sicherung und Steigerung der Qualität schulischer Arbeit ein zentrales Gelenkstück dar. Schule und Unterricht können sich an den Standards orientieren. Den Lehrerinnen und Lehrern geben Bildungsstandards ein Referenzsystem für ihr professionelles Handeln. Die Kompetenzanforderungen einzulösen, so gut dies unter den Ausgangsbedingungen der Schülerinnen und Schüler und der Situation in den Schulen möglich ist, ist der Auftrag der Schulen. Mit Bezug auf die Bildungsstandards kann man die Einlösung der Anforderungen überprüfen. So lässt sich feststellen, inwieweit das Bildungssystem seinen Auftrag erfüllt hat, und die Schulen erhalten eine Rückmeldung über die Ergebnisse ihrer Arbeit.

Diesem Verständnis nach gehen in die Entwicklung von Bildungsstandards (a) gesellschaftliche und pädagogische Zielentscheidungen, (b) wissenschaftliche, insbesondere fachdidaktische und psychologische Aussagen zum Aufbau von Kompetenzen, sowie (c) Konzepte und Verfahren der Testentwicklung ein.

(a) *Bildungsstandards orientieren sich an Bildungszielen, denen schulisches Lernen folgen soll, und setzen diese in konkrete Anforderungen um.*

Bildungsziele sind relativ allgemein gehaltene Aussagen darüber, welche Wissensinhalte, Fähigkeiten und Fertigkeiten, aber auch Einstellungen, Werthaltungen, Interessen und Motive

die Schule vermitteln soll. In den Bildungszielen drückt sich aus, welche Chancen zur Entwicklung ihrer individuellen Persönlichkeit, zur Aneignung von kulturellen und wissenschaftlichen Traditionen, zur Bewältigung praktischer Lebensanforderungen und zur aktiven Teilnahme am gesellschaftlichen Leben wir Kindern und Jugendlichen geben wollen. Schulische Bildungsziele sollten zudem auf Nachhaltigkeit und Transfer ausgerichtet werden. Sie müssen anschlussfähig sein für lebenslanges Weiterlernen, für Anforderungen in Alltag, Beruf und Gesellschaft.

Bildungsziele formulieren somit Erwartungen an die Entwicklung jeder einzelnen Schülerin und jedes Schülers, und zugleich verpflichten sie die Gesellschaft und ihre Bildungseinrichtungen, entsprechende Entwicklungsmöglichkeiten zu schaffen. 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.

Mit Bildungszielen verknüpft sich meist auch ein bestimmtes Verständnis der Bedeutung, die ein Fach oder Lernbereich für die persönliche Entwicklung hat und worin seine gesellschaftliche Funktion besteht. Ist das Erlernen von Fremdsprachen auf kommunikatives Handeln ausgerichtet oder eher auf eine systematische Einführung in Sprache und Kultur? Besteht mathematische Bildung in der Kenntnis von Lösungsmustern und der Beherrschung von Verfahren, oder eher in der Fähigkeit zur Modellierung von Situationen? Welche Bedeutung haben das Verständnis von Grundkonzepten wie „Energie“ und „Stoffwechsel“ und das hypothesengeleitete Denken für eine naturwissenschaftliche Allgemeinbildung? Die Bestimmung von Bildungszielen fordert deshalb auch eine Verständigung darüber, was den Kern von Lernbereichen und Fächern ausmacht.

Bildungsziele geben allerdings nur recht generelle Erwartungen wieder. Damit sie pädagogisch umgesetzt werden können, benötigt man ein Medium, in dem sich die Ziele spezifizieren und definieren lassen. Dieses Medium sind in der Tradition die Lehrpläne, aktuell sollen sie um Kompetenzmodelle ergänzt werden.

*(b) Bildungsstandards konkretisieren die Ziele in Form von Kompetenzanforderungen. 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 Anforderun-*

gen in Kompetenzmodellen, die Aspekte, Abstufungen und Entwicklungsverläufe von Kompetenzen darstellen.

Nach Weinert (2001, S. 27f.) versteht man Kompetenzen als „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen<sup>1</sup> und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“.

Beispielsweise drückt sich die Kompetenz beim Erwerb einer Fremdsprache – wenn man kommunikative Handlungsfähigkeit als Bildungsziel vorgibt – darin aus, wie gut man kommunikative Situationen bewältigt, wie gut man Texte unterschiedlicher Art verstehen und selbst adressatengerecht Texte verfassen kann, aber unter anderem auch in der Fähigkeit, grammatische Strukturen korrekt aufzubauen und bei Bedarf zu korrigieren, oder in der Fähigkeit und Bereitschaft, sich offen und akzeptierend mit anderen Kulturen auseinander zu setzen. Standards für das Fremdsprachenlernen müssen diese Teilkompetenzen darstellen und jeweils verschiedene Niveaustufen unterscheiden (vgl. Anlage a). Hierbei spielen nicht nur kognitive Wissensinhalte eine Rolle; diese sind vielmehr – wie Weinert im obigen Zitat hervorhebt und das zuletzt genannte Beispiel der sog. Interkulturellen Kompetenz besonders deutlich macht – mit Einstellungen, Werten und Motiven verknüpft.

Mit dem Begriff „Kompetenzen“ ist ausgedrückt, dass die Bildungsstandards – anders als Lehrpläne und Rahmenrichtlinien – nicht auf Listen von Lehrstoffen und Lerninhalten zurückgreifen, um Bildungsziele zu konkretisieren. Es geht vielmehr darum, Grunddimensionen der Lernentwicklung in einem Gegenstandsbereich (einer „Domäne“, wie Wissenspsychologen sagen, einem Lernbereich oder einem Fach) zu identifizieren. Kompetenzen spiegeln die grundlegenden Handlungsanforderungen, denen Schülerinnen und Schüler in der Domäne ausgesetzt sind. Durch vielfältige, flexible und variable Nutzung und zunehmende Vernetzung von konkreten, bereichsbezogenen Kompetenzen können sich auch „Schlüsselkompetenzen“ entwickeln, aber der Erwerb von Kompetenzen muss – wie Weinert (2001) hervorhebt – beim systematischen Aufbau von „intelligentem Wissen“ in einer Domäne beginnen.<sup>2</sup>

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 Dimensio-

---

<sup>1</sup> Volition = willentliche Steuerung von Handlungen und Handlungsabsichten.

<sup>2</sup> Der hier verwendete Begriff von „Kompetenzen“ ist daher ausdrücklich abzugrenzen von den aus der Berufspädagogik stammenden und in der Öffentlichkeit viel gebrauchten Konzepten der Sach-, Methoden-, Sozial- und Personalkompetenz. Kompetenzen werden hier verstanden als Leistungsdispositionen in bestimmten Fächern oder „Domänen“ (vgl. Kapitel 6).

nen. 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.

Werden Anforderungen als Stufen eines Kompetenzmodells festgesetzt, so ist dies keinesfalls eine Entscheidung nach Gutdünken. Es handelt sich vielmehr um professionelle Entscheidungen, die sich an Bildungszielen orientieren (z.B. an einem Verständnis des Bildungsauftrags eines Faches), aber auch auf Erfahrungen und Traditionen einschlägiger Disziplinen Bezug nehmen. Vor allem die Fachdidaktik ist gefragt, wenn es festzulegen gilt, welche Anforderungen zumutbar und begründbar sind. Bildungsstandards tragen dazu bei, diese Entscheidungen transparent und damit nach wissenschaftlichen und professionellen Maßstäben prüfbar zu machen. Kompetenzmodelle machen Aussagen über die Dimensionen und Stufen von Kompetenzen, die prinzipiell mit Hilfe passender Aufgaben (siehe unter c) empirisch überprüft werden können. Die Kognitionswissenschaften und die pädagogisch-psychologische Diagnostik stellen hierfür vielfältige Modelle und Methoden zur Verfügung (NRC 2001).

Kompetenzmodelle sollten auch Aussagen darüber machen, in welchen Kontexten, bei welchen Altersstufen und unter welchen Einflüssen sich die einzelnen Kompetenzbereiche entwickeln. Nur so kann von der Schule erwartet werden, dass sie mit geeigneten Maßnahmen zur systematischen Kompetenzentwicklung, zum kumulativen Lernen beiträgt.

Kompetenzmodelle sind wissenschaftliche Konstrukte. Das Erreichen einer Kompetenzstufe sagt etwas darüber aus, welche Handlungen und mentalen Operationen mit hoher Wahrscheinlichkeit korrekt ausgeführt werden können. Für die Umsetzung im Lehrplan und im Unterricht wie auch für die Bewertung von Schülerleistungen braucht man aber konkrete Beispiele und „Operationalisierungen“, bis hin zu Messvorschriften in Gestalt von Testaufgaben.

*(c) Bildungsstandards als Ergebnisse von Lernprozessen werden konkretisiert in Aufgabenstellungen und schließlich Verfahren, mit denen das Kompetenzniveau, das Schülerinnen und Schüler tatsächlich erreicht haben, empirisch zuverlässig erfasst werden kann.*

Diese dritte Ebene ermöglicht, das Erreichen von Standards empirisch zu überprüfen. Ausgehend von den Kompetenzbeschreibungen werden Aufgaben entwickelt, die prüfen, ob eine Person das angestrebte Ergebnis oder Handlungspotential entwickelt hat. Entsprechende Testaufgaben können allerdings nicht einfach aus den Kompetenzbeschreibungen „abgeleitet“ werden. Sie müssen generiert und auf ihre Validität hin geprüft werden. Mit entsprechenden geeigneten Testverfahren (*assessments*) kann das erreichte Kompetenzniveau erfasst werden.

Alle drei Komponenten – Bildungsziele, Kompetenzmodelle und Aufgabenstellungen bzw. Testverfahren – werden benötigt, um Bildungsstandards für die Qualitätsentwicklung an Schulen zu nutzen.

- (a) Ohne Bezug auf allgemeine Bildungsziele wären Kompetenzanforderungen reine Willkür oder bloße Expertenmeinung. Erst die Orientierung an diesen Zielen legitimiert die Bestimmung von erwünschten Niveaustufen und die daraus resultierenden Testverfahren. Insbesondere sollten die Standards von einem Verständnis des Bildungsauftrags der jeweiligen Fächer ausgehen, das expliziert werden muss.
- (b) Kompetenzmodelle und -erwartungen sind unverzichtbar, um zwischen relativ abstrakten, verallgemeinerten Bildungszielen einerseits, konkreten Aufgabensammlungen andererseits zu vermitteln. Vor allem die Kompetenzstufen sind ein zentrales Hilfsmittel für die Konstruktion von Aufgaben, sei es zu Zwecken der Lernplanung oder für die Testentwicklung. Kompetenzmodelle geben Hinweise für die Aufgabenkonstruktion, indem sie kognitive Leistungen mit unterschiedlichem Schwierigkeitsniveau spezifizieren. Umgekehrt ermöglicht erst die Einordnung in ein Kompetenzmodell, zu verstehen, was das Lösen oder Nichtlösen einer Aufgabe bedeutet, und entsprechende Rückmeldungen zu geben. Modellvorstellungen, die den Aufbau von Kompetenzen über mehrere Stufen hinweg charakterisieren, sind demnach wichtige Orientierungen für die Unterrichtspraxis und die Bewertung von Lernergebnissen.
- (c) Ohne Ergebniskontrolle würden die Bildungsziele als statische Größen von der tatsächlichen Entwicklung der Schule abgekoppelt werden; sie wären – wie die Forschung zur Rezeption von Lehrplänen vielfach nachgewiesen hat – kaum handlungsrelevant und würden häufig noch nicht einmal zur Kenntnis genommen. Feedback über den erreichten Kompetenzstand ist daher 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.

## **2.2 Merkmale guter Bildungsstandards**

Es gibt durchaus verschiedene Wege, zu Bildungsstandards zu kommen, und das „Produkt“ kann recht unterschiedlich aussehen. Dies wird weiter unten an vielen Beispielen illustriert (vgl. Abschnitt 2.3, Kapitel 3 sowie den Anhang). Es gibt jedoch eine Reihe von Merkmalen, denen gute

Bildungsstandards genügen müssen, um allen Beteiligten in den Schulen die verbindlichen Ziele und Kompetenzanforderungen möglichst eindeutig zu vermitteln:

1. *Fachlichkeit*: Bildungsstandards sind jeweils auf einen bestimmten Lernbereich bezogen und arbeiten die Grundprinzipien der Disziplin bzw. des Unterrichtsfachs klar heraus.
2. *Fokussierung*: Die Standards decken nicht die gesamte Breite des Lernbereiches bzw. 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 Verlauf der Lerngeschichte aufgebaut worden sind. Damit zielen sie auf kumulatives, systematisch vernetztes Lernen.
4. *Verbindlichkeit für alle*: Sie drücken die *Mindestvoraussetzungen* aus, die von *allen* Lernern erwartet werden. Diese Mindeststandards müssen schulformübergreifend für alle Schülerinnen und Schüler gelten.
5. *Differenzierung*: Die Standards legen aber nicht nur eine „Meßlatte“ an, sondern differenzieren 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, einer Schulform darstellen.
6. *Verständlichkeit*: Die Bildungsstandards sind klar, knapp und nachvollziehbar formuliert.
7. *Realisierbarkeit*: Die Anforderungen stellen eine Herausforderung für die Lernenden und die Lehrenden dar, sind aber mit realistischem Aufwand erreichbar.

Diese Merkmale guter Bildungsstandards sollen im Folgenden näher diskutiert werden.

### **Merkmal 1: Fachlichkeit**

Unterrichtsfächer sind aus gutem Grund das Gerüst, das traditionell die Struktur der Lehr- und Lernaktivitäten in den Schulen bestimmt. Unterrichtsfächer korrespondieren mit wissenschaftlichen Disziplinen, die bestimmte Weltansichten (eine historische, literarisch-kulturelle, naturwissenschaftliche usw.) ausarbeiten und dabei bestimmte „Codes“ einführen (z.B. mathematische Modelle, hermeneutische Textinterpretationen). Die Abgrenzung einzelner Fächer innerhalb der übergreifenden Lernbereiche – beispielsweise die Differenzierung oder Integration der naturwissenschaftlichen Fächer oder das Verhältnis zwischen Politik, Geschichte und Geographie – ist immer wieder diskutiert worden und wird in den Ländern unterschiedlich behandelt, aber im Prin-

zip muss sich die Schule an der Systematik dieser Weltansichten orientieren, wenn sie anschlussfähig sein will an kulturelle Traditionen und an die Diskurse anderer Lebensbereiche.

Auch in den Bildungsstandards muss die Systematik, die lernbereichs- und fachbezogen entwickelt wurde, genutzt werden. Mehr noch: Die Bildungsstandards sollen die Kernideen der Fächer bzw. Fächergruppen besonders klar herausarbeiten, um Lehren und Lernen zu fokussieren. Zu diesen Kernideen gehören: die grundlegenden Begriffsvorstellungen (z.B. die Idee der Epochen in der Geschichte, das Konzept der Funktion in der Mathematik, die Vorstellung von situations- und adressatengerechten Textsorten in den philologischen Fächern), die damit verbundenen Denkoperationen und Verfahren und das ihnen zuzuordnende Grundlagenwissen.

Spezielle Bildungsstandards für „Schlüsselqualifikationen“ wie Lernfähigkeit, problemlösendes und kreatives Denken, Arbeitsorganisation und Kooperation sind daher nicht sinnvoll. Wohl aber sollten solche fächerübergreifenden Bildungsziele in den Standards der einzelnen Lernbereiche berücksichtigt werden. Beispielsweise enthalten Standards für den mathematischen und naturwissenschaftlichen Unterricht Aussagen darüber, wie die Lernenden auf bestimmten Kompetenzstufen mit offenen Problemstellungen umgehen können und inwieweit sie in der Lage sind, ihre Vorstellungen und Lösungswege anderen zu vermitteln.

## **Merkmal 2: Fokussierung**

Einer der größten Nachteile traditioneller Curricula besteht darin, dass sie einen allumfassenden Anspruch haben – auch und gerade, wenn sie den Lehrpersonen die Freiheit zur Auswahl lassen wollen. Ein Lehrplan und übrigens auch ein Lehrbuch, das alle Varianten des Unterrichts zulassen will, verliert eher an Orientierungskraft als sie zu steigern. Der Ruf nach „Verschlankung“ oder „Entrümpelung der Lehrpläne“ begleitet daher die Lehrplanreformen der Länder seit Jahren. Gerade die Bildungsstandards müssen auf zentrale Aspekte fokussiert sein. Sie legen fest, was für alle verbindlich ist, und lassen damit auch Freiraum zu Ausgestaltung und Ergänzung, ohne eine unübersichtliche und beliebige Breite anzubieten.

## **Merkmal 3: Kumulativität**

Bildungsstandards legen fest, welche Kompetenzen *bis zu* einem bestimmten Zeitpunkt *insgesamt* erworben sein müssen. Sie gehen damit über die Anforderungen an die Schülerinnen und Schüler in einem bestimmten Schuljahr oder gar einer einzelnen Unterrichtseinheit, auf die Leistungsbewertungen normalerweise beschränkt sind, hinaus. Es geht eben gerade nicht darum, dass Inhalte für eine Klassenarbeit gelernt und wieder vergessen werden, sondern dass übergreifende Kompetenzen in grundlegenden Bereichen aufgebaut werden und überprüfbar über ei-

nen längeren Zeitraum zur Verfügung stehen. Damit zielen Bildungsstandards auf das „kumulative Lernen“, bei dem Inhalte und Prozesse aufeinander aufbauen, systematisch vernetzt, immer wieder angewandt und aktiv gehalten werden.

Weinert (z.B. 2001) hat in seinen einflussreichen Publikationen immer wieder darauf hingewiesen, dass die Partialisierung von Lernerfahrungen in einzelne, nur wenig verknüpfte Abschnitte eines der wichtigsten Probleme beim schulischen Lernen darstellt. Die Testaufgaben der TIMS- und der PISA-Studie erforderten hingegen häufig Verknüpfungen und verlangten die Anwendung von Wissen und Können aus unterschiedlichsten Teilbereichen der Fächer. Baumert und andere (1997, 2000, 2001) haben in den schlechten Ergebnissen deutscher Schülerinnen und Schüler bei diesen Tests Belege für die mangelnde Kumulativität des schulischen Lernens gesehen.

#### **Merkmal 4: Verbindlichkeit für alle durch Mindeststandards**

Müssen Bildungsstandards eine Stufe festlegen, unter die kein Lernender zurückfallen soll („Mindeststandard“), eine mittlere Niveaustufe, die im Durchschnitt erreicht werden soll („Regelstandard“), oder ein Ideal („Maximalstandard“)? In dieser Frage sind prinzipiell unterschiedliche Lösungen denkbar. Hier wird jedoch nachdrücklich empfohlen, in den nationalen Bildungsstandards für Deutschland ein verbindliches Minimalniveau festzuschreiben. Die Kompetenzmodelle und die zugehörigen Aufgaben (Operationalisierungen) beschreiben dann eindeutig, welche Leistungen ein Schüler/eine Schülerin erbringen muss, um die Mindestanforderungen zu erfüllen. (Was nicht ausschließt, dass auch höhere Anforderungen ausgewiesen werden – als Ziele weiterführender Lernprozesse und zur Differenzierung zwischen Lernergruppen, s. Punkt 4 unten).

Diese Konzentration auf Mindeststandards ist für die Qualitätssicherung im Bildungswesen von entscheidender Bedeutung. Sie zielt darauf ab, dass gerade die Leistungsschwächeren nicht zurückgelassen werden. Jeder Schule, jedem Lehrenden und jedem Lernenden soll klar sein, welche Mindestanforderungen gestellt werden. Angesichts der Tatsache, dass unser Bildungssystem, verglichen mit den Systemen anderer Industriestaaten, Schwächen vor allem im unteren Leistungsbereich zeigt, kommt diesem Merkmal besondere Bedeutung zu. Die national und schulform-übergreifend verbindliche Formulierung von Bildungszielen und Mindestanforderungen – die dann Lehrplanarbeit, Unterrichts- und Förderkonzepte, Schulevaluation und anderes beeinflussen – kann einen entscheidenden Beitrag zum Abbau von Disparitäten in unserem Bildungssystem leisten.

„Regelstandards“, die ein Durchschnittsniveau spezifizieren, enthalten implizit die Botschaft, dass man eine Art Normalverteilung der Kompetenzen erwartet, bei der es im Vergleich zum Regelfall immer Gewinner und Verlierer gibt. Nicht ohne Grund ist man beispielsweise in Schweden von einem solchen Bewertungsmodell dazu übergegangen, für die nationalen Leistungsüberprüfungen ein Mindest-Bestehenskriterium anzugeben. (Und zusätzlich ein Kriterium für „mit Auszeich-

nung bestanden“). Die für die Stützung leistungsschwächerer Schüler entscheidende Frage, was diese wissen und können müssen, um als erfolgreich gelten zu können, lässt sich mit Regelstandards nicht beantworten – jedenfalls nicht positiv. „Maximalstandards“ führen noch stärker dazu, dass Anforderungen an untere Leistungsniveaus bloß negativ, durch den Grad der Abweichung vom Ideal beschrieben werden können.

Auch in der Alltagspraxis der Prüfung und Bewertung von Schülerleistungen wird in den Schulen sehr häufig ein Defizit-orientierter Ansatz gewählt: Wer die allen gemeinsam gestellten Aufgaben voll erfüllt, wird mit „sehr gut“ benotet; darunter liegende Notenstufen sind durch mehr oder weniger zahlreiche und unterschiedlich gravierende Fehler charakterisiert. Dieser Defizit-Orientierung sollten die Bildungsstandards jedoch durch eine positive Beschreibung von Kompetenzen, insbesondere eine positive Darstellung von Mindestanforderungen entgegenwirken.

### **Merkmal 5: Differenzierung**

Für die Qualitätsentwicklung insgesamt ist es selbstverständlich wichtig, über das Mindestkriterium hinaus höhere Anforderungen insbesondere an leistungsstärkere Schülerinnen und Schüler zu richten. Wünschenswert ist auch, dass einzelne Schulen unterschiedliche Profile entwickeln können, also beispielsweise einen Schwerpunkt im mathematisch-naturwissenschaftlichen Bereich legen und hier für sich besonders hohe Erwartungen formulieren. Kompetenzmodelle, die das Spektrum der Kompetenzen in einem Leistungsbereich differenziert nach Teildimensionen und Stufen abbilden, lassen solche Ergänzungen und Profile zu. Schulen können beispielsweise Ziele darüber formulieren, welcher Anteil ihrer Schülerinnen und Schüler die höchsten Kompetenzstufen erreichen sollen, und dies in Evaluationen überprüfen.

Allerdings halten die Autoren dieses Gutachtens es nicht für sinnvoll, solche Zusatzerwartungen in Bildungsstandards selbst, zumal auf nationaler Ebene, festzuschreiben. Standards legen die Grundstruktur der Kompetenzen fest, die vermittelt werden sollen, und fixieren ein verbindliches Minimum. Weitere Spezifikationen bleiben den Ländern und schließlich den einzelnen Schulen vorbehalten – den Institutionen also, die über die konkreten Curricula (Lerninhalte und deren zeitliche Sequenzierung) entscheiden und Abschlüsse vergeben (sei es in landesweit zentralen oder schulischen Abschlussprüfungen). Bildungsstandards sollen darauf angelegt sein, dass weitergehende Erwartungen und Profilierungen, Curricula und Zertifizierungsbestimmungen daran anschließen können; beispielsweise indem sie sich an den zentralen Prinzipien des Lernbereichs und den Kompetenzmodellen orientieren. Die Bildungsstandards selbst beinhalten aber solche Spezifikationen nicht.

Die gemeinsame Sprache der Bildungsstandards, die national verbindliche Formulierung von Bildungszielen und Mindestanforderungen wäre ein großer Fortschritt für die umfassende Sicherung

der Bildungsqualität. Durch ergänzende Profile würde aber auch der Tatsache Rechnung getragen, dass das Schulsystem in Deutschland föderal organisiert und in Schulformen gegliedert ist.

### **Merkmal 6: Verständlichkeit**

Bildungsstandards werden nur dann einen Fortschritt für die Steuerung unseres Schulsystems darstellen, wenn sie eine klarere, verständlichere Orientierung vermitteln als die meisten der herkömmlichen Lehrpläne und Richtlinien. Auch die Gestaltung der Standards muss dafür sorgen, dass sie von Lehrenden ebenso wie in der Öffentlichkeit rezipiert und angenommen werden. Dass dies nicht einfach zu erreichen ist, zeigt beispielsweise die Erfahrung in Großbritannien, wo aus einem national verbindlichen Kerncurriculum schrittweise ein detaillierter Lehrplan geworden ist, der inzwischen von manchen als hinderlich angesehen wird. Überzeugend sind die nationalen Vorgaben (die unterschiedlich benannt sind) in den skandinavischen Staaten und in einigen kanadischen Provinzen, die den Bildungsauftrag der Schulen sehr knapp und präzise zeichnen.

### **Merkmal 7: Realisierbarkeit**

Bildungsstandards sollten für Schüler und Schülerinnen unter den gegebenen schulischen Voraussetzungen realistisch erreichbare Ziele umfassen, und es sollte für den Lehrer möglich sein, die formulierten Ziele in angemessene Unterrichtsaktivitäten zu übersetzen (vgl. auch NRC 1999).

Einer der interessantesten Befunde aus den nationalen Zusatzerhebungen bei TIMSS und PISA war die Erkenntnis, dass Lehrer, Fachdidaktiker und Lehrplanexperten sehr häufig die Wahrscheinlichkeit richtiger Lösungen von Aufgaben überschätzten – am häufigsten in jenen Bereichen, die besonders typisch sind für die Aufgabenstellungen des alltäglichen Fachunterrichts. Man wird also gerade beim Beginn der Entwicklung von Bildungsstandards in Deutschland damit rechnen müssen, dass Anforderungen – seien es Minimalerwartungen oder erweiterte Profile – unrealistisch hoch angesetzt werden. Sicherlich sollen Bildungsstandards eine Herausforderung für Lehrende und Lernende darstellen, aber unrealistisch hohe Erwartungen führen zu Demotivation und gefährden die Akzeptanz der Standards. Man wird daher die Höhe der Kompetenzanforderungen erst nach empirischen Befunden definitiv festlegen können.

Beispiele für unrealistische Erwartungshaltungen finden sich in dem „No child left behind“-Gesetz, das im Frühjahr 2002 in den USA verabschiedet wurde, wie auch in den bildungspolitischen *benchmarks*, die im November 2002 von der Europäischen Kommission vorgeschlagen wurden. Das US-Gesetz sieht vor, dass alle US-Bundesstaaten Standards und zugehörige Testverfahren entwickeln und – jeweils nach eigenen Kriterien – Mindesterwartungen in Grundkompetenzen,

vor allem in der Lesefähigkeit, festsetzen. Das Gesetz erwartet sodann von den einzelnen Schulen, von Jahr zu Jahr den Anteil der Schülerinnen und Schüler, die das Mindestniveau verfehlen, kontinuierlich zu senken. Wissenschaftler haben darauf aufmerksam gemacht, dass derartig geradlinige Entwicklungen auf Schulebene über 10 Jahre hinweg höchst unwahrscheinlich sind. Die Europäische Kommission erwartet von allen EU-Staaten, den Anteil der 15-jährigen, die im PISA-Test zur Lesekompetenz Stufe I nicht überschreiten, bis zum Jahr 2010 zu halbieren. Dies würde in Deutschland einen Leistungsanstieg der schwächeren Schüler bedingen, der mindestens dem Lernzuwachs eines ganzen Schuljahres entspricht. Ein so starker Zuwachs bei Schülerinnen und Schülern, die sich gegenwärtig bereits in der Grundschule befinden, wird selbst mit massiven Reformanstrengungen kaum möglich sein.

Bildungspolitik und Bildungsadministration sind also gut beraten, realistische Ziele zu formulieren, wenn die Output-orientierte Steuerung erfolgreich Veränderungsprozesse anstoßen soll.

### **2.3 Varianten des Begriffs „(Bildungs-) Standard“ in der internationalen Diskussion**

In der Praxis verschiedener Staaten und in der Fachdiskussion sind ganz unterschiedliche Arten von Bildungsstandards entwickelt worden. Im vorliegenden Gutachten legen wir, wie im vorigen Abschnitt erläutert, den Schwerpunkt auf die Klärung pädagogischer Aufgaben der Schule vermittelt konkreter Anforderungen an die Kompetenzen der Schülerinnen und Schüler.

Bis vor kurzem wurde der Begriff „Standard“ in Deutschland im Bildungsbereich kaum gebraucht. In den anglo-amerikanischen Staaten, aber auch etwa in den Niederlanden, ist der Begriff seit langem eingebürgert, allerdings mit vielen unterschiedlichen Bedeutungen. Um Missverständnisse insbesondere im Bezug auf die internationale Diskussion zu vermeiden, soll daher an dieser Stelle ein kurzer Überblick gegeben werden.

Eingeführt wurde der Begriff „Standard“ (eigentlich: „Fahne des König“, dann: „Norm, Maßstab“, vgl. Kluge 1999, S. 787) im Bildungsbereich erstmals um 1860 in England, im Zusammenhang mit einem Gesetz, das die erfolgsabhängige Budgetierung von Elementarschulen vorsah. Es wurden je sechs Leistungsstufen für Lesen, Schreiben und Rechnen festgelegt, nach denen die Schulinspektoren Schüler einzustufen hatten, und diese wurden als „Standard I - VI“ bezeichnet (Aldrich 2000).

Das System der erfolgsabhängigen Budgetierung wurde zwar in England Anfang des 20. Jahrhunderts wieder abgeschafft, aber bis heute spielt der Begriff in der Bildungspolitik in Großbritannien eine zentrale Rolle (vgl. Goldstein & Heath 2000). Seit Ende der 1980er Jahre, im Zuge der von der Thatcher-Regierung eingeleiteten und von new labour weitergeführten output-orientierten Bildungsreform, wird dort an Nationalen Curricula und Schulleistungsmessungen (assessments) gearbeitet. Die zentrale Behörde der Schulinspektoren trägt den Namen „Office for Standards in Education (OFSTED)“. Eines der wichtigsten Ziele der englischen Bildungspolitik ist es, „Standards anzuheben“. Damit ist im Allgemeinen gemeint, dass sich das tatsächliche Leistungsniveau an den Schulen verbessern sollte. Standards werden also – anders als hier vorgeschlagen – nicht als Normen, sondern als das faktisch (im Durchschnitt) von Schülerinnen und Schülern erreichte Leistungsniveau interpretiert.

Dieser auf faktische Leistungen, nicht auf normative Setzungen ausgerichtete Standard-Begriff wurde auch von der OECD in ihrer sehr einflussreichen Schrift „Schulen und Qualität“ (OECD 1991, Original 1989) verwendet, wengleich nicht unerwähnt blieb, dass der Begriff vielschichtig und mehrdeutig sei. Es wurde schon Ende der 80er Jahre ein international recht einheitlicher Trend zu einer an Lernergebnissen orientierten Steuerung des Schulsystems ausgemacht; zugleich wurde die zentrale Bedeutung der Einzelschule als Ort der Qualitätsentwicklung betont.

Entsprechende Reformen sind in der Tat in den 80er und 90er Jahren des 20. Jahrhunderts in vielen Staaten zu verzeichnen gewesen. Das Konzept der Standards spielt dabei eine wichtige Rolle. Dabei hat sich – bis heute mit der Ausnahme Großbritanniens – eine normbezogene Interpretation des Begriffs durchgesetzt. Standards werden inzwischen international als normative Vorgaben für die Steuerung von Bildungssystemen verstanden.

Im Einzelnen finden sich unterschiedliche Varianten, die nach drei Leitfragen unterschieden werden können:

1. Beziehen sich Standards auf den Input und die Prozesse des schulischen Lernens (also z.B. die Ausstattung von Schulen, die Konzeption des Unterrichts, die Qualifikation der Lehrkräfte usw.) oder auf den Output, d.h. vor allem die Lernergebnisse? Im ersten Fall, wenn Lerngelegenheiten für Kinder und Jugendliche festgelegt werden, spricht man im Amerikanischen von opportunity-to-learn standards.

International sind meistens Ergebnis-Normen gemeint, wenn von Standards die Rede ist. Innerhalb dieses Typs gibt es wiederum Unterscheidungen:

2. Auf welcher Abstraktionsebene werden Standards formuliert: Auf der Ebene allgemeiner Bildungsziele und Wertvorstellungen? Auf der Ebene mehr oder weniger bereichsspezifischer Kompetenzen? Auf der Ebene von Lernzielen und -inhalten (content standards)? Oder auf der Ebene von Testwerten, wo der Standard durch einen Grenzpunkt (cut-off point) auf der Testwerteskala markiert werden kann (performance standards)?
3. Welches Zielniveau wird spezifiziert: Ein basales Niveau, das von (fast) allen Schülerinnen und Schülern erreicht werden sollte (Mindeststandards), ein Ideal (Maximalstandards) oder ein eher durchschnittliches Erwartungsniveau? (Im letzteren Fall sprechen wir im Folgenden von Regelstandards.)

Die Expertengruppe hat sich in allen drei Fragen im Rahmen dieses Gutachtens festgelegt: Bildungsstandards in unserem Sinne sind ergebnisbezogen. Sie zielen auf bereichsspezifische Kompetenzen, sind aber gleichsam nach „oben“ und „unten“ anschlussfähig: Sie orientieren sich an allgemeinen Bildungszielen, und sie sind prinzipiell umsetzbar (operationalisierbar) in Aufgaben und Testskalen. Außerdem wird empfohlen, im Rahmen von Kompetenzmodellen verschiedene Stufungen von Kompetenzen zu unterscheiden, darunter aber eine bestimmte Kompetenzstufe als Mindeststandard auszuweisen.

Andere Staaten haben in diesen Fragen unterschiedliche Festlegungen getroffen. Mitunter existieren auch mehrere Arten von Standards nebeneinander, wobei die Begriffe wechseln. So ist das „nationale Curriculum“ in Schweden ein auf Bildungsziele von mittlerem Abstraktionsgrad abzielendes Dokument, das man in unserem Sinne als Bildungsstandard interpretieren kann; zusätzlich werden in Schweden im Rahmen nationaler Schulleistungsuntersuchungen konkrete Testleistungsparameter festgelegt, die einen Mindeststandard (im Englischen „pass“) bzw. ein gehobenes Leistungsniveau („pass with distinction“) kennzeichnen.

In den USA wurde traditionell der Begriff Standard wesentlich synonym zu *Performance Standard* benutzt (vgl. McLeod, Stake, Schapelle, Mellissinos & Gierl 1996), und die Messung von Leistung stand im Vordergrund. Dabei waren es insbesondere Basisfähigkeiten, auf die Bezug genommen wurde. Mit der Publikation der sehr einflussreichen *Curriculum and Evaluation Standards for School Mathematics* der nationalen Mathematiklehrervereinigung NCTM im Jahr 1989 (vgl. Kapitel 3) änderte sich dies – nicht nur für die Mathematik, da die NCTM-Standards zum Vorbild auch für Arbeiten in anderen Fächern wurden (Ravitch 1995). Dem traditionellen, auf Basisfähigkeiten konzentrierten Ansatz stellten die Autoren der NCTM bewusst eine stärker prozessorientierte Sichtweise auf das Mathematiklernen gegenüber. Das Entdecken und Erforschen von mathematischen Verbindungen und Beziehungen sollte ein verständnisbasiertes Lernen ermöglichen. Die NCTM-Standards beinhalten keine Testanforderungen (bis heute hat NCTM keine test-basierten

Standards im klassischen Sinne vorgelegt), sondern sie beinhalten eher eine Vision des guten Mathematikunterrichts, sind also Standards für professionelles Handeln von Mathematiklehrern (opportunity to learn-standards). Zugleich legen sie Inhaltsdimensionen des Mathematikunterrichts fest, sind also auch content standards.

Hinter der Verwendung bestimmter Konzeptionen von Standards stehen meist unterschiedliche kulturelle und politische Traditionen und unterschiedliche Steuerungsmodelle. Dies soll im Folgenden am Beispiel der USA und Skandinaviens gezeigt werden.

In den USA war und ist Schule und Unterricht weniger durch nationale Richtlinien bzw. Regelungen der einzelnen Bundesstaaten bestimmt, vielmehr spielt die Autonomie von Bezirken und Schulen beim Entwurf und der Implementation von Curricula eine wichtige Rolle. Die Standards der NCTM zielen vor diesem Hintergrund auf eine Verbesserung der Qualität des Unterrichts in der Breite durch einheitliche Normen in Bezug auf die Inhalte, aber auch im Hinblick auf Lehrmethoden und Leistungsbewertung. Die hohe Akzeptanz der NCTM-Standards hängt auch damit zusammen, dass sie im Kontext eines gesellschaftlichen Prozesses entstanden sind, in dem standardbasierte Reformen in allen Bereichen des Unterrichts für notwendig erachtet wurden. Im Blickpunkt stand dabei die übergroße Diversivität von Bildung in einem stark dezentral bzw. lokal gesteuerten System, dessen Output beispielsweise in Form von Ergebnissen internationaler Vergleichsstudien als unzureichend angesehen wurde. Die Reformen zielten daher auf die Erfüllung hoher Standards durch alle Schülerinnen und Schüler, die auf der Grundlage interessanter und anspruchsvoller Inhalte erreicht werden sowie die Anwendung theoretischen Wissens auf praktische und lebensnahe Probleme umfassen sollten (McLaughlin & Shepard 1995).

Eine Studie von Swanson und Stevenson (2002) stellt auf der Grundlage von 22 Indikatoren dar, in welchem Maß die Bundesstaaten der USA bis zum Jahr 1996 reformbasierte Standards eingeleitet bzw. umgesetzt haben. Als Indikatoren werden die Anpassung von Curricula, die Beschreibung von Kompetenzstufen, die Einführung von einheitlichen Testverfahren und die Professionalität von Lehrerinnen und Lehrern verwendet. Zunächst zeigt sich, dass die auf den Mathematikunterricht bezogenen Indikatoren insgesamt am besten umgesetzt worden sind. Darüber hinaus belegt die Studie große Unterschiede zwischen einzelnen Bundesstaaten. Maryland, Alabama und Kentucky liegen im Hinblick auf den Umfang der Aktivitäten und Produkte deutlich an der Spitze. Betrachtet man die *Standards* dieser Staaten im Detail, dann fällt allerdings gerade bei ihnen die Orientierung an Testperformanz sowie die Reduktion des Curriculums auf einen minimalen Katalog von Inhalten auf. Die Diskussion standardbasierter Curricula hat also in ihrer konkreten Umsetzung in manchen Bundesstaaten der USA eher zu stark restringierten Standards geführt, die sich dann in Bezug auf den Mathematikunterricht von den *Standards* der NCTM klar wegbewegen. Insbesondere ist die ursprüngliche Intention der Verbindung von anspruchsvollen Inhalten und lebensnahen Problemen dann nicht mehr zu finden.

Die Grundidee des nationalen Curriculums in Schweden lag von Beginn an auf einer anderen Ebene. Hier ging es gerade darum, die Autonomie der einzelnen Schule zu stärken und auf die Individualität von Schülerinnen und Schülern einzugehen. Wichtig waren dabei die Berücksichtigung des kulturellen Erbes, der Internationalisierung von Gesellschaft und Bildungssystem, des technologischen und industriellen Wandels und der Umweltbedingungen. Diese Auffassung ist nicht auf Schweden beschränkt, sondern trifft auch für andere nordeuropäische Staaten zu. So lag in Finnland der angestrebten Änderung des Bildungssystems eine Orientierung an ähnlichen Aspekten zugrunde. Flexibilität und individuelle Wahlmöglichkeiten vor dem Hintergrund einer wachsenden Internationalisierung wurden zu wesentlichen Eckpfeilern der Reform. Gleichzeitig wurde eine Dezentralisierung von Entscheidungsprozessen in Bezug auf curriculare Aspekte angestrebt. Diese Tendenz wird allerdings in letzter Zeit ausgeglichen, indem Kerncurricula, Basiskompetenzen und entsprechende Kompetenzniveaus landesweit definiert werden (vgl. Linnakylä, in Vorbereitung).

Die Beispiele zeigen, dass Standards auf völlig unterschiedlichen theoretischen Annahmen, Wertvorstellungen und gesellschaftlichen Zielen basieren können. Diese Annahmen, Werte und Ziele bestimmen aber ganz wesentlich die Akzeptanz von Standards, ihre Umsetzung in der Schule und die Möglichkeiten der täglichen Arbeit mit diesen Standards.

### 3. Beispiele: Standards und Curricula aus dem Bereich der Mathematik

Die Diskussion über Standards einer mathematischen Grundbildung hat in vielen Staaten bereits wesentlich früher als in Deutschland eingesetzt. Dabei stehen weniger abgegrenzte Lerninhalte im Vordergrund als vielmehr Kompetenzen, die von den Schülerinnen und Schülern erreicht werden sollen. Inzwischen gibt es in mehreren Staaten elaborierte Beispiele für solche kompetenzorientierte Standards, und es liegen Erfahrungen mit ihrer konkreten Umsetzung vor. Der Bereich der Mathematik ist daher sehr gut geeignet, um an dieser Stelle das Konzept der Bildungsstandards konkret zu veranschaulichen.

Zu beachten ist allerdings, dass die Terminologie zur Bezeichnung von Standards und Lehrplänen in verschiedenen Staaten unterschiedlich ausgelegt wird (vgl. oben, Abschnitt 2.3). Was hier, im Kontext unserer Expertise unter den Begriff „Bildungsstandard“ gefasst wird, trägt in anderen Staaten die Bezeichnung „Standard“, „nationales Curriculum“ oder andere Namen. Wichtig ist, ob und wie die zentralen Komponenten ausgefüllt sind: (a) Orientierung an Bildungszielen, (b) Kompetenzmodelle, (c) konkrete, durch Aufgabenstellungen und Testverfahren operationalisierte Anforderungen.

Wir beginnen in Abschnitt 3.1 mit den äußerst einflussreichen *Principles and Standards* des amerikanischen Verbandes der Mathematiklehrer und -didaktiker, NCTM. Anhand zweier kanadischer Provinzen beschreiben wir in Abschnitt 3.2, wie die Ideen der NCTM in Lehrpläne (Alberta) bzw. konkrete Leistungserwartungen (Ontario) übertragen wurden. Abschnitt 3.3 diskutiert schließlich die Entwicklungen in Europa, Abschnitt 3.4 die Tradition der Lehrplanarbeit in Deutschland. Auf den Beitrag der PISA-Studie zur Entwicklung von Kompetenzmodellen gehen wir an anderer Stelle (in Abschnitt 6.2) ein.

#### **3.1 Die Principles and Standards der NCTM als Rahmenkonzept eines modernen Mathematikunterrichts**

Neuere Entwicklungen in Bezug auf den Mathematikunterricht versuchen, mathematische Grundbildung über die Schulstufen hinweg übergreifend zu beschreiben. Das international bekannteste und einflussreichste Beispiel sind die im Jahr 2000 vorgelegten *Principles and Standards* der *National Council of Teachers of Mathematics* (NCTM). Die NCTM ist der wesentliche, mathematikdidaktikbezogene Berufsverband in den USA, dem sowohl Mathematiklehrerinnen und -lehrer als auch Wissenschaftlerinnen und Wissenschaftler aus der Mathematikdidaktik angehören. Die

nach einer breiten Diskussion in der Fachöffentlichkeit, mit Vorarbeiten ab Mitte der 1980er Jahre zusammengestellten *Principles and Standards* stellen vielfach beachtete Rahmenbedingungen für den Unterricht in allen Klassenstufen dar, umfassen also alle Stufen vom Beginn der Primarstufe bzw. der Vorklasse bis hin zum Ende der Sekundarstufe. Ihr Ziel ist es, eine Leitlinie für die Verbesserung des Mathematikunterrichts im Sinne einer Umorientierung hin zu einem verständnisbasierten und problembezogenen Unterricht zu geben. Entsprechend können sie im Grunde als Vision einer idealen Praxis betrachtet werden (Tate, in Vorbereitung). Sie enthalten fachdidaktische, pädagogische und lernpsychologische Kernideen und basieren damit auf einer konkreten Theorie des Lehrens und Lernens.

Die *Principles* sind fachübergreifende Leitlinien guten Unterrichts, die inhaltsspezifisch ausformuliert sind. Sie können also in der oben (Abschnitt 2.3) vorgestellten Terminologie als eine Art *opportunity-to-learn standard* angesehen werden. Die eigentlichen *Standards* hingegen sind typische Beispiele für *content standards*. Sie haben einen stärker fachspezifischen Charakter und betreffen sowohl konkrete Lerninhalte als auch die Ebene der intendierten Prozesse und der anwendbaren Methoden. *Principles* und *Standards* eines zeitgemäßen Mathematikunterrichts bilden dabei in allen Bereichen Einheiten und sind durch umfassende wechselseitige Beziehungen geprägt.

Zu den *Principles* gehören folgende allgemeine und grundlegende Unterrichtsmerkmale:

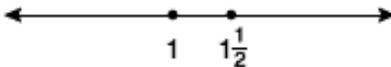
- das *Chancengleichheitsprinzip* („Equity“), über das ein Mathematikunterricht für alle Schülerinnen und Schüler eingefordert wird;
- das *Curriculumprinzip*, welches zum Inhalt hat, dass im Mathematikunterricht fachlich bedeutende Inhalte, die individuell und gesellschaftlich interessant sind, in kohärenter und umfassender Form vermittelt werden;
- das *Lehrprinzip*, mit dem die Notwendigkeit betont wird, Mathematikunterricht von mathematisch, mathematikdidaktisch und pädagogisch hochqualifizierten Personen erteilen zu lassen;
- das *Lernprinzip*, das die Bedeutung eines sinnstiftenden Mathematikunterrichts betont;
- das *Bewertungsprinzip* („Assessment“), das eine ständige Kontrolle des Leistungsstands ermöglicht; eine Bewertung muss allerdings damit vereinbar sein, dass der Mathematikunterricht auf Verstehensprozesse und nicht primär auf Faktenwissen abzielt;
- das *Technologieprinzip*, welches die verantwortungsvolle Nutzung von Technik durch die Schülerinnen und Schüler betont.

Die *Standards* konkretisieren, „what instruction should enable all students to know and be able to do“ (S. 29). Sie gliedern sich in zwei Teile, nämlich einen inhaltlichen Teil, der gängige Themenbereiche des Mathematikunterrichts nennt, und in einen eher methodenorientierten bzw. prozessorientierten Teil, in dem wesentliche Arbeitsfelder des Mathematikunterrichts zusammengestellt sind. Inhaltliche Themenbereiche sind (1) Zahlen und Operationen, (2) Muster, Funktionen und Algebra, (3) Geometrie und Raumorientierung, (4) Messen und (5) Datenanalyse, Statistik und Wahrscheinlichkeit. Im methodenorientierten Teil werden als wesentliche Aspekte Problemlösen, Argumentieren und Beweisen, Kommunikation, Verbindungen und Darstellungen genannt. Der inhaltliche Teil beschreibt entsprechend, was ein Schüler oder eine Schülerin lernen sollte, der prozessorientierte Teil nennt Wege, wie inhaltliches Wissen erworben und angewendet werden kann. Die Aufstellung dieser verschiedenen Bereiche unter dem einen Begriff *Standards* zeigt, dass mathematische Inhalte und fachbezogene Arbeitstechniken eine unterrichtliche Einheit bilden. Wesentlich ist darüber hinaus, dass die *Standards* nicht für eine bestimmte Klassenstufe allein formuliert sind. Vielmehr handelt es sich um Inhalte und Methoden, die den gesamten Mathematikunterricht vom Beginn an bis zum Abschluss der Sekundarstufe II prägen sollten. Die Publikation der NCTM enthält zahlreiche Beispiele, die diesen Grundgedanken für die verschiedenen Klassenstufen konkretisieren. Dabei wird immer auf konkrete Bezüge zwischen den einzelnen Klassenstufen geachtet.

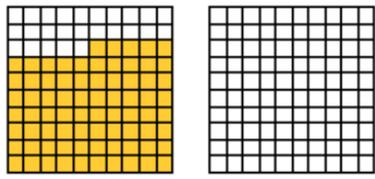
Die *Standards* können im Sinne dieser Expertise als eine Art Kompetenzmodell verstanden werden. Die fünf Inhaltsaspekte und die fünf Prozessaspekte beschreiben insgesamt, was mathematisches Denken und Arbeiten ausmacht. Die konsequente Strukturierung aller zehn Aspekte über alle Jahrgangsstufen vom Vorschulunterricht bis zur Ende der High School hinweg („K-12“) beinhaltet ein Konzept des systematischen, kumulativen Lernens für die gesamte Schullaufbahn.

Die *Principles and Standards* sind von der NCTM nicht in Testverfahren und *performance standards* umgesetzt, auch wenn Empfehlungen zur Gestaltung von Leistungsmessungen gegeben werden. Die umfangreiche Veröffentlichung enthält aber zu allen Bereichen Beispielaufgaben, durch die Ziele, Inhalte und Methoden verdeutlicht werden. Die Aufgabe in Abbildung 1 soll etwa den Bereich „Zahlenverständnis, Darstellung von Zahlen, Beziehungen zwischen Zahlen und Zahlensysteme“ veranschaulichen. Sie illustriert, dass offene Aufgaben, multiple Lösungswege und Begründungen zu den typischen Anforderungen gehören.

a. If  is  $\frac{3}{4}$ , draw the fraction strip for  $\frac{1}{2}$ , for  $\frac{2}{3}$ , for  $\frac{4}{3}$ , and for  $\frac{3}{2}$ . Be prepared to justify your answers.

b. 

Using the points you are given on the number line above, locate  $\frac{1}{2}$ ,  $2\frac{1}{2}$ , and  $\frac{1}{4}$ . Be prepared to justify your answers.

c. 

Use the drawing above to justify in as many different ways as you can that 75% of the square is equal to  $\frac{3}{4}$  of the square. You may reposition the shaded squares if you wish.

Abb. 1: Understand numbers, ways of representing numbers, relationships among numbers, and number systems (NCTM 2000)

### 3.2 Anwendungen in Kanada

Das Curriculum der kanadischen Provinz Alberta aus dem Jahr 1997 bezieht sich explizit auf die NCTM-Standards. Die grundlegende Struktur des Curriculums ist durch mathematische Inhaltsbereiche (z.B. Zahlen, Operationen mit Zahlen, Messen) geprägt, die einzeln für die verschiedenen Klassenstufen beschrieben und jeweils mit mathematischen Prozessen (z.B. Kommunikation über Mathematik, Verbindungen zwischen mathematischen Begriffen, Schätzen und Kopfrechnen) verbunden werden. Auf diese Weise wird einerseits deutlich, wie die Inhaltsbereiche über die verschiedenen Schuljahre hinweg entwickelt werden, andererseits werden Inhalte und Prozesse ständig in Verbindung miteinander gesehen.

Auch hier fällt – in klarer Anlehnung an die *Principles and Standards* der NCTM – die Konstanz des allgemeinen Rahmens über alle Klassenstufen hinweg auf. Durch alle Jahrgänge hindurch ziehen sich die *Inhaltsbereiche* Zahlen, Muster und Relationen, Form und Raum sowie Statistik und Wahrscheinlichkeit, die als formales Gerüst der Mathematik angesehen werden. Diese Inhaltsbereiche werden von mathematischen *Prozessen* und mathematischen *Grundideen* („*Nature of Mathematics*“) beeinflusst (vgl. Abb. 2). Die *Prozesse* umfassen dabei ganz unterschiedliche Aspekte wie etwa die Kommunikation über Mathematik, die Identifikation von Verbindungen zwischen verschiedenen mathematischen Begriffen, die Betrachtung von Zusammenhängen zwischen Mathematik und Alltagswelt, das Schätzen von Ergebnissen, Kopfrechnen, die Sicht auf

die Mathematik als Problemlösen, mathematisches Argumentieren und Begründen, den angemessenen Gebrauch von Computern und Taschenrechnern und die Benutzung geeigneter Visualisierungen. Mathematische Prozesse spielen im Duktus des Curriculums eine wesentliche Rolle, um konkrete Unterrichtsziele zu erreichen sowie lebenslanges Lernen zu fördern. Dabei sind viele dieser Prozesse in ihrer Definition den *Principles and Standards* der NCTM entnommen. Zu den mathematischen *Grundideen* gehören schließlich Begriffe wie Änderung, Konstanz bzw. Invarianz, Dimension, Zahl, Muster, Quantität, Relation, Form und Unsicherheit.

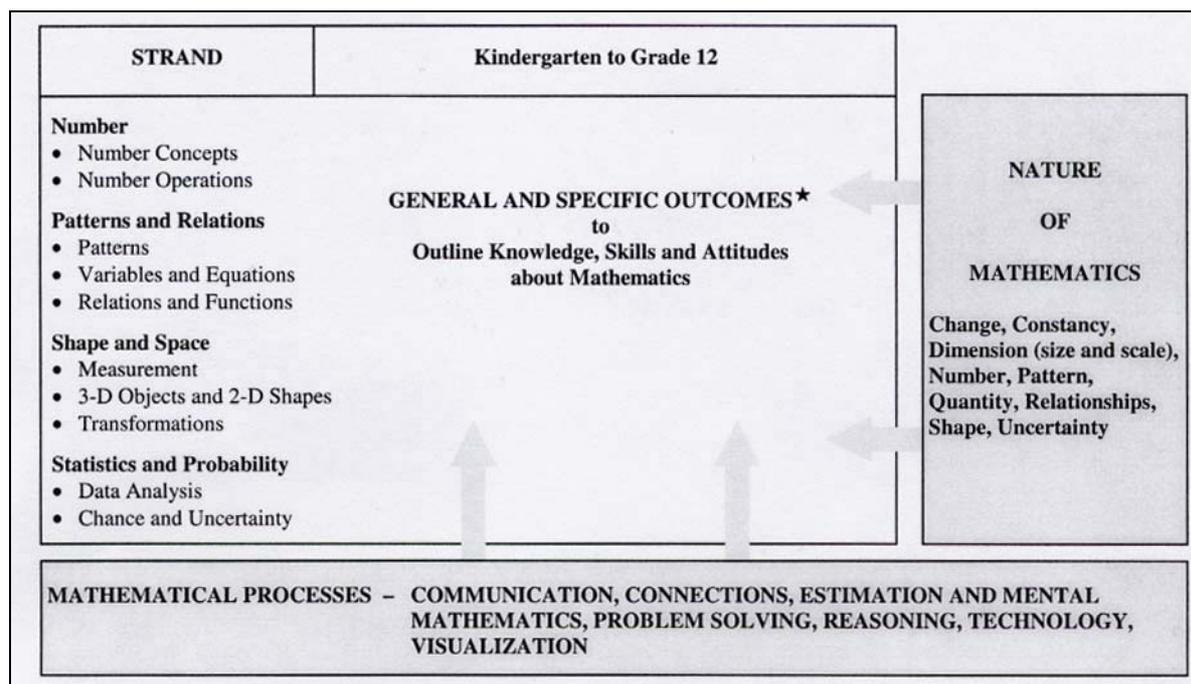


Abb. 2: Alberta Mathematics Program of Studies, Kindergarten to Grade 12  
([http://www.learning.gov.ab.ca/k\\_12/curriculum/bySubject/math/](http://www.learning.gov.ab.ca/k_12/curriculum/bySubject/math/))

Dieses Curriculum gibt konkrete Ziele mathematischer Bildung an. Sie werden einerseits in Form allgemeiner Ergebnisse des Unterrichts formuliert, und sie beschreiben, was Schülerinnen und Schüler am Ende einer bestimmten Klassenstufe wissen und können sollten. Damit umschließt das Curriculum von Alberta Bildungsstandards und ein Kompetenzmodell im Sinne dieser Expertise. Es geht allerdings darüber hinaus, indem es im Sinne eines Lehrplans auch konkrete Zeitvorgaben für Unterrichtsaktivitäten enthält.

Ganz ähnlich, aber in Bezug auf die Operationalisierung spezifischer Anforderungen noch pointierter, ist das Curriculum der Provinz Ontario aus dem Jahr 1997 formuliert. Die wesentlichen Elemente sind dabei Erwartungen an die Schülerinnen und Schüler und ihr Leistungsniveau. Die Erwartungen beschreiben Wissen, das die Schüler entwickeln sollen, und entsprechende Fertigkeiten und Fähigkeiten, die sich auf die Arbeit in der Klasse, auf Tests und auf andere Bewer-

tungsverfahren beziehen. Das Leistungsniveau ist in vier Stufen („Achievement Levels“) aufgeteilt und wird im Hinblick auf die Kategorien Problemlösen, Begriffsverständnis, Anwendung mathematischer Prozeduren und Kommunikation von Wissen betrachtet. Die Niveaus für das Problemlösen zeigt als Beispiel Abbildung 3.

| <b>ACHIEVEMENT LEVELS: MATHEMATICS, GRADES 1-8</b> |                                                                                  |                                                                           |                                                                              |                                                                                                 |
|----------------------------------------------------|----------------------------------------------------------------------------------|---------------------------------------------------------------------------|------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| <b>Knowledge/<br/>Skills</b>                       | <b>Level 1</b>                                                                   | <b>Level 2</b>                                                            | <b>Level 3</b>                                                               | <b>Level 4</b>                                                                                  |
| <b>Problem solving</b>                             | <b>The student solves problems:</b>                                              |                                                                           |                                                                              |                                                                                                 |
|                                                    | with assistance with a limited range of appropriate strategies rarely accurately | with limited assistance with appropriate strategies frequently accurately | independently by choosing the most appropriate strategies usually accurately | independently by modifying known strategies or creating new strategies almost always accurately |

Abb. 3: Ontario Achievement Levels Mathematics  
(<http://www.edu.gov.on.ca/eng/document/curriculum/curr97ma/achievem.html>)

### 3.3 Mathematische Curricula und Standards in Europa

In den letzten Jahren hat es auch in verschiedenen europäischen Staaten Initiativen gegeben, kompetenzorientierte Standards für das Schulsystem festzulegen. Zwei recht unterschiedliche Beispiele sind die nationalen Curricula in Schweden und in England. Das schwedische Curriculum ist seit 1994 verbindlich. Das englische Curriculum geht auf eine Debatte um einheitliche Leistungsstandards in den frühen 80er Jahren zurück und wurde 1988 erstmals implementiert. Seitdem hat es verschiedene Überarbeitungen und Ergänzungen erfahren.

In Bezug auf Schweden kann man im Wesentlichen zwei Ebenen der Zielbestimmung unterscheiden, nämlich die Ebene der allgemeinen, den Unterricht bestimmenden Werte und Normen und die Ebene der Inhalte von Unterricht. Die erste, übergeordnete Ebene ist im sogenannten „nationalen Curriculum“ niedergelegt. Hier sind die fundamentalen, den Unterricht durchziehen-

den Werte, die grundlegenden Inhalte und die Richtlinien des Schulsystems aufgeführt sind. Dazu gehören auch ganz allgemeine Werte, die ein friedliches und produktives Leben von Menschen miteinander ermöglichen. Das nationale Curriculum definiert Minimalziele, die von allen Schülerinnen und Schülern zum Abschluss der Schulausbildung erreicht sein müssen. Dazu gehört u.a. das bewusste Vertreten ethischer Normen, die Fähigkeit zur Kommunikation in einer fremden Sprache und – in Bezug auf die Mathematik – die Kenntnis elementarer mathematischer Konzepte und die Fähigkeit zu ihrer Anwendung. Aber es sind nur diese wenigen Zeilen, die in diesem Curriculum zu den verschiedenen Fächern aufgeführt werden.

Der *Syllabus* als Ergänzung (zweite Ebene) benennt fachspezifische Ziele, beschreibt die Rolle des Fachs im Rahmen der allgemeinen Bildung und weist ihren Bezug zu den fundamentalen Werten und Normen auf. Für die Regelschule führt der *Syllabus* auch einen Minimalkatalog von Wissen auf, das zum Ende der fünften bzw. zum Ende der neunten Jahrgangsstufe erworben sein muss. Abbildung 4 zeigt am Beispiel des Mathematikunterrichts, welche Ziele hier am Ende des neunten Schuljahres erreicht sein sollten.

#### **Goals that pupils should have attained by the end of the ninth year in school**

Pupils should have acquired the knowledge in mathematics needed to be able to describe and manage situations, as well as solve problems that occur regularly in the home and society, which is needed as a foundation for further education.

Within this framework, pupils should

- have developed their understanding of numbers to cover whole and rational numbers in fraction and decimal form,
- have good skills in and be able to make estimates and calculation of natural numbers, numbers in decimal form, as well as percentages and proportions in their head, with the help of written calculation methods and technical aids,
- be able to use methods, measuring systems and instruments to compare, estimate and determine length, area, volume, angles, quantities, points in time and time differences,
- be able to reproduce and describe important properties of some common geometrical objects, as well as be able to interpret and use drawings and maps,
- be able to interpret, compile, analyse, and evaluate data in tables and diagrams,
- be able to use the concept of probability in simple random situations,
- be able to interpret and use simple formulae, solve simple equations, as well as be able to interpret and use graphs for functions describing real relationships and events.

Abb. 4: Syllabuses for the Compulsory School. Skolverket and Fritzes 2001.  
(<http://www.skolverket.se/english/publ.shtml>)

Der *Syllabus* im Fach Mathematik für die Oberstufe umfasst auch Bewertungskriterien, die bestimmen, wann ein Schüler einen Kurs bestanden hat, wann er ihn mit Auszeichnung und wann mit besonderer Auszeichnung bestanden hat. Sie lassen sich als Niveaustufen der mathematischen Kompetenz interpretieren.

Insgesamt kann man die schwedischen Curriculumdokumente durchaus als Bildungsstandards im Sinne dieser Expertise interpretieren. Allerdings ist hier das Kompetenzmodell keineswegs so ausgearbeitet wie in den nordamerikanischen Beispielen.

Noch deutlicher auf Kompetenzniveaus ausgerichtet ist das *National Curriculum* in England. Es umfasst Inhalte und Leistungsziele und bestimmt darüber hinaus, wie die erreichte Leistung bewertet wird. Auch dieses Curriculum basiert auf einem klar formulierten Wertesystem, das Bildung und Erziehung als leitenden Weg zur geistigen, moralischen, gesellschaftlichen, kulturellen, körperlichen und geistigen Entwicklung und damit zum Wohl des Individuums auffasst. In der Präambel werden diese Gedanken ausgeführt, vertieft und in den Kontext von Schule und Familie gestellt. Die fachbezogenen Inhalte („Programmes of Study“) sind eher knapp gehalten und werden im Wesentlichen über die Leistungsziele („Attainment Targets“) und die dazu gehörenden Kompetenzniveaus („Level Descriptions“) operationalisiert. Diese Leistungsziele sind entsprechend sehr konkret gehalten und nennen teilweise einzelne Inhaltsaspekte. Regelmäßige Tests, an denen alle Schulen teilnehmen, operationalisieren diese Ziele und geben den Schulen eine Rückmeldung über das von ihren Schülerinnen und Schülern erreichte Kompetenzniveau.

### **3.4 Entwicklungen in Deutschland: Von detaillierten Lehrplänen zu kompetenzorientierten Rahmenplänen und Prüfungsanforderungen**

Bildungsstandards wurden in Deutschland – bis auf allerneueste Ansätze, die im Schlusskapitel unter 12.1 diskutiert werden – bislang nicht vorgelegt. Umso wichtiger ist es, sich die Traditionen der Lehrplanarbeit vor Augen zu führen.

Die Lehrpläne bzw. Rahmenrichtlinien legen fest, welche Ziele und Inhalte in einem Fach, einer Jahrgangsstufe und einem Schultyp in einem bestimmten Land Gegenstand des Unterrichts sind. Dabei bleiben sie im zeitlichen Verlauf nicht unverändert, sondern sind in gesellschaftliche Strömungen eingebunden und spiegeln pädagogische, didaktische und fachliche Entwicklungen wider. Auch wenn die Lehrpläne bzw. Rahmenrichtlinien sich in den verschiedenen Ländern zum

Teil deutlich voneinander unterscheiden, kann man über die Jahre hinweg ähnliche Tendenzen in ihrer Entwicklung wahrnehmen.

Betrachtet man die Lehrpläne der Länder bzw. auch der DDR in den späten 60er und 70er Jahren, so dominiert ein Bild von Mathematik, das begriffliche Aspekte und systematische Betrachtungen in den Vordergrund stellt. Darüber hinaus zeigen sich die Richtlinien in hohem Maße detailgesteuert. Je nach Land werden mehr oder minder explizit Ziele und Inhalte, aber auch das nach Sicht der Autoren geeignete methodische Vorgehen und entsprechende Lernzielkontrollen vorgeschlagen oder vorgeschrieben.

Ein gutes Jahrzehnt später, in den 1990er Jahren, zeigen die Lehrpläne eine Tendenz, die Detailsteuerung zugunsten allgemeinerer Ziele aufzugeben. Dies zeigt sich schon alleine am Umfang. Im Grunde nennen diese Lehrpläne vor allem die Inhalte und lassen die konkrete Ausarbeitung verstärkt in der Hand von Lehrerinnen und Lehrern. Die Hinweise scheinen allerdings nicht selten eher unsystematisch hinzugefügt worden zu sein und lassen so ein klares Bild (und damit eine pädagogische oder didaktische Theorie des Lehrens und Lernens) vermissen.

Der derzeit in der Diskussion stehende Entwurf eines neuen Lehrplans für das Gymnasium in Bayern durchbricht schließlich die strenge Aufteilung in Inhalte und ihre Explikation. Er beschränkt sich auf die Nennung von Unterrichtsinhalten und Richtstundenzahlen für relativ große Themenblöcke. Die Schüleraktivität gerät nun stärker in den Blickpunkt und drückt sich in Formulierungen aus, die etwa das Entdecken mathematischer Inhalte fordern. Damit gehen einerseits (wenn auch vorsichtig) Strömungen ein, die Lehren und Lernen als einen konstruktiven Prozess begreifen, andererseits findet sich auch eine Auffassung von Mathematik, die mehr prozessorientiert und weniger produktorientiert ist. Auch hier kann man in vielen Ländern ähnliche Tendenzen beobachten. Der Lehrplan von Schleswig-Holstein aus dem Jahr 2002 betrachtet zunächst, welche Kompetenzen ein spezielles Thema vermitteln soll, und unterscheidet dann in Inhalte und fachspezifische Hinweise. Der neue Lehrplan von Mecklenburg-Vorpommern für die Orientierungsstufe betrachtet schließlich den Erwerb von fachübergreifenden Kompetenzen als wesentliches Ziel des Schulunterrichts. Kompetenz wird dabei allerdings – anders als in dieser Expertise – als fächer-übergreifende Handlungskompetenz verstanden, die Sachkompetenz, Methodenkompetenz, Selbstkompetenz und Sozialkompetenz umfasst.

In der Konsequenz zeigen die Lehrpläne in den letzten 25 Jahren durchaus eine positive Entwicklung. Aus stark fachorientierten Rahmenrichtlinien, die Unterricht in hohem Maße zu steuern versuchten, sind Handreichungen geworden, die zunehmend die Schülerinnen und Schüler und ihre Lernprozesse in den Vordergrund stellen. Auch wenn die Entwicklung von Lehrplänen und Rah-

menrichtlinien eine Tendenz zu einer geringeren Detailsteuerung von Schule verdeutlicht, sind damit aber noch keine Bildungsstandards definiert. Lehrpläne beschreiben in der Regel die Anforderungen für ein spezielles Fach und eine spezielle Klassenstufe innerhalb eines Bildungsgangs und können damit allenfalls eine Grundlage für die Definition von Standards sein. Ihnen fehlen grundlegende Komponenten und Merkmale, wie sie in den Abschnitten 2.1 und 2.2 als konstitutiv für Bildungsstandards dargestellt wurden.

Am nächsten kommen dem Konzept der Bildungsstandards in Deutschland die KMK-weiten Vereinbarungen über Bildungsabschlüsse. In Bezug auf den Mathematikunterricht zeigen übrigens auch die *Einheitlichen Prüfungsanforderungen in der Abiturprüfung* aus dem Jahr 2002 einen Weg in Richtung der Festlegung geeigneter Standards. Sie basieren wesentlich auf übergreifenden Inhalten, die im Unterricht im Grunde von Beginn an eine Rolle spielen sollten, und haben ein Stück der oft kritisierten Kleinschrittigkeit und inhaltlichen Gebundenheit zugunsten übergreifender fachlicher Zusammenhänge aufgegeben. So werden unter der Überschrift *Fachliche Inhalte und Qualifikationen* sowohl fachbezogene methodische Kompetenzen als auch fachliche Inhalte subsumiert. Teilaspekte der methodischen Kompetenz sind beispielsweise die Fähigkeit zum mathematischen Modellieren realitätsnaher Probleme, zum mathematischen Argumentieren und Begründen oder das Auswählen, Nutzen und Bewerten von Informationen. Die fachlichen Inhalte sind weiterhin den klassischen Bereichen Analysis, Lineare Algebra und Stochastik entnommen, sie werden aber unter Leitideen gestellt, die ihren Stellenwert und ihre Entwicklung im Curriculum verdeutlichen. Die daraus resultierenden Aufgabenstellungen sind eher offen formuliert und lassen in der Regel explizit mehrere Lösungswege zu. Sie sind jeweils Anforderungsbereichen zugeordnet, die im Sinne eines Kompetenzstufenmodells gesehen werden können. Die Prüfungsanforderungen umfassen außerdem Kriterien, mit denen das Erreichen des angestrebten Kompetenzniveaus überprüft werden kann.

Auch die *Einheitlichen Prüfungsanforderungen in der Abiturprüfung* für das Fach Mathematik sind keine Bildungsstandards und haben diesen Anspruch auch nicht erhoben. Sie enthalten aber bereits wesentliche Elemente und sind sicherlich ein Beispiel, das in der Bundesrepublik Deutschland auf das Fach Mathematik bezogen am konsequentesten ausgearbeitet ist.

## 4. Zur Bedeutung von Standards für die Weiterentwicklung von Schule und Unterricht

Im Grunde lässt sich die Funktion von Bildungsstandards mit einem Satz beschreiben: *Sie arbeiten in klarer und konzentrierter Form heraus, worauf es in unserem Schulsystem ankommt.* Sie konkretisieren den pädagogischen Auftrag der Schule in zentralen Lernfeldern und richten somit das Lehren und Lernen auf gemeinsame Ziele aus.

Diese *Orientierungsfunktion* nutzt den Lernenden und ihren Eltern; sie dient aber auch der Professionalität der Lehrkräfte und der Qualitätsentwicklung auf der institutionellen Ebene.

Konkretisiert in Testverfahren, werden Standards im Rahmen des Bildungsmonitoring und der Evaluation von Schulen angewandt. Sie dienen der Feststellung und Bewertung von Lernergebnissen und haben somit eine *Rückmeldfunktion*, mit der sie zur output-orientierten Steuerung beitragen. Feed-back und *Output*-Orientierung sind aber kein Selbstzweck. Ihr Ziel ist es, die Wirkungen (und Nebenwirkungen) des pädagogischen Handelns in den Blick zu nehmen und so professionelles, rationales Handeln zu ermöglichen. Testerhebungen an Schulen sind also dann und nur dann sinnvoll, wenn sie dazu beitragen, die Professionalität der Lehrenden zu fördern und die Qualität von Schule und Unterricht weiterzuentwickeln. Hierauf gehen wir an anderer Stelle näher ein (Kapitel 9 und 11).

Im Folgenden konzentrieren wir uns auf den Arbeits- und Lebenszusammenhang in den Schulen. Wie wird er durch die Einführung von Bildungsstandards beeinflusst? Welche Konsequenzen haben Bildungsstandards für Schüler, Eltern und Lehrpersonen? Und welche Funktionen sollten Standards und standard-bezogene Tests gerade *nicht* erhalten?

### 4.1 Bildungsstandards aus der Perspektive der Schülerinnen und Schüler und ihrer Eltern

Bildungsstandards sollen so formuliert sein, dass die Eltern und zumindest die älteren Schülerinnen und Schüler sie nachvollziehen können. Die Standards vermitteln diesen Gruppen mit mehr Klarheit als zuvor, auf welche Kompetenzen es in der Schule ankommt und wie diese über die Schullaufbahn hinweg aufgebaut werden. Es sollte auch von Beginn an deutlich werden, dass Bildungsstandards auf eine veränderte Beteiligung und Mitwirkung von Schülerinnen, Schülern und Elternhäusern abzielen, etwa im Sinne einer stärkeren Übernahme der Planung, Begleitung und Überwachung von Lernprozessen. Gleichzeitig verpflichtet sich die Schule ihnen gegenüber, auf bestimmte Ziele hinzuarbeiten und dafür Fördermaßnahmen bereitzustellen. Standards kön-

nen damit einen Bezugspunkt bilden für Gespräche zwischen Eltern, Schülern und Lehrern. Unter Berufung auf klare und verbindliche Erwartungen wird es möglich, den individuellen Lernweg zu planen, Lernhindernisse zu erkennen und bestmögliche Fördermöglichkeiten abzusprechen. Die Orientierung an Grunddimensionen der Kompetenzentwicklung macht solche Gespräche einfacher als ein traditioneller, detaillierter Katalog von Lernzielen und -inhalten.

Die Ziele und Anforderungen der Schule werden somit für alle Beteiligten transparent. Insbesondere wenn Mindeststandards festgelegt werden, kommen Fragen der Bewertung von Lernergebnissen auf den Tisch. Dabei muss jedoch sehr klar unterschieden werden zwischen der Diskussion über die Qualität der Lernergebnisse einerseits sowie der Festlegung von Noten, der Vergabe von Zertifikaten und der *Entscheidung* über die Schullaufbahn von Schülern andererseits. Standards sollen ein Kriterium für die Lernergebnisse jeder Schule, jeder Klasse und konsequenterweise jeder einzelnen Schülerin und jedes einzelnen Schülers sein. Sie dürfen aber nicht mit Benotungs- und Prüfungsregeln oder gar mit Regelungen zum Übergang auf weiterführende Schulen verwechselt werden.

Die Bildungsstandards decken erklärtermaßen nicht das gesamte Curriculum, sondern nur einen Kern in zentralen Domänen des Lernens ab. Zum zweiten ist die Benotung und die Vergabe von Abschlüssen mit gutem Grund Gegenstand von Entscheidungen, die Lehrerinnen und Lehrer nach professionellen Maßstäben, aber selbständig, unter Abwägung verschiedenster Aspekte, in ihrer pädagogischen Verantwortung treffen müssen. Beide Argumente sprechen dafür, Bildungsstandards nicht als Kriterien für Notengebung, Zertifizierung und Selektionsentscheidungen zu missbrauchen.

Besonders wichtig wird diese Abgrenzung, wenn auf die Standards bezogene Tests ins Spiel kommen. Erfahrungen mit dem *high stakes assessment* in den USA haben die negativen Konsequenzen des Missbrauchs von zentralen Tests für Einzelfallentscheidungen in der Schule deutlich gemacht: Der Test beginnt den Unterricht zu dominieren – was bei intelligenten Tests, die Raum für komplexes Denken, Problemlösen und Kreativität lassen, durchaus fruchtbar sein, aber bei schlechten Tests fatale Auswirkungen haben kann. Hinzu kommt, dass Tests, die für Bildungsmonitoring und Schulevaluation entwickelt wurden, zumeist gar nicht für die individuelle Diagnostik geeignet sind, denn jeder einzelne Schüler bearbeitet nur einen kleinen Teil der Aufgaben und die Fehlermarge bei der Abschätzung individueller Leistungskennwerte ist zu hoch (vgl. Kapitel 9). Die Expertengruppe rät daher nachdrücklich zu einer deutlichen Trennung zwischen der Verwendung standard-bezogener Tests für Evaluation, Bildungsmonitoring und (wenn dies methodisch zulässig ist) als Entscheidungshilfe für individuelle Förderung einerseits (vgl. Kapitel 9), Noten und Abschlussprüfungen andererseits. Dies ist mit ein Argument dafür, Testeinsätze nicht in den Abschlussjahren durchzuführen.

Um es ganz deutlich zu formulieren: Diese Expertise sieht die Funktion von Bildungsstandards nicht darin, den individuellen Leistungs- und Selektionsdruck auf Schülerinnen und Schüler zu verstärken. Im Vergleich mit anderen Staaten bzw. Kulturen zeichnet sich Deutschland ohnehin

dadurch aus, dass die Schüler mehr Leistungsdruck als Unterstützung wahrnehmen, während die Schule in den nordischen und den angelsächsischen Ländern, wo Standards und Schulevaluation vertraut sind, eher als unterstützend und aner kennend erlebt wird. Bildungsstandards könnten in unserer Gesellschaft vielmehr dazu führen, dass bestimmte Erwartungen an das Kompetenzniveau als selbstverständlich gelten, für deren Einlösung Schülerinnen und Schüler, Eltern und Schule gemeinsam die Verantwortung übernehmen.

## 4.2 Bildungsstandards aus der Sicht von Lehrerinnen und Lehrern

Das Konzept von Bildungsstandards, das hier empfohlen wird, sieht ausdrücklich ergebnisbezogene Standards vor, keine prozessbezogenen Standards (*opportunity to learn standards*). Bildungsstandards bedeuten daher gerade *keine* Standardisierung oder Normierung des pädagogischen Handelns, keine „Gleichmacherei“. Standards betonen die Verantwortung der Schulen und Lehrkräfte für die Lernergebnisse und schaffen gleichzeitig mehr Raum für eigenständiges professionelles Handeln.

Die Einführung von Bildungsstandards verändert das Bild von Professionalität im Lehrerberuf. Standards bedeuten für Lehrkräfte nicht nur andere Referenzsysteme zur Unterrichtsplanung, sondern auch mehr Freiheitsgrade, Flexibilität und Verantwortung. Die Verantwortung für die Gestaltung von Lehr-Lern-Prozessen bis hin zur Ausarbeitung eines Schulcurriculums wird stärker an die Schulen zurückgegeben, aber nicht unbedingt an den einzelnen Lehrer, sondern beispielsweise an Fachteams. Neben dem individuellen Unterricht wird daher die Abstimmung im Schulkollegium die Tätigkeit von Lehrerinnen und Lehrern zunehmend prägen. Die Lehrkräfte übernehmen außerdem eine Schlüsselposition – im Sinne einer Beratungsfunktion –, wenn es darum geht, Standards für die Lernplanung und eigenständige Lernüberwachung durch die Schülerinnen und Schüler sowie durch die Eltern zu nutzen. Im Zusammenhang mit Evaluation und Bildungsmonitoring kommt der Umgang mit Rückmeldungen über die Arbeit der Schule und über die Wirkung des eigenen Unterrichts hinzu.

Den Lehrenden liefern die Standards ein Referenzsystem für ihr professionelles Handeln. Anders als Lehrpläne, die oft zu sehr ins Detail gehen und eine Vielzahl einzelner Lerninhalte und -ziele auführen, sollen Standards den Unterricht klar fokussieren. Sie gehen von den zentralen Ideen aus, die ein Fach konstituieren, und stellen dar, in welchen Teildimensionen und Niveaustufen sich die Kompetenzen der Schüler entfalten. Damit enthalten Bildungsstandards im Kern eine systematische fachdidaktische Konzeption, an der sich die Lehrkräfte orientieren und die sie eigenständig präzisieren können. Zentrale Ideen des Faches werden im Unterricht herausgearbeitet, und in der Beschreibung der Kompetenzmodelle wird deutlich, auf welche grundlegenden Begriffe und Operationen der Unterricht eingehen muss. Aufgabenbeispiele, mit denen die Kom-

petenzstufen in den Standards illustriert werden, geben – wenn sie intelligent und anspruchsvoll gestaltet sind – Anregungen auch für Unterrichtsaufgaben und für die Unterrichtsplanung. Indem Bildungsstandards kumulatives Lernen über mehrere Jahrgänge und Niveaustufen hinweg darstellen, verändern sie den Blick auf Unterricht von einer reinen fachsystematischen Perspektive hin zu einer stärker „schülerorientierten“, d.h. an der kognitiven Entwicklung der Lernenden ausgerichteten Perspektive.

Mit Blick auf Kompetenzmodelle kann die Lehrperson verschiedenartige Lernwege und interindividuelle Unterschiede einordnen, d.h. die Heterogenität von Lernprozessen und Lernergebnissen verstehen. Durch diesen Vergleichsmaßstab wird die diagnostische Kompetenz der Lehrerin bzw. des Lehrers geschärft. Die Minimal Kriterien, bei deren Erfüllung der Standard gerade eingelöst ist, sind hier besonders wichtig, weil die Lehrperson im Vergleich mit diesen Kriterien einschätzen kann, welche Schülerinnen und Schüler besonderer Förderung bedürfen.

All diese Orientierungen basieren auf den Zielbestimmungen, Kompetenzmodellen und Mindestanforderungen, die zu Bildungsstandards gehören. Sie können – flankiert durch entsprechende Fortbildung – das fachdidaktische Denken und die diagnostische Kompetenz der Lehrer voranbringen. Werden im Rahmen einer Monitoring- oder Evaluationsstudie zudem in der eigenen Schule standard-bezogene Tests eingesetzt, erhalten die Lehrkräfte zusätzlich die Chance, ihr eigenes diagnostisches Urteil mit dem Testergebnis zu vergleichen. Dies sollte – wie oben begründet – keinesfalls dazu führen, dass Noten und Abschlüsse oder Übergangsempfehlungen von zentralen Tests abhängig gemacht werden. Vielmehr geht es darum, dass Lehrerinnen und Lehrer überhaupt mit einem externen Maßstab zur Einschätzung der Schülerkompetenzen konfrontiert und auf mögliche Einseitigkeiten des eigenen Urteils hingewiesen werden. Zudem geben solche Untersuchungen Rückmeldungen über den Erfolg der eigenen pädagogischen Arbeit.

Insgesamt fordern Bildungsstandards Lehrkräfte dazu auf, fachbezogene (und fächerübergreifende) Lehr- und Lernprozesse in Hinblick auf klare Zielstellungen neu zu durchdenken, das Wissen und Können der Schülerinnen und Schüler differenziert wahrzunehmen, eine realistische Vorstellung von der eigenen Wirksamkeit aufzubauen, und an der Weiterentwicklung der eigenen Professionalität und der der eigenen Schule (mit)zu arbeiten.

Die hier skizzierten Anforderungen machen deutlich, dass Bildungsstandards nicht nur neue Aufgaben bedeuten, sondern Lehrkräfte in ihrem Selbstverständnis bestärken können. Ein Teil der Lehrerinnen und Lehrer werden sie jedoch möglicherweise als Irritation oder Zumutung empfinden. Akzeptanz in den Kollegien zu gewinnen ist daher eines der wichtigsten Anliegen bei der Implementation von Standards. Die Lehrkräfte sollten eine klare Vorstellung darüber entwickeln können, welche Chancen auf kürzere und auf lange Sicht mit der Stärkung ihrer professionellen Rolle in der Arbeit mit Bildungsstandards verbunden sind. Vor allem aber müssen ihnen Handlungsgerüste angeboten werden, die Sicherheit vermitteln und eine erfolgreiche Nutzung von Standards gewährleisten. Damit kommen auf die Einrichtungen, die schulische Arbeit vorbereiten und unterstützen, auf die Schulaufsicht, die Landesinstitute und die Lehrerbildung, besondere

Aufgaben zu. Auf lange Sicht bilden diese Einrichtungen jene Infrastruktur, die Lehrkräfte und Schulen bei der Arbeit mit Standards zu unterstützen hätte (vgl. Kapitel 10).

### 4.3 Bildungsstandards und Schulentwicklung

Seit langem wird auch in Deutschland die einzelne Schule als Ort der Qualitätsentwicklung im Bildungssystem wahrgenommen. Viele Initiativen der Länder bemühen sich um die Stärkung der Einzelschule, die Verbesserung des sozialen Klimas und der „Schulkultur“, die Förderung der Kooperation innerhalb der Kollegien, die Profilierung der Schule und die Öffnung nach außen. Schulentwicklung bedeutet, dass an den Schulen selbst eine Verständigung über Ziele, über Verantwortlichkeit, Zusammenarbeit und über Rechenschaft sowie den Umgang mit Rückmeldungen erfolgen muss. Neue Verfahren der gemeinsamen Planung und Abstimmung werden entwickelt, erprobt und wiederum in Routinen umgesetzt. Der Schulleitung kommt hierbei eine zentrale motivierende, koordinierende und integrierende Bedeutung zu, die weit über administrative Funktionen hinausgeht. Sie muss Verantwortung für die Gestaltung der Schule und die Ergebnisse der pädagogischen Arbeit übernehmen und gleichzeitig die angemessene Partizipation von Lehrkräften, Schülerinnen und Schülern und Eltern in demokratischen Prozessen sicherstellen und achten.

*Schulprogramme* sind das zentrale Medium, in dem Stärken und Problembereiche benannt, Entwicklungsziele formuliert und Maßnahmen geplant werden. Allzu oft dominieren in Schulprogrammen jedoch Aussagen über die Schulkultur insgesamt, fächer-übergreifende Aktivitäten und Projekte. In den Initiativen für die innere Schulentwicklung spielen die fachlichen Lehr-Lern-Prozesse, der Aufbau von Kompetenzen innerhalb der Lernbereiche häufig eine eher geringe Rolle – gemessen an der zentralen Bedeutung dieser Fragen für den Erfolg von Schule und Lernen.

Bildungsstandards fordern nun dazu heraus, die systematische Förderung der Kompetenzen von Schülerinnen und Schülern in den Mittelpunkt der Schulentwicklung zu stellen. Sie stellen hierfür eine Sprache und ein analytisches Raster zur Verfügung, das geeignet ist, kontinuierliche Lernprozesse über Jahrgangsstufen hinweg zu konzipieren und zu beschreiben.

Die notwendige Abstimmung zwischen den Lehrkräften einer Schule über Wege zur Einlösung von Standards, zur Planung von Übergängen und Anschlüssen zwischen den Jahrgangsstufen und damit zur Herstellung und Aufrechterhaltung von produktiven Übergängen zwischen Schulformen (etwa zwischen Grundschule und Sekundarstufe I) wird dazu führen, dass neben dem allgemein-pädagogischen Schulprogramm ein *Schulcurriculum* für das fachbezogene Lernen und die Verbindung der Fächer entsteht. Langfristig erscheint nämlich eine Koexistenz von nationalen Bildungsstandards einerseits und detaillierten curricularen Vorgaben auf Länderebene anderer-

seits als unwahrscheinlich. Die Lehrpläne und Rahmenrichtlinien werden sich eher zu Kencurricula hin entwickeln (vgl. Kapitel 8). Damit kommen auf die Lehrkräfte der Schulen Aufgaben der konkretisierenden Lehrplanung zu, die aber auch mehr Freiraum für die Schule bedeuten. Es wird zunehmend der einzelnen Schule überlassen bleiben festzulegen, wie genau – mit welchem Curriculum, mit welcher Stundentafel in den einzelnen Jahrgängen, mit welchen Formen der Unterrichtsorganisation im einzelnen – die Kompetenzziele erreicht werden sollen und können. Insbesondere wird jede Schule eigene, auf ihre Schülerschaft abgestimmte Maßnahmen zur gezielten, individuellen Förderung entwickeln. Die Vorgabe von Mindestanforderungen zwingt jede Schule besonders dazu, für die Förderung der Leistungsschwächeren selbst Verantwortung zu übernehmen und in deren Förderung zu investieren.

Eine neue Herausforderung wird für Schulen der Umgang mit Ergebnismrückmeldungen aus Evaluationen und dem Bildungsmonitoring sein. Wenn Bildungsstandards eine output-orientierte Steuerung des Schulsystems implizieren, werden Schulen regelmäßig – in anderen Staaten hat sich ein Rhythmus von drei bis fünf Jahren bewährt – ihre Stärken und Schwächen analysieren und dabei als zentrales, wenn auch nicht einziges Kriterium die erreichten Kompetenzen der Schülerinnen und Schüler zur Kenntnis nehmen. (Wie dies geschehen kann und welche Fragen dabei zu klären sind, wird in Kapitel 9 diskutiert). Die Expertengruppe sieht in der Messung von Lernergebnissen (Schülerkompetenzen) eine Chance für die Schulen, sich der Ergebnisse der eigenen Arbeit zu vergewissern und darauf professionell zu reagieren. Schulrückmeldungen sollen deshalb an die Kollegien und die Gremien der Schule gerichtet sein, nicht an die Öffentlichkeit. Insofern die Schulaufsicht eine unterstützende und beratende Funktion ausübt, kann sie einbezogen werden. Standardbezogene Kompetenzmessungen, wie sie hier vorgeschlagen werden, münden nicht einfach in Vergleichsdaten, sondern geben inhaltliche Hinweise auf Stärken und Schwächen im Kompetenzprofil der Schülerinnen und Schüler, die von großer Bedeutung für die Ausgestaltung des Schulcurriculums und des Unterrichts sind. Nicht zu unterschätzen ist auch die Bedeutung von Ergebnismrückmeldungen als Anerkennung der Leistung des Kollegiums.

#### **4.4 Bildungsstandards und die Qualitätsentwicklung des Bildungssystems**

Schulübergreifend sind Bildungsstandards als ein Instrument zur Förderung der Bildungsgerechtigkeit zu sehen. Mit der Entscheidung, Bildungsstandards auf nationaler Ebene zu etablieren, haben die Kultusminister sich zu dem Ziel bekannt, vergleichbare Bildungsqualität in allen Ländern herzustellen und damit zur Gleichheit der Chancen auf Bildung bundesweit beizutragen. Sie ziehen damit Konsequenzen aus dem Befund der PISA-Studie, dass das deutsche Bildungssystem durch besonders starke Disparitäten zwischen Regionen und sozialen Gruppen gekennzeichnet ist. Schulformübergreifende Mindeststandards, wie sie hier vorgeschlagen werden, be-

tonen diesen Aspekt, lassen aber – über abgestufte Kompetenzmodelle – auch Raum für besondere Kompetenzprofile und weitergehende Anforderungen.

Bildungsstandards werden für das System weitreichende Folgen haben, die in dieser Expertise nur teilweise entfaltet werden können. Eine konsequente Ergebnisorientierung bei der Steuerung des Bildungssystems wird neue Strukturen erfordern (vgl. Kapitel 11), aber auch die Unterstützungsarbeit bestehender Einrichtungen herausfordern (vgl. Kapitel 10). Die Rolle nicht nur der Lehrkräfte und der Einzelschulen, sondern auch der Schulaufsicht, der Landesinstitute und anderer Akteure werden sich ändern; Schulbuchverlage werden sich auf einen weniger eng regulierten Markt einstellen müssen; Schulträger werden gefordert sein, Ressourcen zur Einlösung von Standards, vor allem für individuelle Förderung bereitzustellen und vieles andere mehr.

Die Einführung von Bildungsstandards ist ein Signal nicht nur für die an Bildungsprozessen unmittelbar Beteiligten, sondern für die gesamte Öffentlichkeit, weil diese Gesellschaft für eine systematische Förderung der Kompetenzen von Kindern und Jugendlichen Sorge tragen muss. Bildungsstandards bringen also das Thema „Bildung“ auf der politischen und gesellschaftlichen Agenda nach vorne, aber sie befördern auch den gesellschaftlichen Diskurs über die Bedeutung von Bildung und deren Inhalte. Dies wird im folgenden Kapitel ausgeführt.

## 5. Bildungsstandards, Kompetenzmodelle und Bildungsziele: Konstruktions- und Legitimationsprobleme

Nationale Bildungsstandards gehören ohne Zweifel in den Kontext der Zielvorgaben für das Bildungssystem, denn sie setzen den Standard für die Leistungen der Schule und zwar so, dass man ihn an Individuen vergleichend messen kann. Bildungsstandards sind insofern aber auch Teil der kontroversen gesellschaftlichen Auseinandersetzungen über die Ziele der Schule und über die konkreten Erwartungen, die man an das Bildungssystem und die Kompetenzen seiner Absolventen haben kann und darf. Damit sind aber auch für nationale Bildungsstandards die schwierigen Fragen der Entwicklung, systematischen Begründung und politisch-gesellschaftlichen Legitimation der Ziele von Bildung und Erziehung ebenso aufgeworfen wie für Bildungsfragen insgesamt.

Den allgemeinen Kontext dieser Fragen kann und will die vorliegende Expertise weder umfassend diskutieren noch in jeder Beziehung zufriedenstellend klären. Im folgenden Kapitel sollen aber die spezifischen Probleme und die besonderen Möglichkeiten erörtert werden, die mit der Formulierung von Bildungsstandards für die Klärung normativer Erwartungen an das Bildungssystem entstehen.

Es sind vor allem zwei Dimensionen, die dabei Beachtung verdienen:

- (1) In kritischer Wendung gegen Bildungsstandards entsteht die Befürchtung, dass mit der Einführung von Bildungsstandards die Zieldiskussion nur noch inhaltlich verkürzt oder sogar nur bildungstheoretisch trivialisiert geführt wird oder geführt werden kann. Orientiert an den Kriterien der Messbarkeit und an den Unwägbarkeiten des Systemvergleichs werden, so die Kritik, die Erwartungen an das Bildungssystem nivelliert und reduziert.
- (2) Auf der anderen Seite werden die Vorteile der Formulierung von exakten und messbaren Bildungsstandards stark betont. Die Konstruktion und Implementation von Bildungsstandards, so diese Gegenposition, kann die Zieldebatte für Bildungsprozesse in konstruktiver Wendung ernüchtern und dazu beitragen, die Möglichkeiten des Bildungssystems und seine Leistungen und Defizite gerade auf der Stufe grundlegender Allgemeinbildung realistisch zu sehen.

Die hier vorgelegte Expertise geht von der systematischen Annahme aus, dass die kritischen Befürchtungen zwar aus der Tradition der gesellschaftspolitischen, pädagogischen und bildungstheoretischen Zieldebatten verständlich sind, dass sie aber nicht nur die Schwächen der traditionellen Vorgaben für das Bildungswesen eher zu gering veranschlagen, sondern auch die Vorteile einer Einführung von Bildungsstandards angesichts der unübersehbaren systematischen Schwierigkeiten der Debatte über Bildungsziele eher verkennen.

## 5.1 Bildungsstandards und Bildungsziele

Orientiert an der Erwartung, Basisfähigkeiten im Bildungssystem zu generalisieren und zugleich der Messbarkeit der Kompetenz und des Kompetenzniveaus zu genügen, sehen sich nationale Bildungsstandards mit einer mehrfachen Kritik konfrontiert. Die Dimensionen dieser Kritik sind in dreifacher Hinsicht systematisierbar: (1) als *Reduktionismuskritik*, (2) als *Vorwurf der Vereinheitlichung und Nivellierung* und (3) als *Vorwurf des Legitimationsdefizits*.

### (1) Reduktionismuskritik

Im Lichte der Bildungstheorie und der gesellschaftspolitischen Zieldebatte über Bildungsprozesse und -systeme ist dieser Vorwurf aktuell vor allem in der Diskussion über die PISA-Ergebnisse vorgetragen worden: Die Orientierung an den „kulturellen Basisfähigkeiten“ und die Operationalisierung in unterschiedlichen Dimensionen von „Literalität“ – z.B. des sprachlichen, mathematischen oder naturwissenschaftlichen Verstehens – repräsentierten, so die Kritiker, ein unzureichendes Konzept der modernen Allgemeinbildung (vgl. für diesen Teil der deutschen Diskussion exemplarisch Benner 2002). Konzentriert auf Basisfähigkeiten würden die Erwartungen unterboten, die im Blick auf die Inhalte, Aufgaben, immanente Stufung und die Praxisformen von Bildungsprozessen notwendig seien und z.B. in der Vorbereitung des wissenschaftsorientierten und wissenschaftspropädeutischen Lernens auch die Stufe der Grundbildung bestimmen müssten. Die systematische Ursache für diesen Reduktionismus sehen die Kritiker einerseits in der falschen, nur pragmatischen oder zu sehr vereinfachenden bildungstheoretischen Grundlegung und andererseits in der – für die Kritiker – irrigen Annahme, die wesentlichen Erwartungen an schulische Bildungsprozesse ließen sich hinreichend so operationalisieren, dass sie wirklich messbar werden.

### (2) Vorwurf der Vereinheitlichung und Nivellierung

Der zweite Vorwurf knüpft hier an und sieht dann in Bildungsstandards den untauglichen Versuch, den wesentlich individuell zu verstehenden Prozess der Bildung übergreifenden Standards zu unterwerfen, und zwar sowohl im Tempo des Kompetenzerwerbs als auch in der Definition und Akzeptanz der Ergebnisse. Der Vorwurf der Vereinheitlichung und Nivellierung wird dabei selbst in zweifacher Weise erhoben, sowohl im Blick auf die lernenden Subjekte als auch im Blick auf die Institutionen, also die Schularten und ihre je spezifischen Adressaten und Aufgaben sowie angesichts der konkreten unterrichtlichen Traditionen und Ansprüche. Nivelliert würde dabei auch in zwei Richtungen, sowohl „nach unten“ als auch „nach oben“: Anspruchsvolle Erwartungen, z.B. im gymnasialen Deutschunterricht, würden zugunsten eines simplifizierenden Konzepts sprachlichen Verstehens quasi ausgeschlossen, ein Modell von Minimalerwartungen würde andererseits zum Standard, der die Rücksicht auf besondere Lebenslagen oder Probleme ignoriere.

### (3) Legitimationsdefizit

Ein Legitimationsdefizit wird den Bildungsstandards schließlich insofern zugeschrieben, als in der Debatte befürchtet wird, dass Expertenherrschaft und übernationale Kommissionen die Bildungsdebatte bestimmen, dass damit aber politische Grundsatzfragen, bildungspolitische Entscheidungen in Parlamenten und die öffentliche Auseinandersetzung über die Ziele, Normen und die Funktion der Schule wenn nicht ihr Recht, so doch weitgehend an Aufmerksamkeit verlieren würden. Wenn man schon Bildungsstandards einführen wolle, dann dürfte diese Ebene der politischen, moralischen und öffentlichen Legitimation der Erwartungen an Schule nicht übersprungen werden; sie allein sei die Gewähr dafür, dass jenseits der Fragen der Messbarkeit auch die als zentral beurteilten Kriterien schulischer Arbeit, z.B. die Frage der Bildungsgerechtigkeit, der individuellen Förderung oder der Zukunftsfähigkeit von Schule angemessen berücksichtigt werden könnten.

Die vorliegende Expertise ist von der Überzeugung getragen, dass bei der Einführung von Bildungsstandards die hier knapp vorgetragenen Bedenken zwar sehr ernst zu nehmen sind, dass die Vorwürfe aber nicht systematisch gegen den Versuch der Einführung von Bildungsstandards und schon gar nicht gegen die systematische Verbindung von Kompetenzmodellen und kerncurricularer Arbeit sprechen, die hier empfohlen wird. Zur Begründung werden wir zunächst noch einmal die grundlegenden, z.T. sogar unlösbaren Schwierigkeiten in Erinnerung rufen, die im Blick auf das allgemein bildende Schulwesen mit allen Debatten über Bildungsziele in modernen Gesellschaften verbunden sind. Danach werden wir darstellen und begründen, aus welchen Gründen Bildungsstandards zwar nicht alle Schwierigkeiten systematisch lösen, aber nicht nur den gegenwärtig besten, sondern auch einen legitimen Weg zu ihrer Bearbeitung eröffnen.

## **5.2 Bildungsziele in modernen Gesellschaften – Probleme ihrer Begründung**

Bildungsziele fallen nicht vom Himmel und sie haben nicht den Status unbefragbarer Gewissheiten, sondern verdanken sich historischen Kontexten und nationalen Traditionen. Wer „Allgemeinbildung“ sagt, der beansprucht eine – erkennbar deutsche – Tradition des „Bildungs“-denkens und der Interpretation von „Kultur“, wer von „literacy“ spricht, der orientiert sich eher am angelsächsischen und internationalen Diskurs, wer von „Basisfähigkeiten“ ausgeht, nimmt die Debatte über Standards und notwendige Erwartungen an Schule zur Kenntnis. Bildungsziele sind also, sichtbar an solchen Traditionen, in ihrer konkreten Gestalt immer Ergebnis gesellschaftlicher Entscheidungen und sozialer Machtlagen: Historisch kontrovers diskutiert und politisch entschieden, in Verfassungen und Schulgesetzen kodifiziert, theoretisch diskutiert und durch pädagogische Arbeit auf den Alltag schulischer Arbeit transformiert, existieren Bildungsziele deshalb heute auch

in großer Vielfalt. Hier werden die schönsten Zukünfte von Mensch und Welt formuliert, und erkennbar wird die Schule als ein Ort gesehen, diese Zukünfte herbeizuführen und die Normen und Ziele, Haltungen und Fähigkeiten zu realisieren, die in der Gesellschaft vermisst werden.

Damit beginnen aber die Probleme von Bildungsdebatten in modernen Gesellschaften, und auch diese Probleme lassen sich knapp charakterisieren, damit man die zugrundeliegenden Schwierigkeiten erkennt. Nach unserem Eindruck sind es zumindest fünf Problemdimensionen, die sich zur Diskussion unterscheiden lassen: (1) Unentscheidbarkeit der anthropologischen und gesellschaftlichen Prämissen für Bildungsprozesse, (2) Offenheit der Zukunft für Individuen und Gesellschaft, (3) Unbestimmtheit der Aufgaben und Anforderungen, (4) Pluralität und Konflikthaftigkeit der Erwartungen sowie (5) Utopieüberschuss und Realisierungsprobleme.

#### (1) Unentscheidbarkeit der anthropologischen und gesellschaftlichen Prämissen

Es gehört zu den Grundüberzeugungen moderner Gesellschaften seit der klassischen Epoche der Aufklärung, dass sich für die Bildungsprozesse der Individuen die anthropologischen Voraussetzungen nicht hinreichend beurteilen oder prognostizieren lassen und dass gesellschaftliche Restriktionen, etwa Nachteile der sozialen Lage, sich im pädagogischen Prozess nicht systematisch auswirken dürfen. Die individuellen Rechte und Ansprüche der Individuen und die Zielvorgabe der Gleichheit würden sonst verletzt.

Anthropologisch muss man von der Unentscheidbarkeit der Voraussetzungen ausgehen, weil erst der Bildungsprozess selbst, nicht eine vorgegebene „Natur“ – oder gar ein Geschlecht – des Menschen die Möglichkeiten konstituiert und die Kompetenzen erzeugt, mit denen die Heranwachsenden ihr Leben selbständig gestalten. Alle Versuche, quasi diagnostisch die Möglichkeiten des Menschen vorab erkennen und zuverlässig prognostizieren zu wollen, führen deshalb in die Irre; das gilt freilich genauso für die anthropologische Allmachtsphantasie mancher Pädagogen, den Menschen jenseits seiner individuellen Möglichkeiten und der Praxis seines Lernens schulisch oder pädagogisch nach ihrem Anspruch bilden zu wollen. „Begabung“ ist eine Dimension, die sich aus dem Zusammenspiel von Kultur und Natur, Disposition und Lernen, individueller Anstrengung und pädagogischer Förderung aufbaut. *Individualisierung*, nicht *Egalisierung* ist dafür die Leitlinie.

*Gleichheit*, die gesellschaftliche Prämisse für diese Arbeit, tritt aber ergänzend, wen auch spannungsreich, als Kriterium hinzu. Sie hat insofern zu gelten, als der Prozess des Heranwachsenden und die Praxis der pädagogischen Arbeit zwei Gütekriterien genügen müssen: Der Prozess obligatorischen schulischen Lernens soll, folgt man dem Selbstverständnis moderner demokratischer Gesellschaften, einerseits für alle Heranwachsenden zu Ergebnissen führen, die ihnen die selbständige Teilhabe an Gesellschaft und Kultur ermöglichen, und er soll zugleich einen Mindeststandard an kultureller Gemeinsamkeit, die „Basisfähigkeiten“, sichern, auf die alle Gesellschaften angewiesen sind – das ist die Idee der grundlegenden Allgemeinbildung; der Prozess muss

andererseits so strukturiert sein, dass die Kompetenz zur Teilhabe an Gesellschaft und Kultur unabhängig von der sozialen Herkunft erworben und kultiviert wird. Zumindest muss man erwarten können, dass weder sozialer Status noch Geschlecht noch Konfession noch Region die Eigenanstrengung der Lernenden so überformen, dass der Lernprozess selbst die Ungleichheit der Herkunft oder des Status bekräftigt.

Erkennbar sind beide Prämissen moderner Bildungsarbeit – die anthropologische wie die soziale, das Gebot der Individualisierung wie das der Gleichheit – in kritischer Funktion stark, konstruktiv aber schwach: Bisher ist keine Pädagogik entwickelt worden, die ohne Brüche und Widersprüche beiden Prämissen, den Prinzipien der Egalisierung und der Individualisierung, in gleicher Weise und zugleich hätte entsprechen können. Der Zielkonflikt ist unübersehbar. Es bedarf daher der politischen Entscheidung, um hier Prioritäten zu setzen, und es bedarf einer Pädagogik, die sich an der Idee der Kompetenzentwicklung und des Kompetenzaufbaus orientiert, um praktisch trotz des Zielkonflikts handlungsfähig zu sein.

## (2) Offenheit der Zukunft für Individuen und Gesellschaft

Der Unentschiedenheit in den anthropologischen und sozialen Voraussetzungen korrespondiert in der Zeitdimension die Offenheit der Zukunft, auf die Erziehung und Bildung vorbereiten.

Bildungsziele werden deshalb häufig unter dem Anspruch formuliert, auf diese Zeitproblematik in eigener Weise zu reagieren, also z.B. den sozialen Wandel zu berücksichtigen, Zukunft zu antizipieren und je aktuell Antworten auf die Herausforderungen zu eröffnen, die sich im historischen Prozess ergeben. Sie sollen Rücksicht nehmen auf die gewandelte Lebenslage, wie sie die Heranwachsenden in Familie, Kindheit und Jugendalter, aber auch in Arbeits- und Berufsverhältnissen in raschem Wechsel erleben.

Bildungsziele und schulische Arbeit werden als Antworten auf gesellschaftliche Probleme verstanden, aber auch als angemessene Reaktion auf den technologischen Wandel oder auf die Veralterung des Wissens und den rapiden Zuwachs an Technologien. Damit sollen die Pädagogen die Zukunft vorbereiten, die sich nur diffus abzeichnet. So plausibel und verständlich solche Erwartungen sind, erkennbar sind nicht nur Pädagogen mit solchen Erwartungen überfordert. Die Pädagogen sind von dieser Offenheit aber deswegen stärker betroffen, weil sie eine Praxis konstruieren müssen, die auch schon hier und jetzt, in der Gegenwart der Lernenden, sinnvoll sein muss und Motivation nicht durch den Verweis auf ferne Zukünfte aufbauen und bewahren kann.

## (3) Unbestimmtheit der Aufgaben und Anforderungen

Schulische Arbeit kann sich schon deswegen nicht einfach an einer gegebenen Struktur von zukünftigen Aufgaben – in Arbeit und Beruf oder in Kultur, Gesellschaft und Politik – orientieren und darauf vorbereiten, weil diese Aufgaben selbst nicht gleich bleiben. Zwar lassen sich generelle

Erwartungen identifizieren, die in allen modernen Gesellschaften für Leben und Arbeit nach der Phase des Lernens warten, die Diskussion über Schlüsselqualifikationen hatte hier ihren sinnvollen Anknüpfungspunkt. Gleichzeitig aber gilt schon spätestens seit dem 20. Jahrhundert die Erfahrung, dass schulisch erworbene Qualifikationen im Lebenslauf nicht hinreichen und dass auch die Orientierung an Schlüsselqualifikationen entweder dürre Abstrakta erzeugt hat oder Scheinsicherheit. Der Berufswechsel z.B. ist ebenso alltäglich wie die Notwendigkeit, bereit zur Weiterbildung zu sein oder angesichts politischer Konflikte und des raschen kulturellen oder sozialen Wandels durch Anpassung an das Neue handlungsfähig zu bleiben. Bildungsziele müssen also zumindest so formuliert sein, dass sie die Zukunft nicht versperren, die sie eröffnen wollen.

#### (4) Pluralität und Konflikthaftigkeit der Erwartungen

Erschwerend kommt hinzu, dass nicht allein Experten über mögliche Zukünfte und den Wandel von Aufgaben forschen und streiten, sondern Politik, Gesellschaft, Individuen und Gruppen selbst unterschiedliche Vorstellungen und Wünsche über die denkbaren Zukunftsoptionen haben. Das sind in der Regel Optionen, die grundlegende gesellschaftliche Konflikte einschließen und die sich nicht einfach technisch überspielen oder ungefiltert in Lernprozesse, Bildungsziele oder Curricula umsetzen lassen. Bereits auf der Ebene der Bildungsziele spiegeln sich diese Kontroversen, grundlegend und konkret. Man kann Welt- und Menschenbilder identifizieren, die hier wirken und fortwirken, politische Ideologien ebenso wie Herrschaftsgelüste, private Wünsche an unterschiedliche Lebensentwürfe ebenso wie konkurrierende Lebensstile und ihre Normen und Werte. Die Bestimmung von Bildungsstandards kann deshalb auch als Versuch interpretiert werden, unbeschadet aller Konflikte einen Bereich von Konsens und Übereinstimmung zu bezeichnen.

#### (5) Utopieüberschuss und Realisierungsprobleme

Es spiegelt die Bedeutsamkeit der Schule als einer Form, der Zukunft bereits in der Gegenwart schon Gestalt zu geben, dass sie immer neu mit diesen Problemen von Unentscheidbarkeit, Offenheit, Unbestimmtheit, Pluralität und Konflikthaftigkeit konfrontiert wird. Es macht freilich auch die Schwäche der Schule und der bildungstheoretischen und bildungspolitischen Debatten aus, dass sie auf die Komplexität und Diffusion der Erwartungen und Ziele, die aus dieser Situation resultieren, eher mit der Übersteigerung von Erwartungen und der Überbietung von Ansprüchen als mit Nüchternheit und Pragmatik reagieren. Der bekannte Überdruß an den einschlägigen pädagogischen Debatten rührt ja auch daher, dass zwischen utopischen Entwürfen und realen Möglichkeiten und Leistungen der Schule ein nicht übersehbares, ja ein schreiendes Missverhältnis existiert. Schule ist geradezu der exemplarische Ort, um den Umschlag von Utopie in Ideologie sichtbar zu machen – und die Formulierung von Bildungszielen ist anscheinend die gesellschaftliche Form, in der man dieses Missverhältnis auf Dauer stellt.

### **5.3 Bildungsstandards und Kompetenzmodelle – die pragmatische Antwort auf die Probleme der Bildungsziele**

Wo können in dieser systematisch schwierigen Situation Lösungswege gefunden werden, und zwar solche, die nicht allein eine abgehobene Zieldebatte wiederbeleben oder stärken, sondern für die Organisation von Bildungsprozessen und für die Konkretisierung von Erwartungen an ihre Ergebnisse, sinnvoll sind? Wir sehen grundsätzlich zwei Wege: (1) den Weg der Identifikation eines latenten Konsenses einerseits, (2) die zielbezogene Orientierung der pädagogischen Arbeit an Kompetenzentwicklung und Kompetenzmodellen andererseits.

#### **(1) Kulturelle Selbstverständlichkeiten und der Konsens über Basisfähigkeiten**

Trotz aller systematischen Probleme und unterhalb der aktuellen Kontroversen über Bildungsziele darf man zunächst eine Zone des Konsenses in der Bildungsarbeit nicht übersehen, die auch nicht allein durch das Grundgesetz oder die Verfassungen der Länder definiert wird. Die in der Praxis geltenden und, vor allem, die wirksamen Normen der Bildungsarbeit beschreiben trotz der Pluralität, die sie spiegeln, in allen modernen Gesellschaften ein vergleichbares Bündel von Themen und Vorgaben für den Prozess öffentlich organisierter Bildung. Es gibt für die Funktion allgemeiner Bildung und für die Erwartungen an ihre Qualität einen praktischen Konsens, den man jenseits der immer neuen Kontroversen nicht übersehen sollte. Dieser Konsens bezieht sich sowohl, gesellschaftlich gesehen, auf die Erwartung, dass das Bildungssystem mit daran arbeitet, auf die Staatsbürgerrolle vorzubereiten, also zur Teilhabe am öffentlichen Leben zu befähigen, als auch, im Blick auf die Subjekte, auf die Erwartung, dass im Bildungswesen die Fähigkeiten erworben werden, das eigene Leben als Lernprozess selbst gestalten zu können, trotz der Unsicherheit von Beruf und Arbeit, Karriere und sozialer Lage.

Generell lässt sich damit erkennen, dass Bildungsziele die Beziehungen von Individuen und Gesellschaften normieren. Sie bestimmen Anspruch und Form der Vergesellschaftung und die Rolle der Subjekte. Für moderne, der Tradition der Aufklärung verpflichtete und demokratisch organisierte Gesellschaften gilt dann ein Bild von Individualität als leitend, in dem – wie es das Grundgesetz sagt – die Würde des Menschen und die freie Entfaltung der Persönlichkeit oberste Maximen sind. Zu allgemeinen Bildungszielen werden diese Prämissen, weil erst im Prozess des Aufwachsens zu sichern ist, dass alle Heranwachsenden einer Generation, und zwar unabhängig von Herkunft und Geschlecht, dazu befähigt werden, in der selbständigen Teilhabe an Politik, Gesellschaft und Kultur und in der Gestaltung der eigenen Lebenswelt diesem Anspruch gemäss zu leben und als mündige Bürger selbstbestimmt zu handeln.

Historisch variabel und kontrovers sind solche allgemeinen Bildungsziele erst in der konkreten Ausgestaltung von Lernprozessen. Systematisch gesehen treten dann die Spannungen zwischen Subjektwerdung und Vergesellschaftung zutage, konkret werden z.B. die Tugenden diskutiert, die

für den Alltag generalisiert werden müssen, damit die individuelle Freiheit verantwortlich wahrgenommen wird. Der Konflikt zwischen eher konservativen oder eher kulturinnovativen Positionen ist damit vorgezeichnet, „die verfassungsmäßige Ordnung oder das Sittengesetz“ bezeichnen innerhalb des Grundgesetzes solche Schranken, die trotz aller Kontroversen gelten. In der pädagogischen Diskussion bleibt dann aber immer noch kontrovers, ob z.B. die sog. „Sekundärtugenden“ Geltung beanspruchen können und damit die Frage, wieweit z.B. Tüchtigkeit, Ordnungsliebe, Sauberkeit oder Pünktlichkeit zu legitimen Zielen der Bildungsarbeit werden können. Das Bildungswesen reagiert auf solche Kontroversen aber nicht mit Vereinheitlichung, sondern macht die Kontroversen selbst zum Thema: Werterziehung wird dann selbst ein – in ihren Möglichkeiten kontrovers diskutierter – Wert und ein Ziel schulischer Arbeit und das Ziel ist nicht die Vermittlung von Doktrinen, sondern die Befähigung der Lernenden zu selbständiger, kritischer Entscheidung angesichts von Konflikten über Ziele und Werte – und in diesem Ziel besteht wiederum Konsens.

Dieser Konsens schließt auch eine Übereinstimmung hinsichtlich der Basisfähigkeiten ein, die vorausgesetzt sein müssen, damit überhaupt der offene Umgang mit Mensch und Welt vollzogen werden kann. Niemand hat je bestritten, dass die kompetente Verfügung über Kulturtechniken – des Lesens, Schreibens, des Rechnens, des Umgangs mit Texten, der Rücksicht auf die Mitmenschen oder den eigenen Körper – zu den Basisfähigkeiten gehört, die im Prozess des Aufwachsens gesichert werden müssen. Das geschieht schon in Familien und in den Altersgruppen der Heranwachsenden, aber die Schule ist notwendig, wenn systematisches Lernen die Voraussetzung zur Beherrschung und Nutzung solcher Basisfähigkeiten ist, wie beim Lesen oder Schreiben oder, aktuell, beim kompetenten Umgang mit Hard- und Software in einer medial bestimmten Kultur. Der Dissens beginnt erst dann, wenn man nicht allein die Fähigkeit zum Gebrauch generalisieren, sondern auch noch die Nutzung der Basisfähigkeiten inhaltlich normieren will, so dass z.B. nur gelesen werden soll, was gelesen werden darf. Aber das, die Beherrschung der Basisfähigkeiten in einem einzigen und dann noch moralisch oder normativ eindeutigen Sinne, das können die Pädagogen nicht garantieren.

Das Problem der Messung von Leistungen des Bildungssystems bleibt dabei erhalten und es muss auch gelöst werden, denn offenbar ist nicht einmal die Verfügung über Basiskompetenzen hinreichend sicher erwartbar. Diese Fähigkeit zum Gebrauch der Basisfähigkeiten kann man auch sehr gut messen, so dass auch Verständigung über die Leistung der Schule an diesem Punkt – im Bereich der Konsenszone der Erwartungen – möglich ist. Auch anspruchsvolle und ambitionierte Bildungsziele – „Mündigkeit“ etwa, also die Fähigkeit zur selbstbestimmten Teilhabe an einer demokratischen Gesellschaft – entziehen sich zwar nicht grundsätzlich einer empirischen, an Messprozeduren ansetzenden Diskussion. Aber es gibt bisher keinen Konsens über die angemessene Operationalisierung und die alte Hoffnung curriculumtheoretischer Debatten ist zerstört, als ließen sich aus allgemeinen Bildungszielen in eindeutiger Weise konkrete, am Verhalten ohne Widerspruch messbare Lernziele ableiten.

Die politische Debatte kann deshalb Konsens nur über die Bedeutsamkeit der Dimensionen schulischer Arbeit erzeugen, weil niemand ernsthaft bestreiten kann, dass Mündigkeit in einer Demokratie ebenso unverzichtbare Erwartung an das Handeln ihrer Akteure ist wie die Verfügung über Basisfähigkeiten. Der Rest bleibt kontrovers und die interessante Frage ist, wie man dennoch in der Diskussion über die Qualität von Bildungsprozessen weiter kommt. Hier setzte traditionell der öffentliche Diskurs über Bildung an, aktuell wird er fortgesetzt in der Orientierung an Kompetenzmodellen.

## (2) Kompetenzentwicklung und -modelle als Lösungsweg in einer normativ konflikthaften Situation

Blickt man auf die Funktion und die Möglichkeiten von Schule in modernen Gesellschaften, dann sind die Fähigkeiten, die über schulische Arbeit vermittelt und in der Gesellschaft als universelle Prämissen für Kommunikation vorausgesetzt werden, offenbar Fähigkeiten, die erst angesichts systematisch kontroverser Außenerwartungen ihre besondere Qualität gewinnen, d.h. Bedeutsamkeit auch jenseits der Konsenszone und trotz der Begründungsschwierigkeiten für Bildungsziele. „Bildung“, die traditionelle deutsche Generalformel für die Erwartungen an (lebenslange, nicht allein schulische) Lernprozesse, bezeichnet bereits exakt die Fähigkeit von Subjekten, unter den Bedingungen der Unentscheidbarkeit und Unbestimmtheit, Offenheit und Pluralität handlungsfähig zu sein; die Erwartung an den Bildungsprozess war es, für ein Handeln unter solchen Bedingungen mündig zu werden, und zwar in einem Prozess, der die Selbstkonstruktion des Subjekts angesichts gesellschaftlich unausweichlicher Erwartungen ermöglicht, Gleichheit und Individualisierung also zugleich eröffnen soll.

„Kompetenzen“ beschreiben aber nichts anderes, also solche Fähigkeiten der Subjekte, die auch der Bildungsbegriff gemeint und unterstellt hatte: Erworbene, also nicht von Natur aus gegebene Fähigkeiten, die an und in bestimmten Dimensionen der gesellschaftlichen Wirklichkeit erfahren wurden und zu ihrer Gestaltung geeignet sind, Fähigkeiten zudem, die der lebenslangen Kultivierung, Steigerung und Verfeinerung zugänglich sind, so, dass sie sich intern graduieren lassen, z.B. von der grundlegenden zur erweiterten Allgemeinbildung; aber auch Fähigkeiten, die einen Prozess des Selbstlernens eröffnen, weil man auf Fähigkeiten zielt, die nicht allein aufgaben- und prozessgebunden erworben werden, sondern ablösbar von der Ursprungssituation, zukunftsfähig und problemoffen.

Man erkennt damit unschwer, dass auch Kompetenzmodelle, theoretische Beschreibungen der Struktur dieser spezifischen Fähigkeiten und der Stufen ihres Erwerbs, als Antwort auf typisch moderne Problemlagen konzipiert sind. Betrachtet man nämlich ihren (eher impliziten als expliziten) normativen Gehalt, dann reagieren sie auf die Schwierigkeiten der Debatte über Bildungsziele nicht durch Fixierung auf eine Gegenwart, sondern mit Offenheit für den Prozess; und auf den notwendigen Zwang zur Revision im Prozess reagieren sie mit Reflexivität: Wie der Bildungsbe-

griff es schon intendierte und im Begriff des Gebildeten noch schichtspezifisch stilisierte, kann deshalb auch in Kompetenzmodellen als Kriterium der Graduierung ebenfalls unterstellt werden, dass die Verfügung über die intendierten Fähigkeiten erst vollständig ist, wenn sie die Fähigkeit zum Umlernen und Neulernen, systematisch gesehen die Fähigkeit zur kritischen Beobachtung der Welt und des eigenen Lernens einschließt.

Die Theorie von Kompetenz und Kompetenzentwicklung nimmt diese interne Unterscheidung also nur insofern auf, als sie in eigenen Modellen der Graduierung Stufen der Kompetenz ebenso unterscheidet wie eine Sequenz des Erwerbs. Parallel zur alten Bildungstheorie ist dabei der Grad der Reflexivität der Problembearbeitung, also die kritische Beobachtung des eigenen Lernens und der Welt, der Aufgaben und der Lösungsprozesse, der beste Indikator, um solche Graduierung zu konstruieren und für konkrete Handlungsbereiche des Menschen zu entwickeln. Die „Domänen“ des Wissens und Handelns, in denen sich Kompetenzen entwickeln, lassen sich deshalb durchaus und ohne systematischen Bruch den Dimensionen einer modernen Allgemeinbildung vergleichen, Dimensionen, die ihren klassischen Ursprung bis heute nicht verbergen können.

Die „Basisfähigkeiten“, die man von allgemeiner Bildung heute erwarten darf, lassen sich auch ganz konkret nennen: Sie beziehen sich auf die Fähigkeit, an Gesellschaft selbstbestimmt teilzunehmen, die unterschiedlichen Dimensionen des Handelns – moralische, kognitive, soziale und individuelle – in ihrer je eigenen Bedeutung zu sehen und nutzen sowie das eigene Handeln an einem allgemeinen Gesetz ausrichten zu können. Bezogen auf den Lebenslauf wird zugleich unterstellt, dass alle Heranwachsenden dabei auch fähig werden, mit neuen Herausforderungen, einer ungewissen Zukunft und alternativen Optionen in der Gestaltung des eigenen Lebens im Modus des Lernens umzugehen. Das „Lernen des Lernens“ ist die grundlegende, für das gesamte Leben unverzichtbare Kompetenz, die in modernen, offenen Gesellschaften in schulischer Arbeit generalisiert werden muss.

Die allgemeine Formel vom „Lernen des Lernens“ verlangt aber nach inhaltlicher Konkretisierung: die Beherrschung der grundlegenden Kulturtechniken – Lesen, Schreiben, Rechnen, wie sie die unterschiedlichen Dimensionen im literacy-Konzept allgemeiner Bildung festhalten und in den PISA-Studien gemessen wurden – beschreibt dann die basale Ebene. Aber als generelle Prämisse für die Teilhabe an gesellschaftlicher Kommunikation reicht die Dimension der einfachen Kulturtechniken nicht mehr aus. Die Heranwachsenden müssen vielmehr fähig werden für den Gebrauch der Computer, für den Umgang mit Medien, für die Herausforderungen einer multikulturellen Welt, und sie müssen zugleich in der Form der Welterfahrung von den einfachen Formen des Ich-zentrierten Umgangs mit Welt auf die grundlegenden wissenschaftlichen Modi der Welterfahrung übergehen können.

Die klassische Bildungstheorie hat vier solcher Modi der Welterfahrung angeboten und als Kern ihrer Idee der allgemeinen Bildung und der für sie notwendigen „Kenntnisse“ festgehalten: „historische“, „mathematische“, „linguistische“ und „ästhetisch-expressive“. Man erkennt in diesen Modi

der Welterfahrung die grundlegenden Dimensionen der historisch-gesellschaftlichen, mathematisch-naturwissenschaftlichen, mutter- und fremdsprachlichen sowie ästhetischen Bildung, wie sie die Grundstruktur des Lehrplans moderner Schulen bis heute bestimmen und die Fächerauswahl und das Fächerangebot organisieren. Überwölbt waren diese Dimensionen durch den Anspruch, dass der Unterricht insgesamt „philosophisch“ zu sein habe, kritisch gegen sich selbst. Damit bezeichnen die klassischen Dimensionen allgemeiner Bildung aber exakt schon die Dimensionen des Wissens und Könnens, der Kompetenzen mithin, die ein „Kerncurriculum moderner Allgemeinbildung“ (Tenorth 1994, Baumert 2002b, bes. S. 113) z.B. auch heute nennt, um die Bezugsfelder und das kanonische Orientierungswissen auszuweisen, für die „Kulturwerkzeuge“ als „basale Sprach- und Regulationskompetenzen“ erworben werden müssen.

| <b>Modi der Weltbegegnung<br/>(Kanonisches Orientierungswissen)</b>                                                                      | <b>Basale Sprach- und Selbstregulationskompetenzen (Kulturwerkzeuge)</b> |                                   |                                |                     |                                            |
|------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|-----------------------------------|--------------------------------|---------------------|--------------------------------------------|
|                                                                                                                                          | <b>Beherrschung der Verkehrssprache</b>                                  | <b>Mathematisierungskompetenz</b> | <b>Fremdsprachl. Kompetenz</b> | <b>IT-Kompetenz</b> | <b>Selbstregulation des Wissenserwerbs</b> |
| <b>Kognitiv-instrumentelle Modellierung der Welt</b><br>Mathematik<br>Naturwissenschaften                                                |                                                                          |                                   |                                |                     |                                            |
| <b>Aesthetisch-expressive Begegnung und Gestaltung</b><br>Sprache/Literatur<br>Musik/Malerei/Bildende Kunst<br>Physische Expression      |                                                                          |                                   |                                |                     |                                            |
| <b>Normativ-evaluative Auseinandersetzung mit Wirtschaft und Gesellschaft</b><br>Geschichte<br>Ökonomie<br>Politik/Gesellschaft<br>Recht |                                                                          |                                   |                                |                     |                                            |
| <b>Probleme konstitutiver Rationalität</b><br>Religion<br>Philosophie                                                                    |                                                                          |                                   |                                |                     |                                            |

Abb. 5: Grundstruktur der Allgemeinbildung und des Kanons in Anlehnung an Baumert (2002).

Kompetenztheoretisch begründete „Bildungsstandards“ werden deshalb auch mit guten Gründen nicht als allgemeine Bildungsziele formuliert, sondern als bereichsspezifische Leistungserwartungen. Sie können in einem eigenen politischen und theoretischen Diskurs aber auf die allgemeinen Erwartungen begründet rückbezogen und an ihnen geprüft, aber weder aus diesen abgeleitet noch mit ihnen gleichgesetzt werden.

## 5.4 Legitimationsprobleme: Öffentlicher Diskurs und Konstruktionen der Experten

Allgemeine Bildungsziele behalten ihre eigene Funktion, aber die ist nicht so sehr konstruktiv oder operativ, sondern kritisch und begleitend. Hartmut von Hentig hat in seinen Reflexionen über „Bildung“ jüngst vorgeschlagen (Hentig 1996, bes. S. 71ff.), solche allgemeinen Ziele nicht als messbare Vorgaben für Bildungsprozesse, sondern als Kriterien zur Prüfung ihrer Geltung zu interpretieren. Sie sollen als „Maßgaben“ fungieren, mit deren Hilfe sich die Wünschbarkeit und Legitimität von Bildungsprozessen und ihren Ergebnissen im Medium der allgemeinen, politisch-gesellschaftlichen Verständigung grundlegend diskutieren lassen.

Diese „möglichen Maßstäbe“ heißen für Hentig im Blick auf die Bildung des Subjekts und seine Fähigkeiten: „Abscheu und Abwehr von Unmenschlichkeit; die Wahrnehmung von Glück; die Fähigkeit und den Willen, sich zu verständigen; ein Bewusstsein von der Geschichtlichkeit der eigenen Existenz; Wachheit für letzte Fragen; und – ein doppeltes Kriterium – die Bereitschaft zu Selbstverantwortung und Verantwortung in der *res publica*.“ Hentig weiß, dass diese „Bildungskriterien“ nicht messbar sind, er unterstellt auch nicht, dass sie als Leistung von der Schule allein erwartet werden können, weil Schule nur einen der „Lernanlässe“ des Lebens darstellt, aber dennoch will er diese Kriterien nutzen, und zwar unter der Frage: „Was auch immer den Menschen bildet – verändert, formt, stärkt, aufklärt, bewegt –, ich werde es daran messen, ob dies eintritt“, und dieses „dies“ ist mit seinen „Maßstäben“ bezeichnet.

Mit solchen Ansprüchen hat man aber die Diskussion über messbare und basale Bildungsstandards verlassen und das Feld der allgemeinen Bildungsziele betreten. Der Ort der Diskussion solcher „Bildungskriterien“ ist – mit Hentig gesprochen – die *polis* selbst, die aufgeklärte, moralisch verantwortlich argumentierende Öffentlichkeit. Der politische Diskurs mit seinen eigenen Bedingungen und Konflikten bildet damit den Rahmen, in den Bildungsstandards eingebunden sind, aber er kann ihn weder systematisch bestimmen noch die Messung der Leistung der Bildungsarbeit ersetzen. Allerdings ist damit das Medium der diskursiven Legitimation bezeichnet, auf das Bildungsstandards angewiesen sind, wenn sie nicht allein der Legitimation durch Verfahren ihre Geltung verdanken sollen, die sich aus dem Prozess ihrer politischen Konstruktion ergibt.

Bildungsstandards müssen sich dann an der Frage messen lassen, ob sie den Zugang zu solchen allgemeinen Bildungszielen und der Diskussion dieser Kriterien eröffnen. Kompetenztheoretisch definierte Bildungsstandards können das schon deswegen, weil sie erkennbar der Logik bildungstheoretischer Debatten selbst folgen und sich in einem modernen Kerncurriculum allgemeiner Bildung systematisch verorten lassen.

Sie werden damit in zwei Richtungen anschlussfähig und produktiv: Für den bildungstheoretischen und bildungs- wie gesellschaftspolitischen Diskurs unterbieten sie die Dimension des Problems nicht, dass Bildung zwar allein als Leistung des Subjekts möglich und denkbar ist, aber zugleich die Erwartungen festhält, die Gesellschaften mit ihrer eigenen Zukunft und den Möglich-

keiten der Individuen verbinden. Für die Frage nach der spezifischen Rolle der Schule in diesem Prozess können Bildungsstandards zugleich eine kriteriengestützte Operationalisierung und Messung von Erwartungen an die Lernenden und die Qualität pädagogischer Arbeit anbieten und dabei auch die Bereiche des Wissens und Könnens ausweisen, an denen solche Qualität lehrbar und steigerbar wird. Anders als die nur philosophische und politische Debatte können solche Modelle aber auch den realistischen Blick auf die Leistungen des Bildungssystems anbahnen, der heute unverzichtbar ist.

## 6. Kompetenzmodelle aus pädagogisch-psychologischer Sicht

Eine wichtige Aufgabe von Bildungsstandards, wie sie in dieser Expertise verstanden werden, ist es, die Kompetenzen zu benennen, die Schülerinnen und Schülern erwerben müssen, damit Bildungsziele, wie sie im vorangegangenen Kapitel 5 dargestellt wurden, als erreicht gelten können. Die Rolle von Kompetenzmodellen ist es, einerseits zu beschreiben, welche Lernergebnisse von Schülerinnen und Schülern in bestimmten Altersstufen in den jeweiligen Fächern erwartet werden, und andererseits wissenschaftlich fundiert aufzuzeigen, welche „Wege zum Wissen und Können“ eingeschlagen werden können. Kompetenzmodelle stellen damit die Grundlage für Operationalisierungen von Bildungszielen dar, die den Output des Bildungssystem über das Erstellen von Testverfahren (siehe Kapitel 7) empirisch zu überprüfen erlauben.

Die Rolle von Kompetenzmodellen besteht darin, zwischen abstrakten Bildungszielen und konkreten Aufgabensammlungen zu vermitteln. Modellvorstellungen über den Erwerb von Kompetenzen helfen nicht nur bei der Gestaltung von Testverfahren, sondern bieten zugleich Anhaltspunkte für eine Unterrichtspraxis, die an den Lernprozessen und Lernergebnissen der Schülerinnen und Schüler im jeweiligen Lernbereich orientiert ist und nicht allein an der fachlichen Systematik von Lehrinhalten.

Die theoretische Analyse und Modellierung wie auch die empirische Untersuchung von Kompetenzen ist ein sehr aktives Feld psychologischer und pädagogischer Forschung. In den vergangenen Jahren ist es gelungen, ein differenzierteres Verständnis des Aufbaus, der Entwicklung und der schulischen Förderung von bereichsspezifischen Kompetenzen zu erzielen. Diese Forschungsergebnisse können die Entwicklung von Bildungsstandards stützen. Gleichwohl liegen ausgearbeitete, empirisch gestützte Kompetenzmodelle nur für einzelne Lernbereiche, Altersgruppen und Schülerpopulationen vor. Mit am differenziertesten ist die Kenntnislage in der Mathematik. Man wird sich daher bei der Arbeit an den Bildungsstandards in vielen Fällen primär auf das Erfahrungswissen der Fachdidaktiken stützen, deren Vorstellungen zu Struktur und Entwicklung von Schülerkompetenzen fachsystematische, psychologische und unterrichtspraktische Aspekte verbinden. Die beiden im Anhang dargestellten Ansätze – der Gemeinsame Europäische Referenzrahmen für Sprachen und das kanadische Rahmenkonzept für naturwissenschaftliche Bildung – stellen solche pragmatischen Ansätze dar.

## 6.1 Der Kompetenzbegriff

In einer Überblicksarbeit zeigte der Erziehungswissenschaftler und Psychologe Franz Weinert (1999), dass eine Vielzahl unterschiedlicher Kompetenzbegriffe verwendet wird, die eine weite Spanne abdeckt von angeborenen Persönlichkeitsmerkmalen (z.B. Begabung, Intelligenz) bis hin zu erworbenem umfangreichem Wissensbesitz, von fächerübergreifenden Schlüsselqualifikationen bis hin zu fachbezogenen Fertigkeiten. Soll der Kompetenzbegriff zur Grundlage für bildungspolitische Veränderungen gemacht werden, wie dies bei der Entwicklung von Bildungsstandards der Fall ist, ist eine Übereinkunft im Sprachgebrauch notwendig. Weinert argumentierte überzeugend, dass die tragfähigste Definition von Kompetenz diejenige ist, die in dem Bereich der Expertiseforschung entwickelt wurde. Die Expertiseforschung beschäftigt sich mit der Untersuchung von leistungsfähigen Experten in einem bestimmten Fach bzw. Gegenstandsbereich – in der Expertiseforschung als „Domäne“ bezeichnet. Der dort verwendete Kompetenzbegriff lässt sich hervorragend auf den schulischen Bereich übertragen. Bei der Beschreibung von Kompetenz und vor allem bei Versuchen ihrer Operationalisierung stehen hauptsächlich kognitive Merkmale (fachbezogenes Gedächtnis, umfangreiches Wissen, automatisierte Fertigkeiten) im Vordergrund. Jedoch gehören ausdrücklich auch motivationale und handlungsbezogene Merkmale zum Kompetenzbegriff.

In Übereinstimmung mit Weinert (2001, S. 27f.) verstehen wir unter *Kompetenzen die bei Individuen verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten, bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können.*

Kompetenz ist nach diesem Verständnis eine Disposition, die Personen befähigt, bestimmte Arten von Problemen erfolgreich zu lösen, also konkrete Anforderungssituationen eines bestimmten Typs zu bewältigen. Die individuelle Ausprägung der Kompetenz wird nach Weinert von verschiedenen Facetten bestimmt:

- Fähigkeit
- Wissen
- Verstehen
- Können
- Handeln
- Erfahrung
- Motivation

Das Zusammenspiel dieser Facetten soll am Beispiel einer Fremdsprachenkompetenz illustriert werden, bei der die kommunikative Handlungsfähigkeit als Bildungsziel vorgegeben ist. Die so verstandene Fremdsprachenkompetenz drückt sich darin aus,

- wie gut man kommunikative Situationen bewältigt (Handeln und Erfahrung),
- wie gut man Texte unterschiedlicher Art versteht (Verstehen) und
- selbst adressatengerecht Texte verfassen kann (Können),
- aber unter anderem auch in der Fähigkeit, grammatische Strukturen korrekt aufzubauen und bei Bedarf zu korrigieren (Fähigkeit und Wissen),
- oder in der Intention und Motivation sich offen und akzeptierend mit anderen Kulturen auseinander zu setzen (Motivation).

Standards für das Fremdsprachenlernen, die dem Bildungsziel der kommunikativen Handlungsfähigkeit verpflichtet sind, müssen alle diese Facetten berücksichtigen (vgl. Anhang a). Hierbei spielen nicht nur kognitive Wissensinhalte eine Rolle sondern – wie etwa die zuletzt genannte „interkulturelle Offenheit“ deutlich macht – auch Einstellungen, Werte und Motive.

Für die Entwicklung von Kompetenzmodellen und die daran anschließende Erarbeitung von illustrativen Aufgabenbeispielen und Testverfahren ergeben sich einige wichtige Folgerungen:

- Kompetenz kann nur leistungsbezogen erfasst und gemessen werden. Kompetenz stellt die Verbindung zwischen Wissen und Können (siehe unten) her und ist als Befähigung zur Bewältigung von Situationen bzw. von Aufgaben zu sehen. Jede Illustration oder Operationalisierung einer Kompetenz muss sich daher auf konkrete Anforderungssituationen beziehen.
- Kompetenzen kann man nicht durch einzelne, isolierte Leistungen darstellen oder erfassen. Der Bereich von Anforderungssituationen, in denen eine bestimmte Kompetenz zum Tragen kommt, umfasst immer ein mehr oder weniger breites Leistungsspektrum. Die Entwicklung und Förderung von Kompetenzen muss daher eine ausreichende Breite von Lernkontexten, Aufgabenstellungen und Transfersituationen umschließen. Entsprechend breit muss auch die Darstellung der Kompetenzen in Bildungsstandards und ihre Umsetzung in Aufgaben und Tests gestaltet sein. Eine eng gefasste Leistungserfassung kann dem Anspruch von Kompetenzmodellen nicht gerecht werden. Gerade wenn die sieben oben genannten Facetten von Kompetenz berücksichtigt werden, wird erkennbar, dass die Erfassung von Kompetenz einen breiten Begriff von Aufgaben bzw. Tests erfordert, der nicht mit Wissensabfragen allein erschöpft ist.

## 6.2 Kompetenzmodelle

Kompetenz als Vermittler von Bildungszielen, konkreten Aufgabenstellungen und Anforderungen enthält eine Vielzahl von Komponenten, die zusammen wirken müssen. „Kompetenzmodelle“ erfüllen in Bezug auf Bildungsstandards zwei Zwecke: erstens beschreiben sie das Gefüge der Anforderungen, deren Bewältigung von Schülerinnen und Schülern erwartet wird (Komponentenmodell); zweitens liefern sie wissenschaftlich begründete Vorstellungen darüber, welche Abstufungen eine Kompetenz annehmen kann bzw. welche Grade oder Niveaustufen sich bei den einzelnen Schülerinnen und Schülern feststellen lassen (Stufenmodell).

Kompetenzmodelle sollten alle sieben der unter 6.1 genannten Facetten berücksichtigen. Demnach kann von Kompetenz dann gesprochen werden,

- wenn gegebene Fähigkeiten der Schülerinnen und Schüler genutzt werden,
- wenn auf vorhandenes Wissen zurückgegriffen werden kann bzw. die Fertigkeit gegeben ist, sich Wissen zu beschaffen,
- wenn zentrale Zusammenhänge der Domäne verstanden werden,
- wenn angemessene Handlungsentscheidungen getroffen werden,
- wenn bei der Durchführung der Handlungen auf verfügbare Fertigkeiten zurückgegriffen wird,
- wenn dies mit der Nutzung von Gelegenheiten zum Sammeln von Erfahrungen verbunden ist und
- wenn aufgrund entsprechender handlungsbegleitender Kognitionen genügend Motivation zu angemessenem Handeln gegeben ist.

Aufgrund der zentralen Rolle fachbezogener Fähigkeiten und fachbezogenen Wissens sind Kompetenzen in hohem Maße domänenspezifisch.

Diese starke Ausrichtung des hier vertretenen Kompetenzbegriffs auf Lernbereiche, Fächer bzw. „Domänen“ mag Leser, die mit pädagogischen Debatten über Kompetenzförderung vertraut sind, überraschen, weil dort der Begriff der Kompetenz häufig für allgemeinere, fächer-übergreifende Fähigkeiten verwendet wird. Die pädagogisch-psychologische Forschung zeigt jedoch, dass es nicht ausreicht, fächerübergreifende „Schlüsselqualifikationen“ als Allheilmittel bzw. als eigenständige Zieldimensionen schulischer Bildung auszuweisen. Auch wenn Komponenten wie Methoden-, Personal- und Sozialkompetenz bedeutsam sind, ersetzen sie doch nicht die starke fachliche Bindung von Kompetenz. Die Forschung legt sogar nahe, dass die Entwicklung fächer-übergreifender Kompetenzen das Vorhandensein gut ausgeprägter fachbezogener Kompetenzen voraussetzt. Die Frage der Reichweite von Kompetenzmodellen ist daher nicht durch die Gegen-

überstellung von „fachbezogen“ versus „fächerübergreifend“ zu beantworten. Vielmehr stellen fachbezogene Kompetenzen eine notwendige Grundlage für fächerübergreifende Kompetenzen dar.

*Eine* Konsequenz ist, dass konkrete Ausformulierungen und Operationalisierungen des Kompetenzbegriffs zunächst in den Domänen bzw. Fächern zu erfolgen haben. Daraus begründet sich *weiterhin* die Notwendigkeit, bei der Entwicklung von Kompetenzmodellen auf dem Theorie- und Erkenntnisstand der Fachdidaktiken aufzubauen. Fachdidaktiken rekonstruieren Lernprozesse in ihrer fachlichen Systematik und zugleich in der je spezifischen, domänenabhängigen Logik des Wissenserwerbs und der Kompetenzentwicklung; beide Aspekte müssen bei der Darstellung von Komponenten und Kompetenzstufen berücksichtigt werden.

Neben der Komponentenstruktur und der Fach- bzw. Domänenspezifität von Kompetenzmodellen ist die Berücksichtigung von Kompetenzstufen ein zentraler Aspekt von Kompetenzmodellen. Diese Stufen bieten die Möglichkeit einer kriteriumsorientierten Interpretation von Testergebnissen und einer Verankerung von Mindeststandards (siehe unten Kapitel 7). Jede Kompetenzstufe ist durch kognitive Prozesse und Handlungen von bestimmter Qualität spezifiziert, die Schülerinnen und Schüler auf dieser Stufe bewältigen können, nicht aber auf niedrigeren Stufen.

Die Systematik der Beschreibung von Kompetenzstufen kann je nach Domäne sehr unterschiedlich aussehen. Im allgemeinen werden die Stufen Mischungen der oben genannten Facetten (Wissen, Können, Verstehen, Handeln, Motivation usw.) darstellen (vgl. unten, Abschnitt 6.3). Beispiele für Stufenmodelle liegen bei TIMSS und PISA vor (Klieme 2000; Klieme, Neubrand & Lüdtke 2001). Dabei ist die Grundvorstellung, dass sich die mathematische Kompetenz einer Person über Aufgaben beschreiben lässt, denen ein entsprechender Schwierigkeitsgrad zugeordnet werden kann. Auf der untersten Kompetenzstufe verfügen Personen über ein arithmetisches Wissen, das abgerufen und unmittelbar angewendet werden kann. Auf der obersten Kompetenzstufe werden hingegen komplexe Modellierungen und mathematische Argumentationen geleistet. Das im Folgenden wiedergegebene Modell wird bei Neubrand et al. 2002 sowie Knoche et al. 2002 weiter ausdifferenziert. Betrachtet man dieses Modell, so wird deutlich, was ein auf dem Erwerb von Kompetenzen basierendes Curriculum leisten muss. Es kann nicht darum gehen, kleinschrittig Inhalte zu beschreiben („sicheres Rechnen im Zahlenraum bis 20“), sondern es kommt darauf an, diese Inhalte in einen Anwendungskontext einzubinden.

Das Kompetenzstufenmodell von PISA (Klieme, Neubrand & Lüdtke 2001, S. 160) sieht folgende 5 Niveaus vor:

- *Stufe I: Rechnen auf Grundschulniveau*  
Personen, die dieser Stufe zugeordnet werden, verfügen lediglich über arithmetisches und geometrisches Wissen auf Grundschulniveau. Sie können dieses Wissen abrufen und unmittelbar anwenden, wenn die Aufgabenstellung von vornherein eine bestimmte Standard-Mathematisierung nahe legt. Begriffliche Modellierungen sind nicht leistbar.
- *Stufe II: Elementare Modellierungen*  
Auf dieser Stufe werden auch einfachste begriffliche Modellierungen vorgenommen, die in einem außermathematischen Kontext eingebettet sind. Personen auf dieser Kompetenzstufe können unter mehreren möglichen Lösungsansätzen den passenden finden, wenn durch Graphiken, Tabellen, Zeichnungen usw. eine Struktur vorgegeben ist, die das Modellieren erleichtert. Auch auf dieser Stufe sind allerdings nur die Wissensinhalte der Grundschulmathematik sicher verfügbar.
- *Stufe III: Modellieren und begriffliches Verknüpfen auf dem Niveau der Sekundarstufe I*  
Mit dieser Stufe findet im Vergleich zu Stufe II in mehrfacher Hinsicht ein qualitativer Sprung statt. Schülerinnen und Schüler auf dieser Kompetenzstufe verfügen auch über einfache Wissensinhalte der Sekundarstufe I, also über den Standardstoff der Lehrpläne aller Schulformen. Sie können Konzepte aus unterschiedlichen mathematischen Bereichen verknüpfen und zur Lösung von Problemstellungen nutzen, wenn visuelle Darstellungen den Lösungsprozess unterstützen.
- *Stufe IV: Umfangreiche Modellierungen auf der Basis anspruchsvoller Begriffe*  
Schülerinnen und Schüler auf dieser Kompetenzstufe bewältigen im technischen Bereich umfangreichere Verarbeitungsprozesse, können also eine Lösung über mehrere Zwischenergebnisse hinweg aufbauen. Auch offene Modellierungsaufgaben werden bewältigt, bei denen man unter vielfältigen Lösungswegen einen eigenen finden muss. Verstärkt können auch innermathematische begriffliche Zusammenhänge modelliert werden.
- *Stufe V: Komplexe Modellierung und innermathematisches Argumentieren*  
Auf dieser letzten Stufe ist auch anspruchsvolles curriculares Wissen verfügbar. Dies Schülerinnen und Schüler, die dieser Kompetenzstufe zugeordnet werden, können auch sehr offen formulierte Aufgaben bewältigen, bei denen ein Modell frei gewählt bzw. selbst konstruiert werden muss. Begriffliche Modellierungsleistungen auf dieser höchsten Stufe umschließen häufig Begründungen und Beweise sowie das Reflektieren über den Modellierungsprozess selbst.

Möglicherweise lassen sich die Stufen eines Kompetenzmodells auch als Schritte beim *Erwerb* von Kompetenzen interpretieren. In diesem Fall würde ein Kompetenzmodell Aussagen darüber machen, wie sich in der Lernbiographie von Kindern und Jugendlichen das Zusammenwirken der verschiedenen Komponenten von Kompetenz entwickelt, wie Kompetenzerwerb also verläuft. Diese Fragen sind jedoch wissenschaftlich wesentlich schwieriger zu beantworten als die Frage nach Niveaustufen innerhalb einer bestimmten Alters- oder Jahrgangskohorte von Schülerinnen und Schülern. Die Kompetenzstufen-Modelle von TIMSS und PISA sind beispielsweise dezidiert nicht als Entwicklungsmodelle gedacht, sondern als Beschreibung von Niveaustufen der mathematischen Kompetenz innerhalb der untersuchten Schülerpopulation.

Die Entfaltung eines wissenschaftlich abgestützten Kompetenzmodells im Sinne von Komponenten, Niveaustufen oder gar Entwicklungsverläufen wird nur selten vollständig möglich sein. Allerdings gehen auch die traditionelle Lehrplanarbeit und die Unterrichtsplanung von Lehrern *implizit* immer von Vorstellungen über Kompetenzstrukturen und deren Erwerb aus. Bildungsstandards sollten hier einen qualitativen Schritt nach vorn ermöglichen, indem sie die zugrundeliegenden Annahmen über den Aufbau von Kompetenzen explizit machen und – soweit möglich – den Anschluss an fachdidaktische und psychologische Forschung herstellen.

### **6.3 Zum Verhältnis von Wissen und Können**

Es wurde bereits aufgezeigt, dass Bildungsstandards Kompetenzmodelle nutzen, in denen der simultane Einsatz von Wissen und Können bei der Bewältigung von Anforderungssituationen eine wichtige Rolle spielt. Im Bildungsbereich herrscht jedoch nicht selten Unklarheit über das Verhältnis von Wissen und Können. Oft wird Wissen gleichgesetzt mit den Fakten, die in Lehrplänen niedergeschrieben sind. Selbst wenn es Schülerinnen und Schülern gelingt, dieses Wissen aufzunehmen, bleibt es doch oft „träge“. Das heißt, das Wissen kann außerhalb der Lernsituation nicht angewendet werden. Damit stellt es auch keine hinreichende Basis für kompetentes Handeln dar – der Auftrag der Förderung lebenslangen Lernens und zur verantwortlichen gesellschaftlichen Beteiligung wird damit untergraben. Soll die Idee von Bildungsstandards zum Tragen kommen, ist ein Bezug auf Theorien über den Zusammenhang von Wissen und Können hilfreich, wie sie in den letzten zehn Jahren in der empirischen Pädagogik und in der Kognitionswissenschaft ausgearbeitet und empirisch überprüft wurden. Diese Theorien gehen von einem allgemeineren Wissensbegriff aus, der Wissen nicht auf Kenntnisse reduziert.

Einige Theorien nehmen an, dass Wissen zunächst als deklaratives Wissen (explizites, verbalisierbares Wissen über Sachverhalte) erworben werden muss und dann zunehmend „prozeduralisiert“, d.h. in automatisch zugängliche Verknüpfungen und Abläufe überführt wird. Prozeduralisiertes Wissen muss nicht in jedem Anwendungsfall explizit und bewusst rekonstruiert werden, sondern ist wie eingeschliffene Bewegungsmuster und Gedankenfolgen automatisch verfügbar. Hiermit ist ein Prinzip genannt, das zur Abstufung von Kompetenzniveaus verwendet werden kann: Höhere Kompetenzniveaus in einer Domäne wären durch immer stärkere „Prozeduralisierung“ von Wissen charakterisierbar. In klassischen Begriffen bedeutet dies: Wissen geht auf höheren Niveaustufen in Können über.

Andere Prinzipien des Kompetenzaufbaus sind die zunehmende Vernetzung von Wissensselementen, die Bildung von Meta-Wissen (Wissen über das eigene Wissen, seinen Aufbau, seine Anwendung usw.) und abstrakterem Wissen. Solche Prinzipien lassen sich nutzen, um die Niveau- oder Entwicklungsstufen von Kompetenzen zu unterscheiden. Es wäre beispielsweise zu erwarten, dass auf höheren Kompetenzstufen in einer Domäne die Fähigkeit vorhanden ist, das eigene Vorgehen beim Bearbeiten von Problemen zu reflektieren und argumentativ darzustellen. In solchen Modellvorstellungen treffen sich, wie in Abschnitt 5.3 hervorgehoben wurde, bildungstheoretische und kompetenztheoretische (eher psychologisch fundierte) Theorien. Ob eine solche Verknüpfung von Wissen und Können, Wissen und Meta-Wissen usw. den Kompetenzaufbau in einer Domäne korrekt beschreibt, ist selbstverständlich eine Frage, die empirisch – oder zumindest anhand von gut begründeten Modellvorstellungen der Fachdidaktik – entschieden werden muss.

Das hier skizzierte Verständnis des Verhältnisses von Wissen und Können hat auch konkrete Konsequenzen für die Gestaltung von Lehrplänen und Unterricht. Der Aufbau von höheren Kompetenzstufen, die mit Handlungskompetenz und Können verbunden sind, gelingt nur, wenn Wissen stets der Bewährungsprobe erfolgreicher Leistung unterzogen ist. Die Verknüpfung von Wissen und Können darf also nicht auf Situationen „jenseits der Schule“ verschoben werden. Vielmehr ist bereits beim Wissenserwerb die Vielfalt möglicher Anwendungs-Situationen mit zu bedenken. Bedeutsam ist hier der Aufbau „domänenspezifischer Schemata“. Dies sind in Anwendungssituationen erworbene Wissensstrukturen, die von den Lernenden (nicht von den Lehrenden!) auf Grund ihrer Lern-Erfahrungen solcherart verallgemeinert und systematisiert werden, dass sie künftig auch auf andere Situationen anwendbar sind. Aus der empirischen pädagogischen Forschung ist bekannt, dass der Erwerb solcher Schemata beispielsweise durch fallbasierte Lehr-Lern-Umgebungen unterstützt werden kann. Diese erfüllen mehrere Kriterien erfolgreichen Kompetenzerwerbs:

- Die Lernenden machen innerhalb des Gegenstandsbereiches (Lern-)Erfahrungen, die für sie selbst relevant sind.

- Das Lehr-Lern-Material bildet die verschiedenen Komponenten von Kompetenz angemessen ab.
- „Überprüft“ wird die „Behandlung von Fällen“, nicht eine isolierte Komponente von Kompetenz wie z.B. Faktenwissen.
- Der kulturelle Kontext, innerhalb dessen Lernprozesse ablaufen, wird bedacht.
- Die Situation, in der der Lernprozess stattfindet, ist bedeutsam für dessen Ergebnis.

Es wird an dieser Stelle deutlich, dass Vorstellungen über Struktur, Stufung und Entwicklung von Kompetenzen, wenn sie in psychologischen und fachdidaktischen Theorien verankert sind, nicht nur die systematische Entwicklung von Aufgaben und Testverfahren stützen, sondern auch Hinweise für die Gestaltung der schulischen Lernumwelt liefern.

## 7. Testentwicklung: Methodische und diagnostische Grundlagen

Bildungsstandards werden zunächst verbal formuliert. Sie benennen die Kompetenzen, die Schülerinnen und Schüler im jeweiligen Lernbereich erwerben sollen, und stützen sich dabei auf Kompetenzmodelle, in denen Teilaspekte (Dimensionen) und Stufen dieser Kompetenzen spezifiziert werden. Bildungsstandards legen zudem normativ eine bestimmte Stufe als dasjenige Kompetenzniveau fest, das als Minimalanforderung von allen Schülerinnen und Schülern erreicht werden soll<sup>3</sup>.

Diese verbal, also qualitativ formulierten Bildungsstandards bedürfen, wie schon in Kapitel 2 ausgeführt, der Konkretisierung durch Aufgabenstellungen und Testverfahren. Die Entwicklung von Aufgaben, mit denen fachliche Bildungsziele konkretisiert werden, gehört auch zum Alltag der Lehrerinnen und Lehrer bei der Unterrichtsvorbereitung und der Leistungsbewertung. Standards können ihnen dabei Orientierung geben. Das Zusammenstellen und Erproben von Tests und schließlich der Einsatz von Tests im Rahmen der schulübergreifenden Qualitätssicherung und -entwicklung sind jedoch weitergehende, spezialisierte Tätigkeiten, die ein Zusammenspiel von Experten aus der Fachdidaktik, der empirischen Bildungsforschung und der pädagogisch-psychologischen Methodenlehre erfordern.

Zu begrüßen ist es, dass an Schulen im Rahmen der Selbstevaluation (siehe dazu Kapitel 9) zunehmend auf kollegialer Ebene innerhalb eines Fachteams Klassenarbeiten koordiniert und als Parallelarbeiten geschrieben werden, um die Ziele des Fachunterrichts im Kollegium abzustimmen und Information über die Leistungsstände paralleler Klassen eines Jahrgangs zu bekommen. Diese Parallelarbeiten können für die Diskussion vor Ort hilfreiche Information liefern, sollten jedoch nicht als Tests im hier verwendeten Sinne verstanden und schulübergreifend eingesetzt werden. (Zur Unterscheidung dieser Parallelarbeiten von schulübergreifenden, professionell entwickelten Testverfahren vgl. auch Helmke & Hosenfeld 2003).

Testentwicklung und -einsatz im Rahmen des Bildungsmonitoring oder einer externen Evaluation sollte professionellen Qualitätsmaßstäben genügen, weil sonst die Wahrscheinlichkeit von Fehl- und Überinterpretationen zu groß ist. Insofern beobachtet die Expertengruppe mit einer gewissen Sorge, dass in der Praxis mitunter schulübergreifende Vergleiche stattfinden, die diesen professionellen Kriterien nicht genügen und Gefahr laufen, mehr Fehlinformation und Schaden als Aufklärung und Orientierung zu erzeugen.

---

<sup>3</sup> Möglicherweise werden auch Aussagen über höhere Kompetenzniveaus gemacht, beispielsweise im Sinne von „Regelstandards“, die festlegen könnten, welches Niveau von mindestens der Hälfte der Lernenden erreicht werden soll. Dieser Aspekt wird jedoch im Folgenden nicht gesondert behandelt, weil die Testentwicklung selbst davon unberührt bleibt.

Auch innerhalb des Bereichs der professionellen, an Standards orientierten und mit den Mitteln der pädagogisch-psychologischen Methodenlehre (Rost 1996) arbeitenden Testentwicklung, mit der sich dieses Kapitel befasst, müssen mehrere Ziele des Umgangs mit den Testaufgaben unterschieden werden. Die jeweiligen Verwendungsziele von Testaufgaben können Implikationen auf die Aufgabenentwicklung, die Zusammenstellung von Testinstrumenten, die Testdurchführung und die Testauswertung haben und bestimmen letztlich die Prinzipien mit, nach denen die Testentwicklung organisiert werden sollte. Es gehört zu den Grundsätzen moderner Diagnostik, dass die Güte von Testverfahren (deren Validität, aber auch die Messgenauigkeit und andere Merkmale) nicht „an sich“ bewertet werden können, sondern immer nur im Hinblick auf die Art der Nutzung und die Art der Schlussfolgerungen, die man aus den Testergebnissen ziehen will.

Es werden im Folgenden vier solcher Einsatzbereiche standardbezogener Tests unterschieden:

1. *Überprüfung von Kompetenzmodellen*: Bildungsstandards sollten sich – wie in dieser Expertise begründet wird – auf differenzierte Kompetenzmodelle beziehen. Empirische Untersuchungen mithilfe von Tests sind erforderlich, um zu prüfen, ob diese Modelle tatsächlich die Aspekte der Kompetenzen von Lernenden, ihre Niveaustufung und ggf. ihre Entwicklung angemessen widerspiegeln. Derartige Untersuchungen stellen die Voraussetzung für einen Einsatz der Testinstrumente in einem der drei folgenden Bereiche dar und können daher auch der Phase der Testentwicklung zugerechnet werden.
2. *Systemmonitoring*: Wie es bei TIMSS und PISA der Fall ist, werden Tests verwendet, um Aussagen über das Kompetenzniveau von Schülerinnen und Schülern zu machen und Zusammenhänge mit schulischen wie außerschulischen Bedingungen aufzudecken. Ergänzende Befragungen, auch Videoerhebungen und andere Verfahren der empirischen Forschung sind denkbar, um Hintergrundfaktoren aufzudecken. Solche Studien liefern Informationen auf der Ebene des Bildungssystems (ggf. für ein Land) – also etwa Aussagen darüber, inwieweit Bildungsstandards tatsächlich eingelöst werden – aber in der Regel nicht auf der Ebene der Einzelschule. Aus den Daten von Systemmonitoring-Untersuchungen können auch Normen abgeleitet werden, die in späteren Untersuchungen eine normorientierte Interpretation der Ergebnisse, also einen Vergleich der Messwerte mit einer Referenzpopulation erlauben.
3. *Schulevaluation*: Hier werden Tests und andere Verfahren – möglicherweise auch interaktive Verfahren zur Reflexion der eigenen Praxis (Stichwort: Selbstevaluation) – eingesetzt, um zu prüfen, inwieweit eine Schule ihre pädagogischen Ziele erreicht. Im Idealfall geht eine solche Schulevaluation von einer genauen Analyse der jeweiligen pädagogischen Programmatik und Problemlage an der Einzelschule aus, um genau passend hierzu Information liefern zu können. Interne und externe Evaluation sollten ineinandergreifen.
4. *Individualdiagnostik und Förderung einzelner Schülerinnen und Schüler*: Mit dem Einsatz von Tests verbindet sich im pädagogischen Alltag die Erwartung, Aussagen über spezifische Stärken und Schwächen und damit den Förderbedarf einzelner Schülerinnen und Schüler zu

machen. Für diesen Zweck ist es in der Regel sinnvoller, einen kleineren Kompetenzbereich detaillierter zu erfassen, als das gesamte Spektrum eines Systemmonitorings mit relativ wenigen Aufgaben pro Schüler abzudecken.

Die dargestellten unterschiedlichen Anlässe der Verwendung von Aufgaben und Tests sollten in der öffentlichen Diskussion deutlich voneinander abgegrenzt werden. Dies gilt unbeschadet der Tatsache, dass es sehr wohl bestimmte Mischformen der o.g. Varianten geben kann. So kann man etwa eine Mischung aus Systemmonitoring und Schulevaluation vorsehen, z.B. wenn Schulen an einem landesweiten Evaluationsprogramm teilnehmen, das ihnen Informationen zu ausgewählten, zentral vorgegebenen Qualitätsaspekten gibt. Zu solchen Qualitätsaspekten wird zukünftig sicherlich die Einlösung der nationalen Bildungsstandards gehören. Häufig wird auch – wie etwa bei PISA – das Systemmonitoring mit der Überprüfung von Kompetenzmodellen verbunden.

Prinzipiell wäre noch eine weitere Art der Nutzung standardbezogener Tests denkbar: die Benotung und Zertifizierung von Schülerleistungen. Im weitesten Sinne können nämlich auch (zentrale) Abschlussprüfungen als Testverfahren angesehen werden, und es erscheint plausibel, Standards – insbesondere wenn sie sich auf Kompetenzanforderungen am Ende eines Bildungsgangs beziehen – als Grundlage solcher Prüfungsverfahren zu nutzen. Die Autoren dieser Expertise sprechen sich aber sehr deutlich gegen eine Verwendung von standardbezogenen Tests für Zwecke der Benotung und Zertifizierung aus. Bildungsstandards und deren Umsetzung in Tests dienen nach dem Verständnis dieser Expertise ausschließlich der Qualitätsentwicklung der Schulen und Schulsysteme, möglicherweise auch der Förderung einzelner Schülerinnen und Schüler, aber sie dienen NICHT der zentralisierten Examinierung (siehe dazu Kapitel 4).

Alle dargestellten Varianten machen sehr unterschiedlichen Gebrauch von den Standards, den zu ihrer Erfassung entwickelten Tests und den damit gewonnenen Ergebnissen. Die unterschiedlichen Zielsetzungen können letztlich auch Implikationen auf die Testentwicklung selbst haben.

Im vorliegenden Kapitel 7 gehen wir auf die Fragen ein, wie die Tests entwickelt werden (7.1) und welche Infrastruktur hierfür erforderlich ist (7.2). In beiden Punkten werden die unterschiedlichen Zielsetzungen von Testapplikationen zu berücksichtigen sein. Die Fragen, wie diese Tests möglicherweise eingesetzt werden und welche Infrastruktur dafür erforderlich ist, werden erst in Kapitel 9 aufgegriffen, da diese Fragen die Konsequenzen und Rahmenbedingungen der Einführung von Standards betreffen.

Die folgenden Überlegungen beziehen sich – dies sei nochmals ausdrücklich betont – nicht auf Leistungsmessungen im Arbeitsalltag der Schulen, sondern auf die Entwicklung zentraler Testinstrumente für einen schulübergreifenden Einsatz. Hier werden daher auch speziellere „technische“ Aspekte angesprochen.

## 7.1 Prinzipien der Testentwicklung

Im Folgenden werden vier grundlegende Prinzipien der Testentwicklung angesprochen, die jeweils einer Entscheidung bedürfen, wenn man ein konkretes Testinstrument für eine bestimmte Zielsetzung entwickelt:

- Sollen die Testergebnisse im Vergleich zu den Testergebnissen anderer Schüler bzw. Schulen interpretiert werden (man spricht dann von „verteilungsorientierter“ – oder „normorientierter“ Testinterpretation) oder in bezug auf ein inhaltliches Kriterium („kriteriumsorientiert“)?
- Soll die Kompetenz innerhalb eines Lernbereichs/Faches auf einer einzigen Gesamtskala („eindimensional“) erfasst werden, oder sollen mehrere Teilkompetenzen unterschieden und getrennt erfasst werden („mehrdimensionale Skalierung“)?
- Bearbeiten alle Schülerinnen und Schüler dieselben Testaufgaben, oder werden – durch sogenannte „Rotation“ von Testversionen innerhalb einer Klasse – den Schülerinnen und Schülern unterschiedliche Aufgaben vorgelegt, damit insgesamt mehr Aufgaben eingesetzt werden können?
- Soll ein und derselbe Test für alle Kompetenzniveaus und Schulformen gelten, oder verwendet man Aufgaben, die an das Leistungsvermögen der Gruppe bzw. des/der Einzelnen angepasst sind („verzweigtes“ bzw. „adaptives Testen“)?

In der ersten Frage impliziert schon die Konzeption der Bildungsstandards eine eindeutige Antwort: Bildungsstandards stellen Kriterien im Sinne des Konzepts der *kriteriumsorientierten* Leistungsmessung dar. In Abgrenzung zu einem reinen „Ranking“ geht es bei der Kontrolle der Einhaltung von Bildungsstandards nicht darum, wie der einzelne Schüler oder Schülergruppen relativ zu anderen Schülern oder anderen Vergleichspopulationen abschneiden. Den Referenzrahmen für die Interpretation der Ergebnisse bilden nicht die Verteilungen der Ergebnisse in der Gesamtpopulation oder in einzelnen Vergleichspopulationen, sondern allein die durch die Bildungsstandards gesetzten Kriterien. Das individuelle Erreichen oder Nichterreichen eines Standards darf nicht davon abhängen, wie viele Schüler diesen Standard erreicht haben.

Im Rahmen der Überprüfung von Bildungsstandards, die in Form von Kompetenzmodellen konkretisiert sind, stellen die als zentraler Bestandteil von Kompetenzmodellen beschriebenen *Kompetenzstufen* die Kriterien dar, an denen das Testergebnis gemessen wird. Kompetenzen sind abstrakte Konzepte einer wünschenswerten Bildung und nicht durch unmittelbar beobachtbares Verhalten definiert. *Kompetenzstufen* werden dann in Form von unterschiedlich anspruchsvollen kognitiven Prozessen und Wissensanforderungen definiert, deren Beherrschung mit bestimmten Niveaus einer Kompetenzdimension korrespondiert (siehe oben Kap. 6). Sie sind der Schlüssel zur Konstruktion und Auswertung kriteriumsorientierter Tests, indem sie eine Alternative zur willkürlichen Setzung von Leistungsmarken auf einem Kontinuum bieten.

Sind die Zielvorgaben, die es in verschiedenen Jahrgangsstufen zu erreichen gilt, in Form von Kompetenzstufen festgehalten, besteht das Ziel der Testentwicklung darin, valide Indikatoren für das Erreichen jeder Kompetenzstufe als Testaufgaben zu konstruieren. Dazu ist es erforderlich, Testaufgaben zu entwickeln, die jeweils genau das erfassen, was für eine bestimmte Kompetenzstufe charakteristisch ist. Eine solche systematische, am Modell der Kompetenzstufen orientierte Testentwicklung ist ohne didaktische und psychologische Expertise nicht möglich. (Viele Beispiele hierzu gibt der Bericht des National Research Council 2001). Die Aufgabenentwickler müssen vorab festlegen, welcher Kompetenzstufe eine Aufgabe entspricht, welche Prozesse und Wissensinhalte sie erfordert. Diese Anforderungsprofile der Aufgaben müssen empirisch geprüft werden, um sicherzustellen, dass tatsächlich das gemessen wird, was im Kompetenzmodell angesprochen ist. Erst nach einer solchen Erprobungsphase kann der Test zur Kompetenzmessung eingesetzt werden. Das individuelle Testergebnis besteht dann in der Feststellung, welche Kompetenzstufe die Schülerin und der Schüler im Rahmen des jeweiligen Kompetenzmodells erreicht.

Empirische Erhebungen im Sinne eines *Systemmonitoring* (s.o. Punkt 2) sollten primär auf die kriteriumsorientierte Interpretation der Testergebnisse ausgerichtet sein, auch wenn eine *zusätzliche* Entwicklung von Verteilungsnormen im Sinne einer Testnormierung nicht ausgeschlossen ist. TIMSS und vor allem PISA haben gezeigt, wie eine solche Verbindung von anspruchsvoller kriteriumsorientierter Testentwicklung und (auch) normorientierter Interpretation aussehen kann.

Eine *mehrdimensionale* Messung in Form von Profilen wünschenswerter Kompetenzen wird der Beschreibung von Bildungsstandards eher gerecht als eine *eindimensionale* Abbildung von Leistungsunterschieden. Die Erfassung möglichst vieler Kompetenzdimensionen ermöglicht Rückschlüsse auf die Angemessenheit von Zieldefinitionen und deckt zugleich Defizitbereiche auf, die für eine Nicht-Einhaltung der Standards verantwortlich sind. Insofern verhindern differenzierte Bildungsstandards und mehrdimensionale Tests ein einfaches Ranking, das für eine Qualitätsentwicklung der Schulen wenig hilfreich wäre.

Werden Tests zum Zweck der Individualdiagnostik der Leistungen einzelner Schüler verwendet etwa mit dem Ziel geeignete Fördermaßnahmen einzusetzen (s.o. Punkt 4), so ist es selbstverständlich, dass jede bzw. jeder einzelne der untersuchten Schüler die zur Beurteilung ihrer/seiner Stärken und Schwächen relevanten Testaufgaben vollständig bearbeitet. Dies ist bei Erhebungen zum Systemmonitoring im allgemeinen nicht der Fall. Um die Aussagekraft der Ergebnisse auf der aggregierten Ebene, d.h. für eine Klasse, eine Schule oder ein Land, zu erhöhen, kann das Spektrum der mit den Tests erfassten Leistungen dadurch drastisch verbreitert werden, dass jeder Schüler und jede Schülerin nur noch einen Teil der Testaufgaben bearbeitet. Die Gesamtheit der erfassten Leistungen sichert eine hohe Validität der Ergebnisse auf dem Aggregationsniveau von Klassen, Schulen oder Ländern. Die Vergleichbarkeit und Zuverlässigkeit der Ergebnisse auf Individualniveau ist aber eingeschränkt.

Im Rahmen der Verwendung moderner, sogenannter probabilistischer Testmodelle (Rost 1996) bei der Testanalyse ist es nicht nur möglich, dass verschiedene Schülergruppen unterschiedliche Testteile bearbeiten, sondern diese Testteile können in ihrem Schwierigkeitsgrad auch dem erwarteten Leistungsniveau von Teilpopulationen angepasst sein. Bei einer gegebenen Vernetzung der Testteile durch Brückentests lassen sich trotzdem die Leistungsparameter aller Schüler auf derselben Skala abbilden. Die Vorteile solcher niveau-angepasster Tests liegen zum einen in einer höheren Messgenauigkeit der Testergebnisse, zum anderen aber auch in einer Aufrechterhaltung der Testmotivation, die bei einer Unter- oder Überforderung der Schüler durch zu leichte oder zu schwere Tests in der Regel nachlässt.

Die eingangs genannten vier unterschiedlichen Zwecke oder Ziele von Testanwendungen und die hier dargestellten Implikationen machen deutlich, dass es nicht um die Entwicklung des einen Testinstrumentes geht, das universell einsetzbar wäre. Der jeweilige Zweck der Testung bestimmt die Optimalitätskriterien mit. Es könnte also sinnvoll sein, für das Systemmonitoring breitere Aufgabensammlungen einzusetzen, während bei der Schulevaluation und insbesondere für die Individualdiagnostik Kompetenzbereiche ausgewählt werden, die dann mit zusätzlichen Aufgaben genauer und differenzierter erfasst werden. Trotzdem liegt den möglichen Testvarianten das gemeinsame Konzept zugrunde, dass die Leistungserhebung inhaltlich nicht an einer willkürlichen Auswahl von Aufgabenstellungen orientiert ist, sondern an den Kompetenzen und Kompetenzmodellen, die die Bildungsstandards definieren.

## **7.2 Zuständigkeit und Institutionalisierung der Testentwicklung**

Die Testentwicklung für schulübergreifende Zwecke sollte in der Verantwortung einer wissenschaftlich qualifizierten Agentur stehen, die die methodische Qualität der Instrumente sicherstellt. Diese Agentur kann und sollte die Entwicklung der Instrumente aber nicht allein durchführen, sondern eng mit den Landesinstituten, fachdidaktischen Vereinigungen, Lehrerverbänden und universitären Zentren zusammenarbeiten. Insbesondere sollte sich eine solche Agentur nicht zu einer übermächtigen und zur Verselbständigung neigenden Institution entwickeln, sondern in relativ schlanker Organisationsstruktur die anstehenden Koordinierungsaufgaben erledigen.

Einer solchen Agentur obliegt die Koordination der Testentwicklung, die Abstimmung zwischen den Ländern, die Qualitätssicherung der Instrumente und die Rückkopplung an die anderen Gremien der Entwicklung von Bildungsstandards (s.o.). Hier muss auch an die Einbindung kommerzieller Anbieter von Testentwicklungen gedacht werden.

Verschiedene Varianten sind denkbar, von denen Variante C nach dem zuvor gesagten als die schwierigste angesehen wird.

- Variante A – *Auftragsentwicklung*. Die Testagentur gibt vor, welche Instrumente zu welchem Zeitpunkt zu entwickeln sind und macht eine Projektausschreibung, die zu einer Beauftragung geeigneter Antragsteller führt.
- Variante B – *Angebotsorientierung/Zertifizierung*. Die Testagentur recherchiert eigenständig, greift vorhandene Möglichkeiten auf und ist offen für Angebote aus Institutionen und Ländern. Ihre Arbeit könnte dann eher darin bestehen, vorhandene – ggf. auch kommerziell entwickelte – Testinstrumente zu zertifizieren, d.h. danach zu bewerten, ob sie den Bildungsstandards entsprechen und den Qualitätsmaßstäben der wissenschaftlich fundierten Testentwicklung genügen.
- Variante C – *Testentwicklungsinstitut*. Die Testagentur hält sich einen Stab professioneller Testentwickler, die in ein Kooperationssystem mit Lehrern und Wissenschaftlern eingebunden sind.

Die Komplexität der anstehenden Entwicklungsaufgaben und die weitgehend fehlende Tradition in der Entwicklung von Testinstrumenten dieser Art lässt eine einfache Lösung unrealistisch erscheinen. Es sollten die genannten gesellschaftlichen und beruflichen Gruppen einbezogen werden, nicht nur um eine breite Kompetenzbasis zu sichern, sondern auch um zu unterstreichen, dass Bildungsmonitoring eine gesamtgesellschaftliche Aufgabe ist.

Unstrittig ist, dass eine solche Agentur erheblichen Finanzierungsbedarf mit sich bringt. Keines der derzeit in diesem Sektor aktiven Institute kann eine solche Aufgabe kostenneutral mit übernehmen. Soll eine solche Agentur eine zentrale Aufgabe in einer output-orientierten Steuerung unseres Bildungssystems übernehmen, so stellt dies eine neuartige Organisationseinheit dar, die weder zum Nulltarif zu haben ist, noch von heute auf morgen erfolgreich funktionieren kann. Es ist mit einer mehrjährigen Aufbau- und Übergangszeit zu rechnen (vgl. dazu auch Kapitel 11, insbesondere S. 107).

## 8. Folgen für Lehrplanentwicklung und schulische Lehrpläne: Standards und Kerncurricula

Bildungsstandards, wie wir sie in den vorstehenden Kapiteln konzipiert haben, können den Lehrerinnen und Lehrern als professioneller Referenzrahmen und den Schulen als Orientierungshorizont ihrer pädagogischen Schulentwicklung dienen. Dies allein kann bereits Folgen für die Planung und Gestaltung des Unterrichts und die Qualitätsentwicklung an Schulen haben. Allerdings werden die Schulen zusätzliche Anleitung, Unterstützung und Beratung benötigen, um produktiv mit den Bildungsstandards umgehen zu können. Daraus folgt, dass die Lehreraus- und -fortbildung, die Lehrplanarbeit, die Schulaufsicht und andere Instanzen der Bildungsadministration die Impulse der Bildungsstandards aufgreifen und neue Funktionen übernehmen müssen.

Zusätzlich zu dieser Orientierungsfunktion können Bildungsstandards der Qualitätssicherung im Bildungssystem dienen: Lernergebnisse werden auf der Ebene des Systems („Bildungsmonitoring“) und auf der Ebene der Einzelschule („Schulevaluation“) empirisch erfasst und den Betroffenen zurückgemeldet, um – im Sinne einer „Output-orientierten Steuerung“ – aus der Erfahrung lernen zu können. Diese Monitoring- und Evaluationsfunktion ist ebenfalls und vermutlich in noch stärkerem Masse mit Änderungen im Bildungssystem verbunden, weil bislang die Qualität der schulischen Bildung eher über Input und Prozess (Stundentafeln und Unterrichtsversorgung, Curricula, Benotungs- und Versetzungsregelungen, Anerkennung von Schulbüchern u.a.m.) sichergestellt wurde.

In den folgenden Kapiteln 8 bis 10 diskutieren wir diese Konsequenzen der Bildungsstandards im einzelnen. Zunächst geht es um die Lehrplanarbeit (mit wachsender Bedeutung von Kerncurricula und schulischem Curriculum anstelle detaillierter Zentralcurricula), sodann um Monitoring und Evaluation und schließlich in Kapitel 10 um die notwendigen Änderungen in der Arbeit der Unterstützungssysteme.

### 8.1 Lehrplanpraxis und Steuerung von Schularbeit

Die Einführung nationaler Bildungsstandards ist selbstverständlich nicht folgenlos für die aktuell dominierende Form der Steuerung von Schule und Lernen mit Hilfe von Lehrplänen und ihnen zugeordneter Materialien wie Lehrbüchern oder Unterrichtshilfen für den Lehrer. Ohne Zweifel wird sich, wie es auch die internationalen Erfahrungen mit Bildungsstandards belegen, die Schularbeit mit der neuen Form der output-orientierten Steuerung verändern, aber diese Erfahrungen mit der Einführung von Bildungsstandards sind hinsichtlich der Konsequenzen für die Lehrpläne

nicht eindeutig, die Optionen entsprechend vielgestaltig und offen. Es ist deshalb notwendig, die bestehenden Möglichkeiten der Lehrplansteuerung von Unterricht zu prüfen, um begründet über die wünschbaren Konsequenzen entscheiden zu können.

Bislang wurde in Lehrplänen für die einzelnen Länder mehr oder minder weitreichend festgeschrieben, *was* (Stoff und Inhalte), *wann* (Klasse), *wie* (Methode) und *wo* (Schulart) zu lehren ist. Die Entwicklung und Vermittlung von Lehrplänen repräsentiert insofern nicht nur einen Prozess der Selektion und Transformation von Kultur und der gesellschaftlichen Bildungsziele in staatliche Vorgaben für den Unterricht (wie in Kapitel 5 ausgeführt), sie stellen zugleich den Versuch dar, die konkrete Gestaltung von Unterricht und das erwünschte Ergebnis pädagogischer Arbeit zu normieren. Die historisch-gesellschaftlich zu beobachtende Definition von Lehrplänen war deshalb auch Gegenstand öffentlicher, bildungspolitischer, rechtlicher, administrativer und schulpraktischer Auseinandersetzungen. Die dominierende Rolle des Staates in diesem Prozess blieb dabei im Grunde unbestritten (Biehl u.a. 1998).

Im internationalen Vergleich moderner Gesellschaften lassen sich idealtypisch vier Grundvarianten der staatlichen Regulierung des Unterrichts bestimmen, die bis heute gelten und unabhängig von der Form und dem Medium der Steuerung – über Lehrpläne, Curricula oder Standards – wirksam sind (Biehl u.a. 1996):

(a) Assessment-Modell

Durch unterschiedliche Abgangskontrollen, wie standardisierte Schulleistungstests, wird den Schulen relativ präzise vorgegeben und zugleich gespiegelt, was sie zu bestimmten Messzeitpunkten im Unterricht erreicht haben sollen und wirklich erreicht haben. Die Lehrkräfte der Schulen tun hier gut daran, ihren Unterricht auf diese Abgangsprüfungen hin auszurichten. Der konkrete Unterricht selbst wird also gleichsam indirekt reguliert.

(b) Examen-artium Modell

Statt Abgangskontrollen haben hier Zugangskontrollen zu den einzelnen Bildungsinstitutionen die Funktion, den Unterricht der jeweils vorausgegangenen Bildungsinstitution zu regulieren. Die Aufnahmeprüfungen an Colleges haben dann z.B. eine regulierende (und wettbewerbssteigernde) Funktion für den Unterricht an den High-Schools. Wie im Assessment-Modell liegt eine indirekte Regulierung des Unterrichts vor.

(c) Philanthropisches Modell

Einerseits schreibt der Staat durch Lehrpläne stoffliche und methodische Vorstellungen von Unterricht verbindlich fest, auch wenn diese Vorgaben meist nur auf allgemeiner Ebene formuliert sind. Andererseits entwickelt der Staat potentielle schulische Innovationen in der Eröffnung von Modellversuchen. Hier liegt also eine direkte Regulierung des Unterrichts in seinen Stoffen und den praktizierten Methoden durch Lehrplanvorgaben vor.

#### (d) Klassisch-bürokratisches Modell

Der Staat konzentriert sich hier auf die Vorgabe eines Lehrplans, der Stoffe und Inhalte schulischen Unterrichts – z.T. äußerst genau – enthält, sowie auf schulartspezifische Normierung zur Steuerung der Selektion. Dem Lehrer bleibt im Unterschied zum philanthropischen Modell nach einer Art Lizenzprinzip die Wahl der Methode und damit die Verantwortung für die Durchführung des Unterrichts überlassen. Die staatliche Kontrolle des Schulwesens geht davon aus, dass Lehrplanvorgaben und die Lizenzierung und Professionalisierung des Personals die Umsetzung garantieren, und er verzichtet daher weitgehend auf eine Output-Kontrolle.

In Deutschland regiert ein Mischsystem, in dem mit der Vorgabe von Lehrplänen, mit Kompetenzprüfungen für das Personal und mit teilstandardisierte Abgangskontrollen, z.B. im Abitur oder in der Regulierung der Anforderungen in den mittleren Abschlüssen, die Qualität der Schularbeit gesichert werden soll. Leistungsuntersuchungen wie TIMSS und PISA haben dann gezeigt, dass dieses Steuerungssystem seine Schwächen hat, Bildungsstandards sollen hier Abhilfe schaffen. Aber was sind die Konsequenzen für die traditionellen Steuerungspraktiken, vor allem: welche Konsequenzen hat das für Lehrpläne und für die Lehrplanarbeit?

In radikalen Modellen der output-orientierten Steuerung verlieren zentrale Lehrpläne auf nationaler oder – wie in Deutschland – Länderebene nahezu vollständig ihre Bedeutung als strukturierendes Element von Unterricht. Hier geht man von der These aus, dass langfristig keine Doppelregelung Bestand haben könne: Lehrpläne der herkömmlichen Art müssen und werden, so die These, bei einer Output-Steuerung des Bildungssystems ihre Funktion an eine zielgerichtete, der Autonomie der Einzelschule verpflichtete Standardorientierung abgeben, damit die angestrebten Kompetenzziele auch wirklich ermöglicht werden.

Schulisches Lernen wird damit durch nationale, einheitliche, verbindliche Bildungsstandards und deren Überprüfung einerseits, durch lokal und innerschulisch bzw. einzelschulisch entwickelte Curricula andererseits gesteuert. Als Begründung für diese Position kann man die gut bestätigte Erfahrung ins Feld führen, dass Lehrpläne nicht sicher und eindeutig jene Qualifizierung und Qualitätssicherung erbringen, die man von ihnen erwartet. Die Schwierigkeiten dieser neuen Steuerungsform sind aber ebenso offenkundig: Die Einzelschule wird, zumal bei einem raschen Übergang zum neuen System, schnell überfordert, und Staat und Gesellschaft verzichten darauf, den Bestand allein an Können und Wissen, aber auch an Werten und Orientierung sichtbar und verbindlich zu machen, der zur Teilhabe an unserer Kultur notwendig und unverzichtbar ist.

Auf der Gegenseite wird deshalb in Modellen nationaler Curricula ein Steuerungssystem entwickelt, in dem sowohl verbindliche Bildungsstandards und Kompetenzerwartungen als auch Fächer, Themen und Inhalte, gelegentlich sogar Lernformen zentral vorgegeben werden. Solche nationalen Curricula sind immerhin geeignet, das Qualitätsbewusstsein für die Schularbeit zu stärken, die Wahl von Schulen nach Leistungskriterien rational zu gestalten und einen nationalen Diskurs über die Erwartungen an Schule zu organisieren. Als Folgeproblem ist andererseits un-

übersehbar, dass die Einzelschulen dabei in ihren Anstrengungen eingeschränkt, wenn nicht sogar gegängelt werden, lokale Problemlagen und spezifische Entwicklungsaufgaben nicht berücksichtigen können und damit eine Normierung eintritt, die die eigene Motivation und Anstrengung vor Ort eher schwächt als stärkt. Die Autonomie der Einzelschule ist aber eine wesentliche Voraussetzung für Leistungssteigerung und Wettbewerb.

Eine fast zweihundert Jahre alte, etablierte Form der Steuerung eines Bildungssystems zu verändern und auf ein anderes Prinzip umzustellen, ist deshalb keine leichte Aufgabe. Zwar stellt sie nicht so sehr ein bildungstheoretisches Problem der Legitimation von Steuerung dar, denn gesteuert wird in jedem Fall, sondern ein Problem der praktischen Umgestaltung, das die Akteure im Bildungsprozess vor ungewohnte Herausforderungen stellt. Ein behutsames Herangehen an die Implementation von Innovationen ist deshalb notwendig, um nicht in bester Absicht kontraintentionale Wirkungen zu erzeugen. Durch klare und überzeugende Perspektiven Akzeptanz für die intendierten Veränderungen zu schaffen, das ist die erste Aufgabe, und diese Akzeptanz muss in der Einzelschule, bei Schülern, Eltern und in den Kollegien, ebenso erreicht werden wie in der Gesellschaft, die mit guten Gründen erwartet, dass ihr Bild der Ziele und Werte, der Gestaltung von Gegenwart und Zukunft von Staat, Gesellschaft und Kultur auch Eingang in die Schule findet und sich im Wissen und in den Kompetenzen der Lernenden manifestiert.

## 8.2 Kerncurricula

Nach unserer Meinung ist deshalb eine Strategie der Steuerung notwendig und angemessen, in der die Leitfunktion nationaler Bildungsstandards und die Orientierungsfunktion von Lehrplänen systematisch gekoppelt werden, und zwar so, dass die Autonomie der Einzelschule gefördert wird. Im Ansatz der fachbezogenen *Kerncurricula* lässt sich ein solches integriertes Steuerungssystem entwickeln, das einerseits anschlussfähig ist an Kompetenzvorgaben, wie sie in Bildungsstandards entwickelt werden, und andererseits offen ist für zeitliche Sequenzierungen und konkrete Unterrichtsempfehlungen, wie sie zur Orientierung der alltäglichen Schularbeit unentbehrlich sind.

Auch Kerncurricula werden, obwohl funktional vergleichbar, international in unterschiedlicher Gestalt entwickelt und staatlich vorgegeben: z.B. kompakt, konzentriert und stufenbezogen, in knappem Umfang, abschluss- oder kompetenzorientiert, an zentralen Themen und großen Ideen ausgerichtet und durch die Arbeit in konkreten Fächern bestimmt, wie z.B. in Norwegen oder in dem in Kapitel 3 vorgestellten Beispiel des NCTM; aber es gibt auch die Versuchung, Kerncurricula (die dann keine „Kerne“ mehr sind) in expansiver Tendenz zu entwickeln, den Fachgeheimen nachzugeben, die Verantwortung der Einzelschule zu dispensieren und auf bekannte Modelle bürokratischer Steuerung und curricularer Allmachtsphantasien zurückzugreifen. Unverkennbar

ist, dass der Begriff des „Kerncurriculums“ noch nicht eindeutig normiert ist: Aber deshalb kann man auf diese curriculare Arbeit nicht verzichten, sondern muss präzise angeben, in welchem Sinne man von Kerncurricula spricht und wie sie sich zu Bildungsstandards verhalten (vgl. Tenorth 2001).

Bildungsstandards und Kerncurricula schließen sich in dem hier vertretenen Konzept nicht gegenseitig aus, sondern überlappen und ergänzen sich. Kerncurricula wie Bildungsstandards sind Elemente innerhalb eines Systems der Steigerung und Steuerung der Qualität des Bildungswesens; ihrer Funktion nach setzen Bildungsstandards am Output an, für den sie Vorgaben spezifizieren, Kerncurricula hingegen am Input, d.h. an der Auswahl der Inhalte und Themen und der Gestaltung von Lehr-Lernprozessen. Den Schnittbereich von Kerncurricula und Bildungsstandards bilden die bildungstheoretischen Leitideen und Kompetenzmodelle. Standards ergänzen dies durch die Stufung und Ausdifferenzierung von Kompetenzanforderungen und letztlich durch konkrete Testverfahren, während Kerncurricula zusätzlich exemplarisch Themen für die inhaltliche Gestaltung schulischer Lehr-/Lernprozesse benennen und modellhaft Anregungen für die Praxis pädagogischer Arbeit geben. Kerncurricula umfassen deshalb mehr als nur Bildungsstandards im Sinne von individuumbezogenen Kompetenzmodellen. Sie eröffnen eine Welt des Lernens.

Das Fazit dieser Überlegungen ist, dass Bildungsstandards die Lehrpläne keineswegs überflüssig machen, weil sie deren inhaltliche und prozessuale Orientierungs- und zeitliche Steuerungsfunktion nicht übernehmen können. Es wird aber sinnvoll sein, die Lehrpläne schrittweise in Richtung auf Kerncurricula umzugestalten – eine Tendenz, der die Curriculumarbeit in Deutschland bereits seit Jahren immer mehr folgt. Die Entwicklung von Bildungsstandards und Kerncurricula sollte in Zukunft Hand in Hand gehen. Detailbestimmungen, z.B. über die möglichen Unterrichtsthemen, die zu behandelnden Texte oder die wichtigsten Arbeitstechniken, können dann zunehmend, aber behutsam von der Landes- auf die Schulebene verlagert werden.

### **8.3 Bildungsstandards, Kerncurricula und der Kanon der Allgemeinbildung**

Die besondere Leistung einer solchen neuen Art der Steuerung kann man im Vergleich mit dem traditionellen Bildungskanon deutlich erkennen. Einerseits, Bildungsstandards und Kerncurricula stellen nicht den Versuch dar, die historisch gewordene Tradition eines klassischen Bildungskanons einfach neu zu beleben. In Deutschland (aber nicht nur hier) repräsentiert der Begriff des „Kanon“ in der öffentlichen Diskussion ja meist eine spezifische Tradition der Konstruktion von Bildung und des Bildungswissens: Programmatisch oder polemisch auf den Kanon der sog. „höheren“ Schulen konzentriert, ist die Kanondiskussion deshalb hierzulande aber nicht nur institutionell eingengt, sondern zugleich auch in ihren Referenzen und Dimensionen spezifisch sozial

geprägt. Der „Bildungskanon“ wird vor dem Hintergrund der europäischen, bürgerlichen, gelehrten Kultur gelesen und überliefert, als Ausdruck eines spezifischen Lebensstils, nicht allein in seiner schulischen Realität, sondern in den vielfachen Ausprägungen kultureller Praxis und einer spezifischen Geselligkeit, ausgestattet mit eigenen Gütekriterien und mit kodifizierten Mustern seiner Aneignung und der Teilhabe an der damit strukturierten Weltsicht – und dann wird der „europäische Bildungskanon“ eher als vergangene Welterfahrung betrauert, denn als eine heute noch gelebte Wirklichkeit wahrgenommen.

Kerncurricula repräsentieren zwar auch das Selbstverständnis einer Kultur, aber nicht schichtspezifisch separiert, sondern als Ausdruck eines Kanons der Allgemeinbildung, der universell, für alle Heranwachsenden einer Gesellschaft gültig sein soll. Ihre Vorgaben sind auf die kulturellen Basisfähigkeiten bezogen, eindeutig schulisch definiert, konkretisiert in obligatorischen Lernprozessen und dort als Instanzen der Normierung und Steuerung verstanden. Wenn man den Kanon der Allgemeinbildung in dieser Weise versteht, also als schulischen Kanon bzw. als Form der Kanonisierung schulischer Lernprozesse, entdeckt man die funktionalen Äquivalenzen zur Tradition und man sieht auch, dass Schulen auf Kanonisierung gar nicht verzichten können. Dann kann die Konstruktion von Kerncurricula aber auch von den Erfahrungen lernen, die mit einem, nämlich dem überlieferten klassischen „Bildungskanon“ gemacht worden sind.

Die Geschichte dieses Kanons und die aktuelle Debatte über die Bestimmung derjenigen Elemente unserer Kultur, die notwendig zu tradieren sind, damit die Heranwachsenden an unserer Kultur selbstbestimmt teilhaben, machen bewusst, dass ein Kerncurriculum zur Normierung und Standardisierung der schulischen Arbeit nicht hinreichend ist, wenn man es allein als Inhaltskomplex definiert. Inhalte, auch konkrete Wissensbestände, die Orientierung in der Welt ermöglichen, sind unentbehrlich, aber die immanente Präsenz von Gütekriterien und der notwendige Prozess der Aneignung eines Kanons lässt sich nicht ignorieren, wenn man verstehen will, wie „Kanonisierung“ funktioniert. Kerncurricula müssen deshalb, um die Kanonisierung von Wissen und Einstellungen, Motiven und Orientierungen zu regulieren, als Einheit von Zielen und Themen, Kompetenzvorgaben und Gütekriterien, Modellen guten Lebens und angemessenen Lernens verstanden werden. In diesem Sinne sind sie als Komplementierung und Konkretisierung eines Systems von Bildungsstandards dann auch unentbehrlich.

Im schulischen Kontext repräsentiert das Kerncurriculum deshalb die Struktur allgemeiner Bildung und die Initiation in die für das Leben notwendigen Modi der Welterschließung: Sprachlich-literarische, mathematisch-naturwissenschaftliche, historisch-sozialwissenschaftliche sowie ästhetisch-expressive Dimensionen grundlegender Allgemeinbildung. Schule als Institution wird damit zu der gesellschaftlichen Form, in der Kulturen, weltweit einander zunehmend gleich, Inhalte und Normen des Lebens definieren und sie wird damit zugleich zur pädagogischen Form, in der unsere Kultur ihre eigene Lehrbarkeit umsetzt und dabei ihren Kern an Wissen und Orientierungen verbindlich macht.

Kerncurricula können, zusammenfassend, diese Leistung ermöglichen und dokumentieren, denn:

- sie bestimmen ein obligatorisches Fächergefüge,
- sie nennen zentrale Themen und Inhalte,
- sie bezeichnen erwartete Kompetenzen der Adressaten schulischer Arbeit und
- das alles „klar, eindeutig und verbindlich“<sup>4</sup>.
- angesichts dieser Art von Vorgaben vertragen sich Kerncurricula sowohl mit einer dezentralisierten, an „Schulautonomie“ ansetzenden Form der Steuerung, als auch mit der zentralen Normierung eines Minimalstandards.
- Kerncurricula bezeichnen nämlich nicht das Totum, sondern allein das unentbehrliche Minimum der Themen, Inhalte und Lehrformen der Schule.
- sie sind damit, lehrplantechnisch und -theoretisch gedacht, sowohl offen für die fachinterne Vertiefung als auch für die Erweiterung durch andere Fächer und für die thematische Kopplung von Lehrgegenständen.
- Kerncurricula nehmen also den produktiven Gehalt bisheriger Lehrpläne und Richtlinien auf, die mit ihren elaborierten Anregungen für die Gestaltung bestimmter Unterrichtsinhalte eine wesentliche Basis eines zu schaffenden Supportsystems für die unterrichtliche Praxis und für schulinterne Curriculuarbeit darstellen.
- sie erlauben Profilbildung auf der Ebene der Einzelschule, verbunden mit einem einheitlichen Niveau der Arbeit im System und an den Alterskohorten der jeweiligen Schulgenerationen.
- die Verknüpfung der Inhalts- mit der Kompetenzdimension der Vorgaben ermöglicht schließlich den unmittelbaren Anschluss an die Entwicklung von Bildungsstandards und die präzise Messung des jeweils erreichten Leistungsstandes in Schulen und von Lernenden.

In einem System deregulierter, offener und dezentraler Steuerung bilden Kerncurricula die Instanz, um lokale Entwürfe und partikulare Ambitionen an einem Modell zu prüfen, das den Anspruch des Allgemeinen mit sich führt, aber der Konkretisierung bedarf, um wirksam zu sein. Kerncurricula und Bildungsstandards gemeinsam stellen insofern den Referenzrahmen dar, der innerschulische Arbeit anregen, unterstützen, orientieren und normieren kann; aber sie sind nicht selbst schon der einzelschulische Lehrplan, sondern sie ermöglichen die von der Schule, dem einzelnen Lehrer und den Kollegien ausgehende, mithin professionsbasierte und mit Hilfe von empirischen Verfahren kontrollierbare Konstruktion von Unterricht.

---

<sup>4</sup> Für den Kontext: Wolfgang Böttcher/Peter E. Kalb (Hrsg.): Kerncurriculum. Was Kinder in der Grundschule lernen sollen. Weinheim/Basel 2002; Heinz-Elmar Tenorth (Hrsg.): Kerncurriculum Oberstufe. Mathematik – Deutsch – Englisch. Expertisen, im Auftrag der KMK. Weinheim/Basel 2001.

## 9. Bildungsmonitoring, Evaluation und Individualdiagnostik auf der Basis von Standards

Bildungsstandards unterscheiden sich, wie im vorigen Kapitel gesagt, von Kerncurricula vor allem dadurch, dass die Standards konkrete Kompetenzanforderungen bestimmen und somit einen Maßstab zur Bewertung von Lernergebnissen beinhalten. Damit spielen sie – über die reine Orientierungsfunktion hinaus, die auch Curricula erfüllen können – eine zentrale Rolle in output-bezogenen Steuerungsmodellen. Bildungsstandards definieren Maßstäbe, nach denen der „Output“ erfasst und bewertet werden kann. Wie Standards und standard-basierte Testverfahren verwendet werden können, um Feedback über Lernergebnisse zu geben, soll im Folgenden dargestellt werden.

In Kapitel 7 wurden vier verschiedene Zielsetzungen von empirischen Untersuchungen des Leistungsstands von Schülern unterschieden und es wurde auf einige Implikationen dieser Zielsetzungen für die Testentwicklung hingewiesen. Rechnet man die empirische Untersuchung von Kompetenzmodellen der Phase der Testentwicklung zu (siehe Kap. 7), so sind im Wesentlichen drei Ziele einer Nutzung von standard-bezogenen Testverfahren in der Bildungspraxis zu unterscheiden. Die Untersuchung von Schülerleistungen kann dazu dienen,

1. das Erreichen der Bildungsstandards zu kontrollieren – zum Beispiel: Werden die Mindeststandards von allen Schülern erreicht? Entspricht die Verteilung der Schüler über die Kompetenzstufen den Erwartungen? etc. (*Bildungsmonitoring*),
2. Rückschlüsse auf den Erfolg schulischer Programme oder unterrichtlicher Maßnahmen zu ziehen und damit eine Grundlage für die Verbesserung von Schule zu haben (*Schulevaluation*),
3. etwas über den einzelnen Schüler zu erfahren, über seine Stärken und Defizite, damit gezielte Fördermaßnahmen getroffen werden können (*Individualdiagnostik*).

Im Folgenden werden diese Varianten der Nutzung von Tests näher beschrieben. Dabei wird auch darauf eingegangen, welche Anforderungen eine regelmäßige Durchführung solcher Untersuchungen an eine entsprechende Infrastruktur in unserem Bildungssystem stellt.

## 9.1 Zur Unterscheidung zwischen Bildungsmonitoring und Evaluation

Die Festlegung von Bildungsstandards macht es möglich, in gewissen Zeitabständen zu prüfen, ob die gesetzten Standards bei den Schülerinnen und Schülern erreicht werden. Diese Aufgabe eines kontinuierlichen Bildungsmonitoring unterscheidet sich von den Anliegen einer Evaluation von Schule oder Unterrichtsprozessen (zu dieser Unterscheidung vgl. Klieme, Baumert & Schwippert 2001).

Bei Studien zum *Bildungsmonitoring* wird der Stand der Bildung im Schulsystem möglichst differenziert erfasst, um ein Bild der aktuellen Situation zu erhalten. Die Testergebnisse werden im Referenzrahmen der vorgegebenen Zielkriterien interpretiert. Dies setzt die Existenz derartiger Zielkriterien voraus. Die Daten können in einer Zeitreihe (z.B. wenn Erhebungen regelmäßig alle drei bis fünf Jahre stattfinden) verknüpft werden, um Veränderungen in der Zeit zu studieren, und sie können nach Teilpopulationen (z.B. Länder oder Schularten) aufgeschlüsselt werden, sofern die Stichprobengröße und die Messgenauigkeit dies zulassen. Das derzeit bekannteste Beispiel für ein Bildungsmonitoring ist die PISA-Studie der OECD, die alle drei Jahre Basiskompetenzen in den Bereichen Lesen, Mathematik und Naturwissenschaften international vergleichend erhebt. Zielkriterien sind hier allerdings keine nationalen Bildungsstandards, sondern Kompetenzmodelle, die von international besetzten Expertengruppen auf der Basis der pädagogischen und fachdidaktischen Forschung entwickelt wurden. Das für die PISA-Studie in Deutschland zuständige Konsortium hat innerhalb dieser Kompetenzmodelle gewissermaßen Minimalstandards bestimmt: bei den Angehörigen der untersten Kompetenzstufen wurde davon gesprochen, dass sie ein hohes Risiko hätten, erfolgreich die Schule abzuschließen und in eine qualifizierte Berufsausbildung überzuwechseln (Baumert et al. 2001).

Auf einem relativ hohen Aggregationsniveau hat das Bildungsmonitoring auch evaluative (bewertende) Aspekte, z.B. im Vergleich der Länder, Schulformen oder maximal auf Schulebene. Der Grad der Zielerreichung eines Landes, einer Schulform oder einer einzelnen Schule macht natürlich eine evaluative Aussage über die betreffende Teilpopulation. In Übereinstimmung mit dem Sprachgebrauch der Bildungsforschung reservieren wir hier jedoch den Begriff „Evaluation“ für den Fall, dass der Erfolg einer *bestimmten* Maßnahme bzw. die Leistungsfähigkeit einer *einzelnen* Institution (z.B. einer Schule) zu beurteilen ist. Außerdem zielen Evaluationsstudien – anders als das Bildungsmonitoring, das breit und unspezifisch angelegt ist – auf konkrete Entscheidungen, z.B. für oder gegen die Fortführung eines bestimmten Unterrichtskonzeptes, für oder gegen die Einführung zusätzlicher Fördermaßnahmen usw. Evaluationsstudien setzen daher eine detaillierte Erfassung der zu evaluierenden Maßnahmen und Faktoren und eine Einigung bei anstehenden Entscheidungen voraus. Sie setzen aber auch voraus, dass die eingesetzten Messinstrumente genau zu den Zielen der evaluierten Institution passen bzw. diejenigen Leistungsaspekte erfassen, die mit der evaluierten Maßnahme angesprochen und gefördert werden sollten.

Im Allgemeinen ist es daher nicht möglich, beiden Anliegen (Monitoring und Evaluation) im Rahmen derselben Studie nachzukommen. Bildungsmonitoring erfordert eine möglichst breite Erfassung verschiedener Kompetenzbereiche, Evaluation eine maßgeschneiderte, relativ enge Erfassung der kritischen Kompetenzaspekte (und sonstigen Ziele), auf die der Unterricht bzw. die Schule insgesamt ausgerichtet war. Bildungsmonitoring sollte auf Systemebene stattfinden und möglichst viele, auch heterogene Schulen einbeziehen. Evaluationsstudien beziehen sich auf ganz bestimmte Schulen und Klassen, denen ein maßgeschneidertes Feed-back gegeben werden soll. Bildungsmonitoring muss zentral, von übergeordneten Behörden und Ministerien, geplant werden, während Evaluation von den Schulen ausgehen kann.

Allerdings ist es durchaus möglich, ein im Bildungsmonitoring bewährtes Instrument auch bei der Evaluation einer Einzelschule zu nutzen – vorausgesetzt, der Test bzw. die Fragebogenskala erfasst einen Aspekt, der für die zu evaluierende Schule bedeutsam ist. Unter Nutzung von Daten des Bildungsmonitoring, das in diesem Fall als Normierungsstudie verwendet würde, könnte sich eine Schule im Vergleich zu anderen Schulen verorten. Beispielsweise könnte es für eine Versuchsschule mit einem besonderen Profil im ästhetisch-künstlerischen Bereich wichtig sein, nachzuweisen, dass sie in den „harten“ Fächern wie Mathematik und Naturwissenschaften nicht hinter vergleichbare Schulen zurückfällt. Dieses Ziel ließe sich mit Hilfe etwa des TIMSS-Tests, der bundesweit Mathematik- und Naturwissenschaftsleistungen gemessen hat, überprüfen.

## **9.2 Nutzung von Standards im Bildungsmonitoring**

In den allermeisten Industriestaaten gehört ein regelmäßiges Bildungsmonitoring seit Jahren zur Routine im Bildungswesen. Eine Ausnahme bilden praktisch nur noch die deutschsprachigen Länder, in denen bislang eine reine Input-Steuerung stattfindet. Hier besteht traditionell eine starke Zurückhaltung gegenüber standardisierten Leistungsmessungen, die häufig als Übergriff der Bildungsverwaltung, als Angriff auf die professionelle Verantwortung der Lehrkräfte empfunden werden. Bis etwa Mitte der 80er Jahre galt dies ähnlich für die skandinavischen und etliche angelsächsische Staaten, die inzwischen jedoch ihre Bildungspolitik geändert haben. Unter Experten gilt dies als einer der wesentlichen Gründe für das bessere Abschneiden dieser Staaten in internationalen Leistungsvergleichen – und zwar nicht bloß wegen der Vertrautheit der Schüler und Schülerinnen mit Tests, sondern aus substantziellen Gründen. Regelmäßiges Bildungsmonitoring gibt auf der Ebene des Gesamtsystems wesentliche Informationen zur Effektivität von Schulen und zu den Wirkungen von Reformmaßnahmen, auf die Bildungspolitik und Bildungspraxis reagieren können.

Bei der Ausgestaltung von Monitoring-Programmen gibt es eine große Zahl von Freiheitsgraden, die in den Staaten sehr unterschiedlich genutzt werden. Einige der Fragen, die durchaus auf verschiedene Weise beantwortet werden können, lauten:

- Werden nur Stichproben untersucht oder werden alle Schulen in das Monitoring-Programm einbezogen?
- Ist die Teilnahme für Schulen bzw. für einzelne Schülerinnen und Schüler, Lehrerinnen und Lehrer freiwillig oder verpflichtend?
- In welchem Rhythmus werden Erhebungen durchgeführt, auf welche Altersstufen und Fächer sind sie bezogen?
- Wie wird mit Ergebnissen der einzelnen Schule bzw. einer einzelnen Klasse umgegangen? Werden sie überhaupt berechnet, oder bleibt es bei der Auswertung auf Systemebene? Werden sie veröffentlicht? Ist intendiert, dass sich Eltern bei der Schulwahl an den Ergebnissen orientieren? Erhält die Schulaufsicht Einsicht in die Daten oder werden sie nur der Schule bzw. den Lehrkräften selbst als feed-back zur Verfügung gestellt? Werden die Schulergebnisse als „Rohwerte“, etwa als durchschnittlich erreichte Punktzahl in einem Test, mitgeteilt, oder werden sie „adjustiert“, um die Rahmenbedingungen der Schule (die soziale Herkunft der Schülerinnen und Schüler, die Eingangsleistungen, den Anteil von Schülern mit Migrationshintergrund usw.) zu berücksichtigen und so eine „faire“ Beurteilung zu erreichen?
- Welche Art von Tests wird eingesetzt? Sind es reine Kenntnisprüfungen mit einfachen Multiple-choice-Fragen, wie es etwa traditionell in den USA der Fall war, oder werden Kompetenzorientierte Verfahren eingesetzt, die auch komplexe, handlungsorientierte Aufgaben, Gruppenaufgaben und ähnliches einschließen? Werden die Ergebnisse rein normorientiert oder kriteriumsorientiert interpretiert (vgl. Kapitel 7)?
- Wer ist für Planung, Durchführung und Auswertung der Tests verantwortlich? Welche Arbeitsteilung besteht hierbei zwischen den Schulen selbst, den staatlichen Behörden, unabhängigen Testagenturen und kommerziellen Anbietern?

Es lohnt sich, die Erfahrungen anderer Staaten in diesem Zusammenhang genau und systematisch aufzuarbeiten. Beispielsweise hat sich in den USA, dem Staat mit der größten Dichte an groß angelegten Leistungsmessungen (*large scale assessments*), inzwischen die Einsicht verbreitet, dass zu häufiges Testen und zu oberflächliches Messen negative Folgen für die Qualität des Unterrichts haben können. Gleichzeitig sind in den USA sehr differenzierte Modelle für Leistungserhebung und -bewertung (*assessment*) und Schulrückmeldung (*accountability*) entwickelt worden, die jene Fehler vermeiden (vgl. National Research Council 2001). In Großbritannien hat man erkannt, dass ein öffentliches *Ranking* von Schulen auf der Basis von Testmittelwerten unfair und häufig kontraproduktiv ist.

In Deutschland ist das Thema „Bildungsmonitoring“ durch TIMSS und PISA wieder aktuell geworden. Im Anschluss an TIMSS haben mehrere Länder Monitoring-Studien durchgeführt, die hinsichtlich der oben genannten strategischen Fragen durchaus unterschiedliche Wege gingen. Die Unterschiede sind in der deutschen Öffentlichkeit, aber auch unter Fachleuten bislang unzureichend beachtet worden. Immer wieder wird beispielsweise der Unterschied zwischen Stichprobenuntersuchungen, wie TIMSS und PISA sie darstellen, und flächendeckenden Erhebungen (z.B. die Lernausgangslagenuntersuchung LAU in Hamburg oder die rheinland-pfälzische MARKUS-Studie) oder der Unterschied zwischen *Ranking* und kriteriumsorientierter Rückmeldung an die Schulen nicht wahrgenommen.

Diese Expertise ist nicht der geeignete Ort, um die Fragen der Gestaltung von Monitoring-Studien umfassend zu beantworten (vgl. dazu aber den von Weinert 2001 im Auftrag der Kultusministerkonferenz herausgegebenen Band „Leistungsmessungen in Schulen“). Sinnvollerweise wird man in den Ländern der Bundesrepublik Deutschland auch weiterhin unterschiedliche Wege gehen, verschiedene Varianten des Bildungsmonitoring ausprobieren und voneinander lernen. Dabei geben die internationalen Leistungsvergleiche, an denen sich Deutschland inzwischen regelmäßig beteiligt, einen Rahmen und Rhythmus vor, an die nationale oder landesspezifische Untersuchungen „angehängt“ werden können. Beispielsweise bietet es sich an, zukünftige Erhebungen im Bereich der Sekundarstufe I an die alle drei Jahre stattfindende PISA-Studie der OECD zu koppeln, wie es bereits bei PISA/PISA-E 2000 der Fall war und für PISA 2003 geplant ist (vgl. unten, S. 104f.).

Im Kontext dieser Expertise ist es von zentraler Bedeutung, dass mit der Einführung nationaler Bildungsstandards eine tragfähige Basis für aussagefähige Studien gegeben ist. Sind erst einmal (wie in Kapitel 7 beschrieben) Testverfahren entwickelt, welche die Standards operationalisieren und eine Diagnose der erreichten Kompetenzstufe ermöglichen, so können die Verfahren auch im Bildungsmonitoring eingesetzt werden. Damit wäre das nationale Bildungsmonitoring nicht mehr, wie bislang, darauf angewiesen, die PISA-, TIMSS- oder IGLU-Skalen als Vergleichskriterium zu nutzen, sondern es könnte sich auf deutschlandspezifische, bundesweit gültige Messinstrumente und Bewertungsvorschriften stützen. Die Verankerung dieser Tests in fachdidaktisch und psychologisch fundierten Kompetenzmodellen gewährleistet, dass die Resultate nicht zu einem bloß normorientierten Vergleich zwischen Staaten, Ländern oder Schulen führen, sondern kriteriumsorientiert, als Aussagen über den Stand erworbener Kompetenzen interpretiert werden. PISA hat hier Qualitätsmaßstäbe gesetzt, die auch international anerkannt sind und von nationalen Testinstrumenten nicht unterboten werden sollten. Auf dieser Grundlage können unerwünschte Wirkungen, wie sie mit oberflächlichen *large scale assessments* vor allem in den USA verbunden waren, vermieden werden.

Nicht zwingend, aber naheliegend und für die Qualitätssicherung in Deutschland hilfreich wäre es, *nationale* Bildungsstandards durch ein *nationales* Bildungsmonitoring zu ergänzen. So könnten alle drei bis fünf Jahre, an wechselnden Stichproben, die in den Standards festgelegten Kom-

petenzen bundesweit und ggf. auch im Ländervergleich überprüft werden. Wie bislang bei internationalen und nationalen Schulleistungsstudien üblich, könnte der Auftrag, ein solches Monitoring zu gestalten und durchzuführen, aufgrund einer offenen Ausschreibung an wechselnde wissenschaftliche Institute bzw. Konsortien vergeben werden. Denkbar wäre aber auch die Einrichtung eines eigenen Instituts für diese Aufgaben. Zumindest sollte es eine nationale Instanz, eine „Agentur für Bildungsmonitoring“ geben, die für kontinuierliche Erhebungen sorgt, Aufträge an Forschergruppen vergibt und für Qualitätssicherung zuständig ist. Diese Agentur kann, muss aber nicht mit der in Abschnitt 7.3 genannten Testentwicklungs-Agentur identisch sein (siehe dazu auch Kapitel 11, S. 107).

Nach jeder Erhebung im Rahmen des nationalen Bildungsmonitorings kann ein Teil der verwendeten Aufgaben für Monitoring- und Evaluationsstudien in den Ländern freigegeben werden. So wurde beispielsweise in Österreich aus freigegebenen TIMSS-Aufgaben ein Test zusammengestellt, mit dem Schulen ihre Lernergebnisse selbst prüfen können. Um die freigegebenen Aufgaben zu ersetzen, findet eine „rollende Testkonstruktion“ statt: Bei jeder Erhebung laufen neue Aufgaben zur Erprobung mit, deren Passung zum Test und zum Kompetenzmodell empirisch geprüft wird und die dann beim nächsten Erhebungstermin den eigentlichen Test bilden. Auf diese Weise lässt sich ein Testverfahren laufend aktualisieren und der Fundus der für unterschiedliche Zwecke nutzbaren Aufgaben laufend erweitern, während zugleich die Vergleichbarkeit der Messungen über Jahre hinweg garantiert ist.

Das Design einer Monitoringstudie (Zusammenstellung von Test- und Fragebogenkomponenten, Bestimmung der Stichprobe u.a.m.) und die wissenschaftliche Auswertung bedürfen des Beitrags unterschiedlicher wissenschaftlicher Disziplinen: empirische Bildungsforschung, pädagogisch-psychologische Methodenlehre, Schulpädagogik, Fachdidaktiken. Daher hat es sich in der Vergangenheit bewährt, interdisziplinäre Konsortien mit solchen Vorhaben zu betrauen. Die Logistik der Testapplikation kann demgegenüber in die Hände kommerzieller Institute gegeben werden. Derzeit ist es bei Monitoringstudien in Deutschland üblich, schulexterne, besonders geschulte Kräfte als Testleiter einzusetzen. Diese Funktion könnten bei innerdeutschen Studien – wie in anderen Staaten üblich – Lehrer und Lehrerinnen einnehmen, was für die Akzeptanz und Identifikation mit den Testinstrumenten von entscheidender Bedeutung sein kann. Die Testauswertung und Verarbeitung der Ergebnisse (Kodierung der Antworten, Kontrolle der Auswertungsobjektivität, die Prüfung der Messgenauigkeit und Aussagekraft der Tests sowie die Bereitstellung deskriptiver Ergebnisse zur Einhaltung von Bildungsstandards für die Länder, Schulformen und einzelne Schulen) muss wiederum zentral erfolgen, möglicherweise unter Beteiligung von Dienstleistungsunternehmen, die auf das Datenmanagement in solchen Studien spezialisiert sind. Die Auswertung sollte aber insgesamt in der Verantwortung der beauftragten Wissenschaftler verbleiben.

### 9.3 Evaluation auf Schulebene

Monitoring- und Evaluationsstudien sollten grundsätzlich als zwei verschiedene Typen von empirischen Studien angesehen werden. Jeder Forschungstyp ist seinen eigenen Qualitätsstandards verpflichtet; es können jedoch streckenweise dieselben Testinstrumente verwendet werden. Während das Bildungsmonitoring in die Zuständigkeit einer zentralen Agentur fällt (siehe oben), sollten Evaluationsstudien von den Schulen ausgehen, in Modellversuchsprogrammen angesiedelt sein oder im Rahmen (fach-)didaktischer Forschung stattfinden.

Verschiedene Strategien sind denkbar und zum Teil in Deutschland auch bereits erprobt worden: Die Länder können Berater für die Selbstevaluation von Schulen bereitstellen oder gesonderte Evaluationsprogramme auflegen, an denen sich Schulen freiwillig beteiligen können. Denkbar ist auch ein freier Markt, auf dem verschiedene Anbieter Evaluationsinstrumente bereitstellen, die zudem Durchführung, Auswertung und Beratung der Schulen übernehmen. Schulen müssten in diesem Fall über ein Budget verfügen, das es ihnen erlaubt, solche Evaluationen „einzukaufen“. Die im Zusammenhang mit dem Bildungsmonitoring auf S. 84 genannten Fragen zur konkreten Ausgestaltung gelten auch für die Schulevaluation. Von entscheidender Bedeutung ist hier, ob eine regelmäßige Evaluation (wie etwa in Skandinavien und den Niederlanden) für die Schulen verpflichtend gemacht wird, wer wie mit den Daten umgehen soll und welche Konsequenzen (z.B. im Sinne von zusätzlichen Beratungsangeboten) ein für die Schule problematisches Evaluationsergebnis hat. Mit diesen Entscheidungen sind eine Reihe von pädagogischen, innovationsstrategischen, juristischen und ethischen Fragen verbunden, die sorgfältig abgewogen werden müssen. Betrachtet man jedoch die Erfahrungen der vergangenen Jahre aus dem Ausland, so ist zu erwarten, dass Schulevaluation – wenn sie tatsächlich auf die pädagogischen Fragen der Schulen abgestimmt ist – als Rückmeldung willkommen sein wird und wichtige Impulse für die pädagogische Schulentwicklung geben kann.

Durch die mit einer Rückmeldung aus der Evaluation zur Verfügung gestellten Daten erhalten die Schulen nicht nur zuverlässige Informationen über ihren gegenwärtigen Stand, sie können auch genauer festlegen, welchen Stand sie (in Zukunft) erreichen wollen, und – bei wiederholter Teilnahme – in welchem Umfang sie sich verändert haben.

Die Qualität der Rückmeldung ist mit entscheidend für

- die Akzeptanz der Ergebnisse in den Schulen;
- die schulspezifische Analyse von Ursachen der Ergebnisse;
- eine zielgerichtete Planung und Umsetzung von Verbesserungen;
- wirksame Vereinbarungen zwischen der Schule und ihren Unterstützungssystemen.

Die Veröffentlichung von einzelschulischen Ergebnissen im Sinne eines „Rankings“ oder in „Ligatabellen“ hat sich als kontraproduktiv erwiesen und sollte deshalb unterbleiben.

Eine Rückmeldung dürfte als „fair“ eingestuft werden, wenn bei der Schulrückmeldung neben den reinen Leistungsdaten der Schülerinnen und Schüler zumindest deren Muttersprache, das Geschlecht, Schul- und Berufsbildung der Eltern, sozioökonomischer Status der Eltern und andere Kontextfaktoren berücksichtigt werden. Solche „Adjustierungen“ gehören inzwischen zum selbstverständlichen Repertoire der Schulforschung und -evaluation. Als beispielhaft für Rückmeldungen über die Ergebnisse von Schulen mit vergleichbarer Schülerschaft kann die PISA-Studie gelten (Stanat et al. 2002).

Erfahrungen mit Schulevaluation und Rückmeldungen aus empirischen Erhebungen zeigen jedoch auch, dass die meisten Schulen externe Beratung brauchen, um eigene Fragestellungen für eine Evaluation zu entwickeln, Ergebnisse zu interpretieren und Strategien für die weitere Schulentwicklung aufzustellen. Ohne entsprechende Unterstützungs- und Beratungsangebote (vgl. Kapitel 10) besteht die Gefahr, dass Evaluationsstudien unproduktiv bleiben und langfristig von den Betroffenen in den Schulen als Ballast empfunden werden.

#### **9.4 Konsequenzen für die Diagnose und Förderung einzelner Schüler**

Das Erhebungsdesign einer bundesweiten Monitoringstudie ist im Allgemeinen *nicht* so ausgelegt, dass es Aussagen auf Individualebene erlaubt. Das kleinste Aggregat, zu dem hinreichend genaue und valide Messergebnisse abgeleitet werden können, ist die Schule oder unter Umständen die Klasse. Dasselbe gilt für schulbezogene Evaluationen. Dies hat mehrere Gründe. Zum einen werden Bildungsstandards stets ein weites Spektrum von Kompetenzen umfassen, die schon allein zeitlich nicht bei jedem Schüler getestet werden können. Um das Spektrum auf Schulebene erfassen zu können, wird es notwendig sein, dass verschiedene Schüler unterschiedliche Testteile bearbeiten (sog. Rotationsdesign; vgl. Kapitel 7). Damit ist aber eine Vergleichsbasis für Individualvergleiche nicht mehr gegeben. Ein weiterer Grund ist die Messgenauigkeit, die auf der Ebene des einzelnen Schülers bzw. der einzelnen Schülerin zu gering ist. Aus diesen Gründen sollte eine Weitergabe von Individualergebnissen aus den Monitoring- und Evaluationsstudien an die Schulen und Lehrer in der Regel nicht erfolgen.

Andererseits erfassen die Tests natürlich *individuelle* Kompetenzen, und ihr Einsatz für individualdiagnostische Zwecke ist im Prinzip möglich. Inwieweit Schulevaluationen Aussagen über einzelne Schülerinnen und Schüler erlauben, hängt von dem Design der Studie ab. Zielt eine Evaluation darauf ab, die Leistungen von Schulklassen hinsichtlich bestimmter Kriterien

zu vergleichen, können die Testinstrumente eher breit angelegt sein, um alle Aspekte der Klassenunterschiede zu erfassen, Unter Umständen sind sie dann wenig aussagekräftig hinsichtlich des individuellen Profils eines einzelnen Schülers. Ist die Evaluation dagegen so angelegt, dass sie den Lernerfolg der Schüler unter anderem auf deren individuelle Lernvoraussetzungen zurückführen will, so sollte die individualdiagnostische Qualität der Testergebnisse natürlich gegeben sein.

Grundsätzlich ist es wichtig, die Grenzen der individualdiagnostischen Aussagekraft einzelner Testanwendungen zu beachten. Viele Probleme, die im Zusammenhang mit dem sogenannten *high stakes assessments* in den USA entstanden sind (also Testerhebungen, die für die teilnehmenden Personen unmittelbare Konsequenzen haben, etwa zur Steuerung von Bildungslaufbahnen, oder die zur Mittelvergabe an Schulen herangezogen werden), hätten sich vermeiden lassen, wenn dort sorgfältiger zwischen Bildungsmonitoring, Evaluation auf Schul- und Klassenebene und Individualdiagnostik unterschieden worden wäre (vgl. auch National Research Council 2001). Jede Art von individueller Leistungsrückmeldung an die Schüler (und sei es auch nur zu Händen des Lehrers) hat psychologische Implikationen auf das Selbstkonzept der Schüler und auf ihre weitere Entwicklung. Ein Testinstrument, das für Zwecke des Monitorings und der Evaluation gut ist, muss nicht unbedingt individuelle Testfairness garantieren. Andererseits beruhen die standardbezogenen Tests auf Modellen der individuellen Kompetenzentwicklung und sind daher hervorragend geeignet, auch für individualdiagnostische Zwecke eingesetzt zu werden. Dies muss aber im Rahmen eines adäquaten Testdesigns geschehen und mit der gebotenen pädagogischen Verantwortung erfolgen. Die Konsequenzen, die aus den Ergebnissen von Monitorings und Schulevaluationen unter Zugrundelegung von Kompetenzmodellen für die Verbesserung von Schule folgen, kommen dem einzelnen Schüler auch dann zugute, wenn das individuelle Testergebnis gar nicht rückgemeldet wird. Testapplikationen mit individualdiagnostischer Zielsetzung sollten in der Verantwortung des Lehrers stehen.

## 10. Aufgaben der Unterstützungssysteme

Mit der Einführung von Bildungsstandards wird ein weitreichender Eingriff in ein eingespieltes Schulsystem unternommen. Freilich sind umfassende und weitreichende Veränderungen erforderlich, wie die aktuellen Befunde über die Leistungsfähigkeit von deutschen Schulen und ihren Schülerinnen und Schülern zeigen. Die Konzeption von Bildungsstandards, die in dieser Expertise begründet wurde, verlangt ein Umdenken hinsichtlich der Steuerung des Schulsystems. Damit sind einerseits neue Anforderungen und Aufgaben, andererseits aber erweiterte Handlungsspielräume und Gestaltungsmöglichkeiten für die Schulen verbunden (vgl. Kapitel 4).

Die Einführung und sukzessive Umsetzung von Bildungsstandards ist ein langfristiger Prozess, der vorbereitet, begleitet und unterstützt werden muss. Ob Bildungsstandards die erwünschten Effekte im Bildungssystem haben werden, hängt nicht nur von der Qualität der Kompetenzmodelle und Testverfahren ab. Mindestens ebenso wichtig ist eine professionelle Nutzung der Bildungsstandards. Die Zielbeschreibungen legen die Richtung fest; sie binden die Schulen und Lehrkräfte. Dennoch kommt es darauf an, dass die Lehrkräfte – und mittelbar auch die Schülerinnen und Schüler bzw. die Eltern – diese Ziele durchdringen und sich zu eigen machen. Auch Kompetenzmodelle und Testverfahren legen das pädagogische Handeln nicht im Einzelnen fest. Sie sind vielmehr Hilfsmittel, besser gesagt mächtige Werkzeuge, um den Unterricht auf die Ziele auszurichten und gegebenenfalls nachzusteuern. Doch will der fachgerechte Einsatz und der geschickte Umgang mit diesen neuen Werkzeugen gelernt sein. Bildungsstandards können somit nur unter aktiver Mitwirkung aller Beteiligten ihre Wirkung entfalten. Eine Schlüsselrolle haben dabei die Schulleitungen und die Lehrkräfte.

Ein Blick auf Erfahrungen anderer Staaten mit der Einführung nationaler (Kern-) Curricula und Standards zeigt, wie wichtig es dabei ist, die Schulen angemessen zu unterstützen. Die Einführung des nationalen Curriculums in England beispielsweise wurde von Anfang an im Zusammenhang mit der Professionalisierung von Lehrerinnen und Lehrern gesehen. Schriftliche Materialien, die in großem Umfang erstellt wurden, wurden durch Lehrerfortbildungsveranstaltungen ergänzt. Es liegt auf der Hand, dass ein Curriculum, das wenig inhaltsbezogene Anteile enthält und kaum etwas über Unterrichtsmethoden und didaktische Grundlagen aussagt, nur mit einem relativ großen Aufwand in konkreten Unterricht umgesetzt werden kann. Es stellt sich außerdem die Frage, ob Fortbildung alleine genügt. Gerade im Bereich der Primarstufe wurde in England lebhaft diskutiert, ob die fachlichen Grundlagen von Lehrerinnen und Lehrern ausreichen, um letztendlich in allen Gegenstandsbereichen aus den wenigen Vorgaben ein eigenes Curriculum für den Unterricht in der eigenen Schule und Klasse zu entwickeln.

Die Entscheidung, Bildungsstandards zu entwickeln und im deutschen Schulsystem zur Wirkung zu bringen, setzt ein Konzept zur Implementation voraus. Diese Implementation zielt darauf ab,

1. das Verständnis und die Akzeptanz von Bildungsstandards zu sichern,
2. in die Arbeit mit Bildungsstandards einzuführen und
3. mit den Möglichkeiten der professionellen Nutzung und Auswertung von Bildungsstandards vertraut zu machen.

Für die Schulen müssen dementsprechend Information, Anleitung, Fortbildung und Beratung vorgehalten und angeboten werden. Sie müssen insbesondere damit rechnen können, Hilfe zu erhalten, wenn die Auswertungen von Tests Probleme und Handlungsbedarf erkennen lassen. Insgesamt müssen also frühzeitig vielfältige Unterstützungen bereitstehen, damit die Standards von den Schulen angenommen und wie geplant umgesetzt werden können.

Es liegt auf der Hand, hier systematisch die Einrichtungen mit einzubeziehen, deren Auftrag es ist, die Arbeit der Schulen zu unterstützen und für die Qualifikation und Professionalität der Lehrkräfte zu sorgen. So werden

- die Einrichtungen der Lehrerbildung (Universitäten, Studienseminare, Fortbildungsinstitute),
- die Landesinstitute und
- die Schulaufsicht

zu wichtigen Unterstützungssystemen, zunächst für die Einführung von Bildungsstandards, dann aber dauerhaft für die Umsetzung im schulischen Alltag und für die Lösung von auftretenden Problemen. Im Folgenden werden wichtige Funktionen und Aufgaben dieser Unterstützungssysteme für die Arbeit mit Bildungsstandards skizziert.

Allerdings hängen die Möglichkeiten dieser Einrichtungen, entsprechende Unterstützungen zu geben, von *Rahmenbedingungen* ab, die von Anfang an berücksichtigt werden müssen. Dazu zählen nicht nur die erforderlichen Ressourcen und Qualifikationen oder frühzeitige Absprachen, Vereinbarungen und Planungen. Eine grundlegende Rahmenbedingung für die Arbeit der Unterstützungssysteme ist die Ausformulierung der bildungspolitischen Begründung für die Einführung von Bildungsstandards und die Darlegung der langfristigen Perspektiven. Für die Akzeptanz entscheidend werden offizielle 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. Die Schulen sollten Kompetenzmodelle, Tests und die darauf bezogenen Informationen, Handreichungen und Fortbildungen als Wegbeschreibungen, als Orientierungshilfen, als Werkzeuge und Handlungsgerüste verstehen können und nicht als neuerliche Vorgaben oder verordnete Aufträge, die ab-

zuarbeiten sind. 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. Das bedeutet, die Lehrkräfte in ihrer Professionalität anzusprechen und auf diese – mit klaren Erwartungen – zu vertrauen. Vor diesem Hintergrund wird es den Schulen sehr viel leichter fallen, Angebote und Aktivitäten der Unterstützungssysteme zu Bildungsstandards als Herausforderung und Hilfe zu sehen und anzunehmen.

## 10.1 Unterstützung der Akzeptanz

Für die Akzeptanz von Bildungsstandards und ihren vielfältigen Folgen für die tägliche Arbeit spielen vor allem vier Faktoren eine wichtige Rolle. Der erste und sicher wichtigste Faktor betrifft die inhaltliche Begründung des ganzen Unternehmens. Hinreichende Gründe für die Einführung von Bildungsstandards findet man nicht zuletzt in dieser Expertise. Entsprechende Begründungen werden dann leichter nachvollzogen, wenn sie Probleme zum Ausgangspunkt nehmen, die den Lehrkräften vertraut sind bzw. sie ebenfalls in ihrer alltäglichen Arbeit belasten. Während dieser erste Faktor für die Akzeptanz rational oder kognitiv verarbeitet wird, beeinflussen die anderen drei Bedingungen eher die Motivation und die Einstellung gegenüber Bildungsstandards. Diese werden, so der zweite Faktor, dann leichter akzeptiert, wenn sie aus der Perspektive der Lehrkräfte ihre Gestaltungsmöglichkeiten erweitern und ihnen mehr Chancen bieten, Unterricht nach ihrem professionellen Sach- und Selbstverständnis zu planen und durchzuführen. Die Öffnung der Handlungsmöglichkeiten darf allerdings nicht Unsicherheit produzieren (im Sinne einer völligen Beliebigkeit) oder mit dem Eindruck verbunden sein, dass Aufgaben abgewälzt werden („nun müssen die Lehrkräfte/Schulen auch noch die Lehrpläne selbst entwickeln“). Spielräume gewinnen die Lehrkräfte auf der Basis einer guten Orientierung über Handlungsmöglichkeiten und Vorgehensweisen, die ihnen viele Freiheitsgrade bietet und die Wahl lässt. Der dritte Faktor betrifft die Erwartung, dieses neue System der Bildungsstandards mit all seinen Aspekten in den Griff zu bekommen und mit ihm kompetent umgehen zu lernen. Bildungsstandards konfrontieren die Schulen und Lehrkräfte mit Entwicklungsaufgaben. Für die Schulen und Lehrkräfte muss erkennbar sein, was sie zu tun haben, was auf sie zukommt, dass sie die Anforderungen bewältigen können und wie dies möglich ist. Bereits der Einstieg in die Arbeit mit Bildungsstandards sollte zur Erfahrung führen, dass diese Arbeit einen neuen Blick ermöglicht, dass sie nützlich ist und gemeistert werden kann. Die vierte Faktor schließlich betrifft die Einschätzung, dass die Einführung von Standards eine gemeinsame professionelle Aktivität ist, bei der nicht nur die Schulen und Lehrkräfte zusammenarbeiten, sondern weitere engagierte Gruppen einbezogen sind, um das Unternehmen zu einem Erfolg zu bringen.

Was können nun die einzelnen Unterstützungssysteme zur Förderung der Akzeptanz leisten? Es liegt auf der Hand, dass die Einrichtungen der Lehrerbildung – in allen drei Phasen – insbesondere zu einem differenzierten Verständnis von Bildungsstandards beitragen können. Das bedeutet zunächst, die erforderliche Erstinformation über Standards und ihre Funktion im Schulsystem bereit zu stellen. Die Einrichtungen der Lehrerbildung sollten dann aber von Anfang an darlegen, wie sie die Lehrkräfte mit dem erforderlichen *Know-how* zur Arbeit mit Standards versorgen wollen. Es wird dann eher Akzeptanz gewonnen werden, wenn systematische Fortbildungsprogramme zur Arbeit mit Standards angekündigt und aufgelegt werden. Einrichtungen der Lehrerbildung können somit zur rationalen Begründung und zum Verständnis der Standards beitragen, auf gesichertes Wissen verweisen (Forschungsbefunde, Kompetenzmodelle), das die neuen Aufgaben strukturiert, gezielte Ausbildungsprogramme vorbereiten und so unterstreichen, dass Bildungsstandards eine gemeinsame Herausforderung darstellen, die gleichermaßen Forschungs- und Fortbildungsanstrengungen verlangt.

Auch die Landesinstitute können maßgeblich bei der Erstinformation über Bildungsstandards mitwirken. Ihr Beitrag dürfte sich besonders darauf richten zu zeigen, wie sie die Arbeit der Schulen unterstützen werden, durch Handreichungen, durch die Modifikation von Lehrplänen, durch die Bereitstellung von flexibel einsetzbaren Materialien, durch ein intelligentes Informationsmanagement und vor allem durch eine bedarfsbezogene Beratung zu speziellen Fragen der Standards (z.B. Kompetenzmodelle oder Tests) und zum breiteren Fragenkreis der Schulentwicklung.

Der Beitrag der Schulaufsicht kann sich demgegenüber weniger auf Informationsaufgaben richten. Um die Akzeptanz zu fördern, sollte sie frühzeitig ihre neue Rolle in der Arbeit mit Bildungsstandards klären und darlegen. Auch für die Schulaufsicht wird der Output der entscheidende Bezugspunkt werden und sie in der Doppelfunktion von Aufsicht und Beratung stark beschäftigen. Dies betrifft aber vor allem den Umgang mit den Ergebnissen aus den Tests im Rahmen des Bildungsmonitoring. Hier sollte von Anfang an erkennbar sein, dass sich das Anliegen der Schulaufsicht vor allem darauf richten muss, konstruktiv zur Lösung von eventuellen Problemen beizutragen und gegebenenfalls Hilfen bereitzustellen. Ansonsten unterstützt die Schulaufsicht die Akzeptanz von Bildungsstandards, wenn sie deren Einführung begrüßt; sich aber Zurückhaltung im Prozess der Implementation auferlegt. Das bedeutet etwa, Beratung für die Umsetzung der Standards anzubieten, die von Schulen angefordert werden kann.

## **10.2 Einführung in die Arbeit mit Bildungsstandards**

Die Arbeit mit Bildungsstandards an Schulen richtet sich primär auf den Fachunterricht und die dort aufzubauenden Kompetenzen. Sie umfasst didaktische Klärungen und Entscheidungen, zum Beispiel über Auswahl, Schwerpunktsetzungen und Sequenzierungen im Lehrstoff in Hinblick auf

Zielbeschreibungen und Kompetenzmodelle. Sie befasst sich mit Unterrichtskonzeptionen und Unterrichtsstrategien, etwa unter der Frage, wie unterschiedliche Lernvoraussetzungen angesprochen, das Verständnis gesichert und Routinen geübt oder eine flexible Wissensanwendung unterstützt werden kann. Zur Arbeit mit Standards gehört dann aber auch, Lernvoraussetzungen und Lernfortschritte, Stärken und Schwächen der Lerner und Lernerinnen zu erkennen bzw. zuverlässig zu diagnostizieren und geeignete Förderungsmaßnahmen zuzuordnen. Diese Anforderungen betreffen auf einer ersten Ebene jede Lehrkraft in Hinblick auf ihren individuellen Unterricht; auf einer zweiten Ebene jedoch das Fachkollegium bzw. die gesamte Schule. Die hier vertretene Konzeption der Bildungsstandards sieht eine enge Abstimmung bei didaktischen Entscheidungen auf der Schulebene vor. Insbesondere sollen Anforderungen über die Klassenstufen hinweg koordiniert werden. Schulintern abzusprechen sind nicht nur die Schnittstellen und Kriterien. Auf längere Sicht gilt es, ein schulinternes Curriculum auszuformen, das die besondere Ausgangslage der Schule und ein mögliches Schulprofil berücksichtigt und bisherige Lehrpläne gezielt weiter entwickelt. Abzustimmen sind weiterhin Vorgehensweisen zur schulinternen Prüfung von Lehr- und Lernerfolgen sowie die Einrichtung besonderer Fördermaßnahmen. Eine zentrale und schwierige Aufgabe schließlich erwächst aus den Rückmeldungen über die Stärken und Schwächen der Schule durch das regelmäßige Bildungsmonitoring oder die spezielleren schulbezogenen Evaluationen.

Die Anforderungen, die auf die Lehrkräfte und Schulen zukommen, lassen sich sehr klar einpassen in die Auflistung von zentralen Aufgaben und Kompetenzen des Lehrerberufs, die im Abschlussbericht der KMK-Kommission zu Lehrerbildung (Terhart 2000) beschrieben wurden. Sie reichen vom Kerngeschäft des Unterrichtens über das Diagnostizieren und Evaluieren von Schülerleistungen zur Weiterentwicklung der eigenen Professionalität bis zur Mitwirkung an der Entwicklung der Schule. Allerdings erkennt die Expertengruppe hier, dass für diese Aufgaben in allen Phasen der Lehrerbildung bisher kaum und zu wenig systematisch vorbereitet wird. Dennoch sind die Einrichtungen der Lehrerbildung dasjenige Unterstützungssystem, das für die Einführung in die Arbeit mit Bildungsstandards in einem besonderen Maße gefordert ist. Im Rahmen der grundständigen Lehrerbildung verfügen die Universitäten mit ihren fachdidaktischen und pädagogisch-psychologischen Instituten an vielen Standorten über das wissenschaftliche Potential, die zukünftigen Lehrkräfte auf die veränderten Anforderungen vorzubereiten. Die Einführung von Bildungsstandards bedeutet also eine konkrete Nachfrage nach wissenschaftlichen Erkenntnissen (z.B. Kompetenzmodelle) und problembezogener Qualifizierung (z.B. Unterrichtsplanung, Diagnostik, Evaluation). Die Universitäten selbst, aber auch die wissenschaftlichen Fachverbände können darauf hinwirken, dass diese Nachfrage bedient wird. An den Universitäten, an denen Zentren der Lehrerbildung oder solche für Lehr-Lern-Forschung eingerichtet wurden, bietet es sich an, über einen gewissen Zeitraum hinweg die Arbeit mit Standards in das Zentrum zu rücken. Qualifizierungsbedarf hinsichtlich einer anspruchsvollen Nutzung der Möglichkeiten von Bildungsstandards besteht dann vor allem in der zweiten Phase der Lehrerbildung. Die Qualifizierungslücken können kurzfristig nur durch eine enge Zusammenarbeit mit universitären Einrich-

tungen geschlossen werden. Dringend vorzubereiten sind Fortbildungsprogramme zur Arbeit mit Bildungsstandards für die Ausbilder in der zweiten Phase und Rahmenkonzeptionen für die Qualifizierung angehender Lehrer und Lehrerinnen während des Referendariats. Besondere Fortbildungsprogramme werden ebenfalls frühzeitig für die Schulleitungen anzubieten sein, die in besonderer Weise durch die Arbeit mit Bildungsstandards gefordert werden. Hauptlast an Qualifizierungsaufgaben wird insgesamt die Lehrerfortbildung zu tragen haben, wenn die Schulen flächendeckend auf die neuen Aufgaben vorbereitet werden sollen.

Die Landesinstitute können die Schulen und Lehrkräfte bei der Einführung von Standards auf sehr vielfältige Weise unterstützen, durch Materialien, Handreichungen, Informationsbörsen, Clearingstellen und durch unterschiedliche weitreichende Beratungsangebote. Besondere Aufmerksamkeit wird dabei dem Übergang von der bisherigen Lehrplanarbeit zur Arbeit mit Standards zu widmen sein. Die Schulen benötigen hier Handlungsgerüste, die Orientierung bieten, aber Wahlmöglichkeiten lassen und begründete didaktische Entscheidungen verlangen. In Anbetracht insgesamt knapper Ressourcen wird es darauf ankommen, die Unterstützungsleistungen von Landesinstituten und Einrichtungen der Lehrerbildung (inklusive Universitäten) so gut wie möglich aufeinander abzustimmen. Hervorzuheben ist jedoch, dass die Qualifizierung aller Lehrkräfte für eine professionelle Nutzung von Bildungsstandards nicht mit traditionellen Mitteln der Lehrerfortbildung erreicht werden kann. Herkömmliche Formen der Lehrerfortbildung können zunächst vor allem zu einer Vermittlung grundlegender Informationen dienen. Sodann wird es aber darauf ankommen, andere Zugänge zu entwickeln und zu verwenden. Verfahren der Schulentwicklung und die Arbeit an Schulprogrammen können genutzt werden, wenn sie tatsächlich auf den Fachunterricht und die Zusammenarbeit in der Fachgruppe fokussieren. Vor allem könnten Ansätze zur unterrichtsbezogenen Qualitätsentwicklung genutzt werden, die sich im Modellversuchsprogramm zur Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts bewährt haben (Prenzel 2000). Für die Arbeit mit Bildungsstandards könnten Module zugeschnitten werden, die von den Schulen bzw. Fachkollegien selbstständig bearbeitet werden könnten. Dabei könnten mehrere Module des Programms (etwa zum kumulativen Lernen, zur Sicherung des Grundwissens und zum verständnisvollen Lernen auf unterschiedlichen Ebenen, zum selbstständigen Arbeiten, zum Prüfen und Rückmelden von Kompetenzzuwachs oder zu schulinternen Standards) direkt aufgegriffen werden. Weitere Merkmale dieses Ansatzes könnten genutzt werden, etwa die Zusammenarbeit von Schulen in regionalen Netzwerken mit Unterstützung durch Koordinatoren, Landesinstitute und Universitäten, und die Bereitstellung von Ressourcen für den Austausch von Arbeitsergebnissen (z.B. betreute Server).

Wenn, wie hier vorgeschlagen wird, Landesinstitute und Einrichtungen der Lehrerbildung eng zusammenarbeiten und deren Beiträge aufeinander abgestimmt werden, dann ist selbstverständlich die Schulaufsicht mit einzubeziehen. Ansonsten sollte sie bei der Einführung in die Arbeit mit den Bildungsstandards keine Kontrollfunktionen ausüben, sondern vorwiegend für Beratung (speziell auch bei eventuell auftretenden Schwierigkeiten oder Konflikten) zur Verfügung stehen.

### 10.3 Unterstützung bei der Nutzung von Rückmeldungen

Ein erster Höhepunkt der Arbeit mit Standards wird dann erreicht sein, wenn die Schulen die Ergebnisse aus Bildungsmonitoring und Evaluation erhalten. Sie erfahren dann, inwieweit es ihnen gelungen ist, die angestrebten Kompetenzen bei ihren Schülerinnen und Schülern zu entwickeln. Sie erhalten gegebenenfalls weitere Rückmeldungen über Leistungsprofile, über Stärken und Schwächen auf der Schul- und eventuell auf der Klassenebene oder sonstige Anhaltspunkte für eine Reflexion ihrer Leistungsbeurteilung. In dieser Situation gilt es, die Information sorgfältig zu interpretieren und die Rückmeldung, wie auch immer sie ausfallen mag, konstruktiv zu werten. Als Prinzip muss gelten, dass besonders dann Unterstützung angeboten wird, wenn im Bildungsmonitoring Probleme erkennbar werden. Kritik und Schuldzuweisungen sind fehl am Platz. Im Rahmen der Ergebnissrückmeldung und ihrer Aufarbeitung werden also die Unterstützungssysteme in besonderer Weise beansprucht.

Es mag sein, dass hier auch ein besonderer Qualifizierungsbedarf an der Schule festgestellt wird, der durch Fortbildungsmaßnahmen zu decken ist. Ansonsten spielen Einrichtungen der Lehrerbildung in dieser Phase für die beteiligten Schulen eine nachgeordnete Rolle. Die Unterstützungsangebote müssen hier in erster Linie von den Landesinstituten in enger Zusammenarbeit mit der Schulaufsicht und der betroffenen Schule kommen. Die Landesinstitute können insbesondere Beratung zu Standards, Unterricht, Schulprogramm etc. anbieten. Die Schulaufsicht wird sich vor allem auch mit der Frage beschäftigen müssen, inwieweit die Schulen zusätzliche und spezifische personelle Unterstützung benötigen, und in Abstimmung mit der Schule ein Programm zur Behebung der Schwächen bzw. zur weiteren Schulentwicklung ausarbeiten müssen.

Insgesamt betrachtet bringt die Einführung von Bildungsstandards hohe Anforderungen an die Schulen mit sich, die von den Lehrkräften nicht ohne weiteres bewältigt werden können. Wenn langfristig die Qualität von Bildungsergebnissen durch Standards angehoben werden soll, dann müssen die Möglichkeiten, die Standards bieten, professionell genutzt werden. Der professionelle Umgang mit Standards muss seinerseits gelernt werden, in einem Entwicklungsprozess, der ebenso sukzessive voranschreitet wie die Ausarbeitung der Bildungsstandards selbst mit ihren genannten Komponenten. Die Schulen bzw. Lehrkräfte, die mit den Standards arbeiten sollen, brauchen also Unterstützung von außen, und zwar in einem beträchtlichen Umfang. Die Unterstützung kann und muss von den dafür vorgesehenen Einrichtungen, der Schulaufsicht, den Landesinstituten und den Einrichtungen der Lehrerbildung aller Phasen, bereitgestellt werden. Wenn man jedoch die momentanen Ausgangsbedingungen realistisch betrachtet, dann sind die genannten schulischen Supportsysteme derzeit noch nicht in der Lage, flächendeckend den Schulen die Unterstützung in der erforderlichen Qualität anbieten zu können. Bei der Einführung von Bildungsstandards wird man also auch darauf zu achten haben, dass die Unterstützungssysteme diesen neuen Aufgabenbereich auch wirklich übernehmen können. Dabei stellen sich sehr unter-

schiedliche Probleme, wenn die einzelnen Einrichtungen auf diese Aufgaben ausgerichtet und für die erforderlichen Unterstützungen qualifiziert werden sollen. Die Lage an Universitäten (und die Möglichkeiten der Beeinflussung) unterscheidet sich grundlegend etwa von der Situation der Schulaufsicht. Dennoch ist es unumgänglich, die Unterstützungssysteme auf diese neuen Anforderungen systematisch vorzubereiten und einzustellen. Ohne ihre qualifizierte Mitwirkung wird es kaum gelingen, Akzeptanz bei den Lehrkräften zu finden und entsprechende Qualifikationen auf breiter Basis zu erzeugen. Vorrangiges Ziel der Unterstützungssysteme ist es, in die Arbeit mit Bildungsstandards einzuführen und dazu beizutragen, dass die Standards effizient genutzt werden.

# 11. Die Infrastruktur für Entwicklung und Implementation von Bildungsstandards

In den vorangegangenen Kapiteln haben wir dargelegt,

- was Bildungsstandards sind (Kapitel 2, mit Beispielen in Kapitel 3 und Vertiefungen in den Kapiteln 5 bis 7), und
- welche Funktionen sie im Rahmen der pädagogischen Schulentwicklung und darüber hinaus für die Qualitätsentwicklung im Bildungssystem besitzen (Kapitel 4, mit Vertiefungen in den Kapiteln 8 bis 10).

In den beiden abschließenden Kapiteln der Expertise soll es nun darum gehen,

- wie Bildungsstandards entwickelt und in der Praxis eingeführt (implementiert) werden. Kapitel 11 fasst die Aufgaben, die in diesem Zusammenhang bewältigt werden müssen, systematisch zusammen und skizziert die Institutionen und Prozeduren, die hierfür benötigt werden. Kapitel 12 schließlich gibt einen Überblick über bisherige und laufende Arbeiten an Bildungsstandards in Deutschland und schlägt einen Zeit- und Arbeitsplan für die nächsten Jahre vor. Wo dies sinnvoll erscheint, werden in beiden Kapiteln Varianten zur Diskussion gestellt.

Im Zusammenhang mit der Einführung von Bildungsstandards kommen auf das Bildungssystem verschiedene Aufgaben zu. Eine erste Gruppe von Aufgaben betrifft die Festlegung von Kompetenzanforderungen und deren Akzeptanz sowie curriculare Umsetzung an Schulen, eine zweite Gruppe von Aufgaben betrifft Fragen der Operationalisierung von Standards (Testentwicklung) und der Nutzung von Tests für die Qualitätsentwicklung im Bildungswesen. Insgesamt geht es um die folgenden sechs Aufgabenbereiche:

- (1) *Entwicklung von Bildungsstandards*: Formulierung von bildungstheoretischen Vorstellungen, Kompetenzmodellen und (Minimal-) Anforderungen;
- (2) *Verbindliche Festsetzung von Bildungsstandards*: „Akkreditierung“ von Bildungsstandards und schulrechtlich verbindliche Einsetzung der Anforderungen;
- (3) *Implementation in Schulen und Unterstützungssystemen*: Umsetzung der Bildungsstandards in der Lehrplanarbeit (zentral und in den einzelnen Schulen); Konsequenzen für die pädagogische Arbeit an den Schulen, für Lehrerbildung, Schulaufsicht usw.;
- (4) *Testentwicklung*: Entwicklung und Zusammenstellung von Aufgaben, empirische Erprobung, Überprüfung von Kompetenzmodellen, Definition von Testskalen, Bereitstellung von Testinstrumenten für verschiedene Zwecke;

- (5) *Bildungsmonitoring*: Überprüfung der Einlösung von Bildungsstandards auf der Ebene des Bildungssystems, ggf. im Zusammenhang mit internationalen Schulleistungsvergleichen und eingebettet in eine nationale Bildungsberichterstattung; Untersuchung von Bedingungsfaktoren für schulisches Lernen;
- (6) *Evaluation von Schulen*: Rückmeldung über die Einlösung von Bildungsstandards an Schulen und ggf. an einzelne Klassen bzw. Lehrer, als Teil einer Untersuchung von Stärken und Schwächen der Schule im Blick auf ihre pädagogischen Aufgaben und Ziele, ihre Selbstüberprüfung und professionelle Weiterentwicklung.

Jeder dieser Aufgabenbereiche erfordert spezifische wissenschaftliche, administrative und pädagogische Qualifikationen, ggf. auch Zuständigkeiten im juristisch-administrativen Sinne. Keine Institution und kein Gremium kann diese Aufgaben in ihrer Gesamtheit bewältigen oder auch nur steuern; die Zusammenarbeit einer Vielzahl von Personen und Einrichtungen ist deshalb nötig. Im föderalen System der Bundesrepublik Deutschland ist auch zu prüfen, welche Aufgaben auf nationaler Ebene angegangen oder zumindest auf nationaler Ebene koordiniert werden sollten, und welche Aufgaben den einzelnen Ländern vorbehalten bleiben.

Im Folgenden gehen wir die Aufgabenstellungen durch und erläutern jeweils – mit möglichen Varianten – wie und von wem sie in einer zukünftigen Bildungslandschaft in Deutschland bewältigt werden können. Dabei lassen wir uns vom Prinzip „Output- statt Inputsteuerung“ leiten, das in der vorliegenden Expertise als Grundsatz der Steuerung staatlicher Schulsysteme empfohlen wird. Es kann analog auch auf die Entwicklung und Implementation der Bildungsstandards selbst angewandt werden. Dies bedeutet: Bildungsstandards und Testverfahren lassen sich vermutlich schneller, effizienter und mit höherer Akzeptanz entwickeln, wenn der Prozess nicht „deduktiv“ von oben durchgeplant wird, sondern wenn unterschiedliche Akteure mit einem hohen Grad an fachlicher Kompetenz und Autonomie „Produkte“ erstellen, die dann öffentlich diskutiert, verbessert und schließlich von zuständigen staatlichen Gremien – nach Maßstäben, die allerdings zentral festgesetzt werden müssen – zur regulären Nutzung freigegeben werden. Durch eine solche Vorgehensweise kann die Stärke eines föderalen Systems, die in der Vielfalt von Ansätzen und der Verfügbarkeit von Expertise an vielen Orten und in vielen Einrichtungen besteht, genutzt und mit der notwendigen Einheitlichkeit der Standards selbst verbunden werden.

## **Aufgabe 1: Entwicklung von Bildungsstandards**

Das Erarbeiten und Formulieren von Bildungsstandards ist nach unserer Konzeption eine Aufgabe, die primär fachdidaktische Expertise erfordert. 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. Sinnvoll ist es, Kompetenzen und Kompetenzstufen schon auf dieser Ebene durch Beispielaufgaben zu illustrieren. Schließlich muss anhand fachdidaktischer Befunde und Erfahrungen Einigkeit darüber erzielt werden, welche Kompetenzanforderungen verbindlich gemacht werden können. All dies erfordert einschlägige fachdidaktische, zum Teil auch fachliche und schulpraktische oder allgemeinpädagogische Expertise.

Innerhalb der Kultusministerkonferenz (KMK) sind für diese Art von Tätigkeiten in jüngster Zeit zwei Modelle entwickelt worden: Für die Entwicklung von Standards im Primar- und Sekundarstufen I - Bereich wurden länderübergreifende Arbeitsgruppen gebildet, denen im Wesentlichen Lehrplanexperten aus Ministerien und pädagogischen Landesinstituten angehören; diese werden von Fachdidaktikern beraten (vgl. Kapitel 12). Für die Erarbeitung von Kerncurricula der gymnasialen Oberstufe wurden hingegen „externe“ Aufträge an kleine Gruppen von Fachdidaktikern und Fachwissenschaftlern vergeben, die Bestimmungselemente von Kerncurricula in Form von Expertisen vorlegen (Tenorth 2001).

Aus einer internationalen Perspektive ist interessant anzumerken, dass etliche sehr einflussreiche Dokumente wie etwa die Mathematik-Standards des NCTM und die Empfehlungen der American Association for the Advancement of Science für naturwissenschaftliche Bildung nicht von staatlichen Gremien, sondern von Fachverbänden erarbeitet und erst danach staatlich akzeptiert wurden. Viele Staaten – wie etwa die Niederlande, Großbritannien und Schweden – haben in den vergangenen Jahren Aufgaben der Steuerung des Bildungswesens an staatlich budgetierte und beaufsichtigte, aber rechtlich selbständige Institutionen delegiert.

Mittelfristig wird die Entwicklung von nationalen Bildungsstandards auch in Deutschland einer Form der Institutionalisierung bedürfen. Die Koordination von Entwicklungs- und ggf. Revisionsarbeiten, die konzeptionelle Fundierung, die Einbindung wissenschaftlicher und insbesondere fachdidaktischer Expertise sind auf Dauer durch *ad hoc* gebildete Gremien kaum zu bewältigen. Dies bedeutet allerdings nicht, dass eine eigene Bürokratie für die Entwicklung von Bildungsstandards entstehen sollte. Sinnvoll erscheint ein rechtlich eigenständiges, wissenschaftlich qualifiziertes Institut, das die Arbeiten konzipiert und koordiniert, aber die „operativen Aufgaben“ in der

Regel an Dritte vergibt. So können – wie beim Kerncurriculum Oberstufe – Expertengruppen beauftragt werden, Bildungskonzepte und Kompetenzmodelle für einzelne Lernbereiche und Fächer zu konzipieren.

Diese grundlegende Arbeit wird sinnvollerweise über Schulstufen und -formen hinweg organisiert. Beispielsweise könnte die Arbeit an Bildungsstandards für Naturwissenschaften von einem Kompetenzmodell profitieren, das die Entwicklung naturwissenschaftlichen Denkens und die Stufen entsprechender Tätigkeiten über die gesamte Bildungslaufbahn hinweg abbildet (vgl. Anhang b). Auf dieser Grundlage können dann Arbeitsgruppen unter Beteiligung von Praktikern Kompetenzanforderungen für bestimmte Jahrgangsstufen erstellen. Nach dem Vorbild etwa der NCTM-Standards wäre anzustreben, solche Kompetenzanforderungen für jede zweite oder dritte Jahrgangsstufe anzulegen. Damit würde sowohl für die Schnittstellen des Bildungsverlaufs (z.B. Ende der Grundschulzeit) und die Abschlussjahrgänge der Sekundarschule als auch für Jahrgänge innerhalb der Bildungsgänge eine Orientierung gegeben. Ähnliches könnte auf der Basis Gemeinsamen Europäischen Referenzrahmen (GER, vgl. Anhang a) für den Bereich der Sprachen insgesamt, zumindest aber für den der Fremdsprachen erfolgen.

## **Aufgabe 2: Verbindliche Festsetzung von Bildungsstandards**

Die Festlegung von Bildungsstandards und die Überprüfung ihrer Anwendung sind essenzieller Bestandteil der staatlichen Schulaufsicht (Art. 7 Abs. 1 GG). „Staat“ im Sinne des Art. 7 Abs. 1 GG ist im Bereich des Schulwesens nach der bundesstaatlichen Zuständigkeitsverteilung (Art. 30, 70 ff. GG) das jeweilige Land. Wie beim Erlass von Lehrplänen ist bei Bildungsstandards – für die sich die Rechtsform der Verwaltungsvorschrift oder auch der Rechtsverordnung anbietet – die Exekutive aufgrund ihrer größeren Sachnähe und Flexibilität zuständig. Ihr fällt auch die Verantwortung für die Überprüfung der Einhaltung der Standards zu.

Die Etablierung von Bildungsstandards als verbindliche Vorgaben für die Arbeit der Schulen ist daher Aufgabe des jeweiligen Landesministeriums. Ein gangbarer und etwa bei den *Einheitlichen Prüfungsanforderungen* für das Abitur erprobter Weg besteht darin, dass die zuständigen KMK-Gremien sich – basierend auf Vorlagen, die von Expertengruppen und ggf. einer nationalen Agentur stammen – auf gemeinsame Dokumente einigen, die dann durch gleich gerichtetes Handeln der Ministerien rechtlich verbindlichen Status erhalten.

Nur wenn die Festlegung und Überprüfung von nationalen Bildungsstandards ein hoheitlicher Akt wäre, müsste dafür eigens eine gesetzliche Grundlage geschaffen werden. Im Verständnis dieser Expertise handelt es sich aber bei den Standards um nicht mehr und nicht weniger als ein Instrument zur Orientierung und Ergebnismeldung für das staatliche Schulwesen, auf das sich die Kultusminister einigen. Bildungsstandards sollen gerade nicht die Funktion erhalten, individuelle Bildungsverläufe zu lenken oder gar Noten und Zertifikate zu begründen.

### **Aufgabe 3: Implementation in Schulen und Unterstützungssystemen**

Die Einführung von Bildungsstandards hat Konsequenzen für die Lehrplanarbeit (vgl. Kapitel 8), die Lehrerbildung – insbesondere die Lehrerfortbildung –, für die Schulaufsicht und andere Unterstützungssysteme (vgl. Kapitel 10). Hierfür sind keine zusätzlichen Institutionen erforderlich, aber die Arbeit der bestehenden Einrichtungen, insbesondere der pädagogischen Landesinstitute, wird sich verändern.

Ein zentraler Grundsatz der vorliegenden Expertise ist es, Schulen verstärkte Eigenständigkeit bei der pädagogischen Arbeit zu geben. Bildungsstandards legen klare und verständliche Maßstäbe für das Lernergebnis fest; sie können und sollen gerade deshalb den Schulen bei der Gestaltung der Wege zur Erreichung der Standards Freiräume lassen. Konsequenterweise weitergedacht, bedeutet dies, dass die Feingliederung des Curriculums, die Abstimmung von inhaltlichen und methodischen Details und die zeitliche Anordnung vor Ort in einem Schulcurriculum festgelegt werden sollten. Allerdings darf nicht verkannt werden, dass eine solche Verlagerung curricularer Entscheidungen zusätzliche Qualifikationen und Ressourcen an den Schulen erfordern. Den Landesinstituten, die derzeit zentrale Lehrpläne und Rahmenrichtlinien erstellen, käme daher verstärkt eine beratende Funktion für die curriculare Arbeit der Schulen zu. Ein solcher grundsätzlicher Strategiewechsel ist nur schrittweise zu realisieren. Daher wurde in Kapitel 8 vorgeschlagen, die Lehrpläne auf Landesebene zunächst beizubehalten, aber stärker zu fokussieren, auf Kompetenzmodelle zu beziehen und somit im Sinne von Kerncurricula zu gestalten.

Verändern wird sich auch die Rolle der Schulaufsicht. Sie nimmt gegenwärtig in Deutschland sowohl kontrollierende als auch beratende Funktionen wahr. Mehrere europäische Staaten sind in jüngerer Zeit dazu übergegangen, für die eigentliche Überprüfung der Schulen („Inspektion“) und für die Unterstützung bei der pädagogischen Entwicklung und der Bewältigung von Problemen („Schulentwicklungsberatung“) unterschiedliche Institutionen zu schaffen, um so letztlich beide

Funktionen zu optimieren. In eine ähnliche Richtung weisen auch Überlegungen zur Neuorientierung der Schulaufsicht in einigen Ländern der Bundesrepublik.

#### **Aufgabe 4: Operationalisierung von Kompetenzmodellen und Testentwicklung**

„Musteraufgaben“, die Kompetenzdimensionen und -stufen exemplarisch illustrieren, werden sinnvollerweise zusammen mit den Standards vorgelegt (siehe Aufgabe 1). Die Konstruktion eines größeren „Pools“ von Aufgaben, die zu den inhaltlichen und kognitiven Anforderungen der Kompetenzstufen passen und sie in einer Vielzahl von Kontexten variieren, erfordert jedoch ein sehr systematisches Vorgehen. Traditionelle Leistungstests wurden zumeist erstellt, indem eine große Menge von Aufgaben eher unsystematisch gesammelt und erprobt wurde. Solche Tests lassen allerdings in der Regel nur eine normorientierte Interpretation zu, keine kriteriumsorientierte (auf die inhaltlichen und kognitiven Anforderungen der Kompetenzstufen ausgerichtete) Interpretation, wie sie hier befürwortet wird. Aussagefähige Tests, die tatsächlich das Kompetenzmodell der Bildungsstandards operationalisieren, bedürfen also einer gezielten Entwicklungsarbeit. Die Sammlung von Aufgaben, wie sie beispielsweise in Vergleichsarbeiten einzelner Länder verwendet werden, kann als Startpunkt geeignet sein, ersetzt aber nicht die gezielte Konstruktion von Aufgaben entlang der identifizierten Kompetenzmodelle. Die Operationalisierung nationaler Bildungsstandards muss in der Tat als eine nationale Aufgabe verstanden werden.

Aufgabenentwürfe werden informell in einzelnen Klassen vorerprobt – z.B. im Hinblick auf Verständlichkeit, Klarheit der Instruktion und der Auswertungsregeln – und dann einer systematischen Erprobungsstudie („Pilotierung“) unterzogen, um ihre Messqualität abzusichern. Eine solche Pilotierung muss nicht an repräsentativen Schülergruppen erfolgen, aber die Teilnehmenden müssen über Schulformen und Leistungsniveaus gestreut sein, und jede einzelne Aufgabe sollte von mehreren Hundert Schülerinnen und Schülern bearbeitet werden. Ergebnis der Pilotierung ist (a) eine empirische Überprüfung und ggf. Modifikation des Kompetenzmodells, (b) die Möglichkeit, Kompetenzstufen mit Beispielen zu illustrieren, die nicht einfach am Schreibtisch konstruiert, sondern empirisch validiert wurden, und (c) ein Testinstrument, das anschließend für Bildungsmonitoring, Evaluation usw. genutzt werden kann.

Es ist offensichtlich, dass sowohl die Konstruktion von Aufgaben als auch die empirische Erprobung spezielle wissenschaftliche Qualifikationen erfordern. Andere Staaten haben daher für entsprechende Tätigkeiten seit Jahren spezielle Institutionen eingerichtet, die privatwirtschaftlich organisiert sind (Educational Testing Service, USA, Cito, Niederlande), staatlich budgetierte Institu-

tionen darstellen (Qualifications and curriculum authority, England), als Abteilungen der nationalen Bildungsbehörden firmieren (Schweden, Finnland) oder unmittelbar im Bildungsministerium angesiedelt sind (Frankreich). In Deutschland gibt es keine vergleichbare Institution; auch die pädagogischen Landesinstitute sind auf diese Aufgaben bislang nicht oder nur unzulänglich vorbereitet. Allerdings haben in den vergangenen Jahren eine Reihe von Hochschulen und außeruniversitären Forschungsinstituten, die im Kontext von internationalen Vergleichsstudien und Länderstudien tätig waren, das einschlägige know-how entwickelt, und eine wachsende Zahl von Fachdidaktikern befasst sich mit Fragen der Aufgabenerstellung und Testentwicklung. Die „technischen Arbeiten“ (Stichprobenziehung, Erstellen von Testunterlagen, Schulung des zuständigen Personals, Datenmanagement und Kodierung von Antworten) werden teils von diesen Arbeitsgruppen übernommen, teils von spezialisierten Dienstleistern, unter anderem dem Data Processing Center der International Association for the Evaluation of Educational Achievement (IEA) in Hamburg, das neben seinen internationalen Aufgaben zunehmend auch bei nationalen Projekten beteiligt ist. Diese Vielfalt der Institutionen, die Aufgaben der Testentwicklung übernehmen können, sollte in der Zukunft erhalten bleiben und nach Möglichkeit ausgebaut werden.

Die Autoren dieser Expertise halten es daher für sinnvoll, für die Arbeiten zur Operationalisierung von Bildungsstandards auf nationaler Ebene eine koordinierende Institution aufzubauen, die Aufträge jeweils an fachlich und methodisch qualifizierte Arbeitsgruppen von Bildungsforschern und Didaktikern vergibt, aber für die Einhaltung von Qualitätsmaßstäben und die konzeptuelle Verbindung verschiedener Projekte sorgt und soweit notwendig die Verwendung der Tests kontrolliert. Diese „Testagentur“ sollte wissenschaftlich qualifiziert arbeiten, bedarf also der Anbindung an eine Hochschule. Sie muss über Aufsichtsgremien verfügen, die aus der Wissenschaft und der Bildungsadministration besetzt werden.

## **Aufgabe 5: Bildungsmonitoring**

Aussagen über das in den Ländern erreichte Kompetenzniveau der Schülerinnen und Schüler sind eine wichtige Grundlage für bildungspolitische Steuerungsmaßnahmen. Regelmäßige repräsentative Erhebungen mit standard-bezogenen Tests bilden die Basis dieser *output*-orientierten Steuerung.

Im Bereich der Sekundarstufe I beteiligt sich Deutschland bereits seit dem Jahr 2000 an den PISA-Erhebungen, die alle drei Jahre – mit wechselnden Schwerpunkten – bei 15-jährigen bzw. in der Jahrgangsstufe 9 stattfinden. Für die Erhebungen in den Jahren 2000 und 2003 wurde das

Design in Deutschland erweitert, so dass Ländervergleiche möglich sind. Die OECD plant, die internationalen Erhebungen auch über 2009 hinaus weiterzuführen. Die Aussagekraft des Bildungsmonitoring für Deutschland würde wesentlich erweitert, wenn zukünftige nationale, standard-bezogenen Tests gemeinsam mit den PISA-Tests bzw. in einem koordinierten Untersuchungsdesign eingesetzt würden. Beispielsweise könnte damit die in Deutschland geführte Debatte um den Bezug zwischen curricularen Standards und literacy-Konzepten empirisch fundiert werden. Zugleich ließen sich differenzierte Erklärungsmodelle der PISA-Studie auf die neuen Kompetenzbereiche erweitern. Diese Verknüpfung mit PISA ist allerdings nur möglich, wenn Standards für die Jahrgangsstufe 9 ausgewiesen werden.

Zu einem voll 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 repräsentative Erhebung mit jeweils neuen Stichproben durchzuführen. Länder wie Kanada und die USA führen ihre nationalen Erhebungen zwar jährlich durch, wechseln aber im festen Rhythmus die Lernbereiche (Fächer). Häufigere *large scale assessments* oder gar flächendeckende jährliche Untersuchungen an allen Schulen, wie sie etwa in einigen Bundesstaaten der USA üblich sind, erscheinen nach Ansicht der Expertengruppe unnötig und letztlich kontraproduktiv, weil der Leistungskontrolle ein zu großes Gewicht gegeben wird.

Ein Bildungsmonitoring muss über Kompetenzverteilungen hinaus auch Ressourcen und Prozesse (z.B. Bildungsverläufe) darstellen und Bedingungsfaktoren für erfolgreiches schulisches Lernen analytisch herausarbeiten. Dies ist eine Aufgabe, die hohe Qualifikationen in der empirischen Bildungsforschung voraussetzt. Nicht zuletzt aufgrund der Teilnahme an internationalen Schulleistungsstudien hat die Schulforschung in Deutschland wieder Anschluss an die internationale Entwicklung gefunden. Auch weiterhin sollten entsprechend erfahrene Institute mit Monitoring-Studien betraut werden. Die Verstetigung des PISA-Prozesses sowie die Einbeziehung zusätzlicher Fächer und Jahrgänge werden allerdings allein aus Effizienzgründen erfordern, dass ein wissenschaftliches Kompetenzzentrum geschaffen wird, das diese Studien zentral koordiniert, Zusammenhänge in Konzeption und Auswertung herstellt (z.B. durch Einfügung längsschnittlicher Komponenten, bei denen dieselben Schülerinnen und Schüler über mehrere Erhebungen hinweg beobachtet werden) und für Einhaltung von Qualitätsmaßstäben sorgt.

## **Aufgabe 6: Evaluation von Schulen**

Schulevaluation ist in Deutschland mittlerweile in der einen oder anderen Form etabliert, allerdings eher in einem qualitativen Sinne. Mehrere Länder haben den Schulen die Erstellung von Schulprogrammen und eine regelmäßige Bestandsaufnahme des im Verhältnis zu diesem Programm Erreichten zur Pflicht gemacht. Im Rahmen von Schulversuchen und Modellprogrammen ist Evaluation mittlerweile eine Selbstverständlichkeit geworden. Auf dem freien Markt sind Dienstleistungsangebote für Schulevaluation ebenso zu haben wie Anleitungen und Materialien zur Selbstevaluation.

Mit nationalen Bildungsstandards werden in diese Praxis keine neuen Strukturen oder Zuständigkeiten, wohl aber neue Kriterien eingebracht. Soll die tatsächliche Erreichung von Standards überprüft werden, ist die einzelne Schule, wie in Kapitel 7 ausgeführt, allerdings überfordert. Sie muss Zugang zu professionell entwickelten, standard-bezogenen Tests erhalten. Hier sind unterschiedliche Wege denkbar:

*Eine* Möglichkeit besteht darin, dass Landesinstitute die standard-bezogenen Tests, nachdem sie im Bildungsmonitoring eingesetzt worden sind, „verwalten“ und für Programme der Schulevaluation nutzen. Eine zweite Möglichkeit besteht darin, Tests im Anschluss an ein Bildungsmonitoring freizugeben und daraus Materialien zu entwickeln, die Schulen selbständig und auf eigene Initiative einsetzen können. So ist man beispielsweise in Brandenburg mit Aufgaben des landesweiten Mathematiktests QUASUM umgegangen. Allerdings sind dieser „Selbsttestung“ aus methodischen und praktischen Gründen Grenzen gesetzt. Beispielsweise ist es so kaum möglich, adjustierte Ergebnisse zu ermitteln, die Eingangsbedingungen der Schülerinnen und Schüler berücksichtigen (vgl. Kapitel 9). Prinzipiell ist es auch denkbar, dass entsprechende Testinstrumente zur Evaluation der Schulen einschließlich der Selbstevaluation kommerziell entwickelt und angeboten werden.

Bislang spielen quantifizierte Leistungsmessungen innerhalb der Schulevaluation in Deutschland praktisch keine Rolle – unter anderem, weil es an Instrumenten fehlte. Daher muss die Erhebung und der Umgang mit solchen Daten sehr sorgfältig geplant werden. Kriteriumsorientierte Informationen sind hilfreicher als normorientierte „Rankings“, und die Grenzen der Aussagekraft müssen berücksichtigt werden. Fatal wäre es, wenn Bildungsstandards durch unprofessionelle Evaluation diskreditiert würden. Die Länder sind gefordert, Verfahren der Schulevaluation und entsprechende „Angebote“ so zu gestalten, dass professionelle Qualitätsmaßstäbe eingehalten werden und ein qualifizierter Umgang mit Standards und Tests garantiert wird. Es könnte auch zu den Aufgaben einer nationalen „Testagentur“ gehören, Tests und Evaluationsverfahren, soweit sie sich auf

Bildungsstandards beziehen, zu zertifizieren und sie damit einer einschlägigen Qualitätskontrolle im Sinne dieser Expertise zu unterziehen.

## **Fazit**

Für die verbindliche *Festsetzung* von Bildungsstandards, ihre Umsetzung in Schulen und Unterstützungssystemen sowie für die Evaluation von Einzelschulen (Aufgabenbereiche 2, 3, 6) werden die Länder weiterhin einzeln Verantwortung übernehmen müssen, auch wenn sie möglicherweise unterschiedliche Strategien verfolgen. Eine zentrale Rolle werden hierbei die pädagogischen Landesinstitute zu übernehmen haben. Die Länder tauschen sich im Rahmen der KMK aus und können das Instrumentarium der Bund-Länder-Kommission für Bildungsplanung und Forschungsförderung (BLK) nutzen, um wechselseitig aus Erfahrungen zu lernen, koordinierte Modellprogramme durchzuführen und dadurch auch Ressourcen besser zu nutzen.

Für die *Entwicklung* und Revision von Standards, eine darauf aufbauende Testentwicklung und das nationale Bildungsmonitoring (Aufgabenbereiche 1, 4 und 5) werden jedoch Strukturen auf nationaler Ebene benötigt. Die Expertengruppe schlägt vor, für möglichst viele dieser Aufgaben befristete Aufträge an wissenschaftliche Institute, Hochschulen oder Konsortien von Wissenschaftlern zu vergeben. Ein solches Verfahren würde Flexibilität und föderale Vielfalt sichern und mannigfache Impulse für eine Weiterentwicklung der empirischen Bildungsforschung geben, die u.a. von der Deutschen Forschungsgemeinschaft für notwendig gehalten wird. Die konzeptuelle Planung, Koordination, Vernetzung und Qualitätskontrolle solcher Aufträge kann jedoch auf Dauer nicht quasi nebenbei von Gremien der Bildungsadministration geleistet werden. Es bedarf einer kontinuierlich arbeitenden, koordinierenden Instanz, die selbst wissenschaftlich qualifiziert, forschend tätig und an eine Hochschule angebunden ist. Es ist daher zu begrüßen, dass Bildungspolitiker derzeit an Konzeptionen für eine entsprechende nationale Agentur arbeiten. Diese Agentur würde im Bereich der Forschung und der Bildungsplanung tätig und könnte daher nach Artikel 91b des Grundgesetzes von Bund und Ländern gemeinsam etabliert und finanziert werden. Zu überlegen wäre auch, ob die drei hier genannten Aufgaben in *einer* Institution zusammengefasst oder auf verschiedene Einrichtungen verteilt werden sollen. Für eine Trennung spricht, dass die Entwicklung von Standards und Tests eher fachdidaktische und testmethodische Qualifikationen erfordert, das Bildungsmonitoring hingegen Qualifikationen im Bereich der empirischen Schulforschung und der Bildungsstatistik. Effizienzgesichtspunkte sprechen andererseits für eine Zusammenfassung der Funktionen.

Eine wesentliche strukturelle Voraussetzung für Bildungsreformen ist die Stärkung der Forschung und die Förderung des wissenschaftlichen Nachwuchses, der für die hier beschriebenen Aufgaben gebraucht wird. Insbesondere die Fachdidaktiken und die empirische, interdisziplinär angelegte Bildungsforschung bedürfen des weiteren Ausbaus.

Dies alles ist nicht zum Null-Tarif zu haben. Allerdings darf man nicht verkennen, dass national eingeführte Bildungsstandards den Aufwand für Lehrplanarbeit innerhalb der Länder erheblich verringern werden: es müssen nicht 16 Länder, in der Regel mit je vier Schularten, z.B. die Grundkonzepte eines Faches und die Kompetenzmodelle mehrfach parallel entwickeln. Kosten, die bislang im Verwaltungsaufwand der Bildungsadministration verborgen waren, würden durch neue Institute und die Vergabe von Aufträgen transparenter. Vor allem aber ist zu vermuten, dass im Vergleich zur bisherigen curricularen Arbeit (vgl. etwa Vollstädt & Tillmann 1998, Biel, Ohlha-ver & Riquarts 1999) die Entwicklung von Bildungsstandards für das Bildungssystem als Ganzes effizienter ist.

## 12. Stand und Perspektiven der Entwicklung von Bildungsstandards in Deutschland

Die Konferenz der Kultusminister und -senatoren der Länder (KMK) hat sich selbst mit ihrem Beschluss zur Einführung von Bildungsstandards vom Juni 2002 ehrgeizige Ziele gesetzt. Im Herbst 2003 sollen der Öffentlichkeit erste Arbeitsergebnisse vorgestellt werden. Bis zum Frühjahr 2004 sollen dann folgende Dokumente vorliegen:

- Standards für die Grundschule (Ende der 4. Jahrgangsstufe) in den Fächern Deutsch und Mathematik,
- Standards für den Hauptschulabschluss in Deutsch, Mathematik und der Fremdsprache,
- neu gefasste Standards für den Mittleren Schulabschluss (Abschluss der Realschule und vergleichbarer Bildungsgänge) in Deutsch, Mathematik und der Fremdsprache, sowie ergänzend
- Überarbeitungen der „Einheitlichen Prüfungsanforderungen“ (EPA) für verschiedene Abiturfächer.

Die Standards sollen länderübergreifend verbindlich festlegen, welche Lernergebnisse am Ende einer bestimmten Jahrgangsstufe (hier: der Jahrgangsstufen 4, 9, 10) erwartet werden. Sie sollen zudem die Basis für Orientierungs- bzw. Vergleichsarbeiten bilden, die länderintern das Erreichen von Standards überprüfen.

Bei der Entwicklung dieser Dokumente kann die KMK auf eine Reihe von Vorarbeiten zurückgreifen, die in den letzten Jahren bereits in einzelnen Ländern oder durch Gruppen von Ländern geleistet wurden. Im Folgenden geben wir einen Überblick über diese Entwicklungen, diskutieren Übereinstimmungen mit und Abweichungen von den Empfehlungen dieser Expertise und skizzieren, welche Schritte in den nächsten Jahren zur Weiterentwicklung von Bildungsstandards notwendig sind.

### 12.1 Vorarbeiten in den Ländern

Viele Lehrpläne der Länder definieren bis heute Bildungs- bzw. Lernziele durch Auflistung einzelner Inhalte, die im Fachunterricht in einem bestimmten Schuljahr gelernt werden sollen. Die Einbindung in einen größeren Rahmen von Unterricht erfolgt eher selten, sieht man von einigen we-

nigen fächerübergreifenden Bezügen ab. Insbesondere werden die Aufgaben des Fachunterrichts in den verschiedenen Klassen- und Schulstufen kaum durch einheitliche und übergreifende Ziele beschrieben, zumeist wird auf eine Einordnung in Lerntheorien oder didaktische Prinzipien verzichtet. Das so festgelegte Curriculum gibt einen Rahmen vor, der im Grunde nur durch ein entsprechend angepasstes Schulbuch für die Lehrerinnen und Lehrer praktikabel wird.

Mehrere Länder haben in jüngster Zeit jedoch neue Lehrpläne oder Rahmenrichtlinien entwickelt, mit denen pädagogische und fachdidaktische Konzepte aufgegriffen werden, die seit TIMSS und PISA verstärkt öffentlich diskutiert werden. Diese neueren Lehrpläne sind schon stärker auf zentrale Aufgaben der Fächer konzentriert; sie definieren verbindliche Kernbereiche und lassen gleichzeitig Freiraum für die einzelne Schule. Häufig werden in der Einleitung zum Lehrplan oder in seiner Aufgliederung Kompetenzen und fachliche Leitideen angesprochen, wie sie auch in dieser Expertise als Grundlage für Bildungsstandards empfohlen werden. Mathematik-Curricula beispielweise verweisen zumindest in ihren Einleitungsabschnitten häufig auf Unterrichtsprinzipien, wie sie auch in den „Standards and Principles“ der amerikanischen NCTM genannt werden. Sie dürfen allerdings nicht (auch wenn die Autoren solcher neuen Lehrpläne teilweise die Begrifflichkeit verwenden) mit Kompetenzen bzw. Kompetenzmodellen verwechselt werden, wie sie in TIMSS und PISA beschrieben sind (vgl. Abschnitt 6.2; zur Diskussion siehe Abschnitt 12.2).

Einige Länder haben sich sogar bereits auf den Weg weg von klassischen Lehrplänen und Rahmenrichtlinien hin zu expliziten Bildungsstandards gemacht. Das Konzept wird durchweg so verstanden, dass es gilt, über die Ziel- und Inhaltsformulierungen sowie über Leitideen der Lehrpläne hinaus zu bestimmen, welche Leistungen man von Schülerinnen und Schülern an bestimmten Punkten ihrer Schullaufbahn erwartet. Hierzu wurden von den Ländern unterschiedliche Wege beschritten:

- In einzelnen Ländern haben Lehrplanexperten als Vorarbeit für Standards versucht, aus den bestehenden Curricula Leistungserwartungen für bestimmte Jahrgänge oder Abschlüsse abzuleiten.
- Ein zweiter Ansatz besteht darin, Rahmenlehrpläne neu zu schreiben und darin konkrete Leistungserwartungen zu definieren. Standards sind hier also unmittelbar Teil des Curriculums. Die neuen Pläne für die Sekundarstufe I in Brandenburg beispielweise enthalten „Qualifikationserwartungen“, die als „Indikatoren für den Lernerfolg“ und zugleich als „Qualitätsstandards von Unterricht“ verwendet und auf drei Stufen (grundlegende, erweiterte und vertiefte allgemeine Bildung) ausgeführt werden.
- Ein dritter Weg besteht darin, neben den bestehenden Curricula Bildungsstandards zu entwickeln. Dieser Weg wurde z.B. schon vor dem KMK-Beschluss vom Juni 2002 in Baden-Württemberg gewählt. Dort hat man sich auf folgende Arbeitsdefinition verständigt: „Bildungs-

standards legen konkret fest, welches verbindliche Wissen und welche Kompetenzen Schülerinnen und Schüler zu einem bestimmten Zeitpunkt zur Verfügung haben müssen. ... Bildungsstandards enthalten Leitgedanken zu einem speziellen Fach oder zu einer Fächergruppe, Zielformulierungen und Inhalte sowie Musteraufgaben zur Evaluation von Schülerleistung und Unterrichtsqualität. Sie beschreiben verbindliche Zielvorgaben der fachlichen Fertigkeiten und Fähigkeiten sowie der überfachlichen – also methodischen, sozialen und personalen – Kompetenzen.“

Gemeinsame Entwürfe für Standards im Grundschul- und Sekundarstufenbereich haben zudem die CDU-regierten Länder auf der Arbeitsebene vorgelegt.

- Ein vierter Weg besteht darin, unmittelbar vom Lehrplan zur Entwicklung von Testverfahren überzugehen. So hat Rheinland-Pfalz das Projekt VERA („Vergleichsarbeiten“) gestartet, das einen Aufgabenpool für die Schulevaluation am Ende der Grundschulzeit entwickeln soll (Helmke & Hosenfeld 2003); an dem Vorhaben sind inzwischen eine Reihe weiterer Länder beteiligt. Hier werden keine inhalts- oder kompetenzbezogenen Standards vorab definiert, sondern es werden unter Beteiligung vieler erfahrener Lehrer Aufgaben entworfen und nach psychometrischen Verfahren Tests zusammengestellt, die dann landesweit normiert werden sollen. Ähnlich ist man in Bayern vorgegangen, wo bereits jetzt landesweit einheitliche Orientierungsarbeiten in den Jahrgangsstufen 2 und 3 geschrieben werden, die sich an neuen Lehrplänen ausrichten.

In der englischsprachigen Terminologie entwickelt man also hier *performance standards* ohne vorherige – über die Lehrpläne hinausgehende – *content standards*. Dieser pragmatische Weg erleichtert vermutlich einen schnellen Start für systematische Schulevaluation und Bildungsmonitoring, bietet den Lehrern aber möglicherweise eingeschränktere Orientierungsmöglichkeiten, weil Kompetenzmodelle, fokussierte Beschreibungen des Kerncurriculums und qualitative Leistungserwartungen fehlen, die nach dem hier vorgestellten Konzept zu Bildungsstandards gehören. Allerdings ist sowohl in Rheinland-Pfalz als auch in Bayern beabsichtigt, diese Tests und Aufgabensammlungen an die bundesweit erarbeiteten Standards und die Ergebnisse der internationalen Grundschuluntersuchung IGLU anzubinden, sobald diese vorliegen.

Viele der in den Ländern erarbeiteten Dokumente zeigen noch eine sehr große Nähe zu herkömmlichen Lehrplänen. Im Anschluss an allgemeine und fachspezifische Präambeln listen sie Zielformulierungen und Inhalte auf, die jeweils schulform- bzw. bildungsgangspezifisch ausgewählt und häufig im Stil klassischer Lernzieltaxonomien gefasst sind. Neu ist an diesen Entwürfen vor allem, dass sie einen verbindlichen Kern beschreiben und durch „Musteraufgaben“, Orientierungsarbeiten usw. ergänzt werden, welche die Anforderungen konkretisieren und langfristig die Basis für Evaluation von Schule und Unterricht bilden sollen.

## 12.2 Das Verständnis von Kompetenzen als entscheidende Herausforderung

Bei näherer Betrachtung der neueren Lehrpläne, Richtlinien und ersten Entwürfe für Standards wird im Vergleich zu den Empfehlungen dieser Expertise vor allem deutlich, dass in der Praxis recht unterschiedlich und unsicher mit dem Begriff „Kompetenz“ und mit Kompetenzmodellen umgegangen wird.

Die oben zitierte baden-württembergische Definition von Standards deutet an, dass man den Begriff „Kompetenzen“ eher für die sogenannten Schlüsselqualifikationen verwendet, während man im Kontext eines Faches von Wissen, Fähigkeiten und Fertigkeiten spricht. Dies gilt ausdrücklich für die neuen Rahmenpläne in Mecklenburg-Vorpommern und die Rahmenlehrpläne in Brandenburg, die sich an einem aus der Berufsbildung übernommenen Konzept von „Handlungskompetenz“ orientieren. Für jedes Fach wird hier aufgeführt, welchen Beitrag es zur Entwicklung von Sachkompetenz, Methodenkompetenz, Sozialkompetenz und Selbstkompetenz leistet. Als „Kompetenzen“ werden hier also sogenannte Schlüsselqualifikationen genannt, die gleichsam als allgemeine, überfachliche Bildungsziele verstanden werden, zu denen jedes Fach beitragen soll.

Die Expertengruppe geht hingegen 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. Diese Inhalts- und Erfahrungsbereiche sind im Schulsystem von heute vor allem durch Fächer repräsentiert. Nach diesem Verständnis haben Kompetenzmodelle die Aufgaben, die Ziele, die Struktur und die Ergebnisse fachlicher Lernprozesse zu beschreiben. Sie bilden die Komponenten und Stufen der Kompetenzentwicklung von Schülerinnen und Schülern ab und bieten somit eine Orientierung für schulisches Lehren und Lernen. Kompetenzen werden in diesem Sinne als Verbindung von Inhalten einerseits und Operationen oder „Tätigkeiten“ an bzw. mit diesen Inhalten andererseits verstanden. Umgesetzt in Aufgaben und Tests erlauben es solche Kompetenzmodelle, den Leistungsstand von Schülerinnen und Schülern kriteriumsorientiert zu beschreiben, und das nicht nur im Vergleich mit dem Leistungsstand anderer, sondern bezogen auf das definierte Kriterium. Eine solche kriteriumsorientierte Kompetenzbeschreibung benennt also konkrete Anforderungen, die auf einer bestimmten Kompetenzstufe bewältigt werden können, und beschreibt zugleich jene Tätigkeiten und Wissensinhalte, die noch nicht beherrscht oder erreicht sind.

Erst diese Kompetenzmodelle geben den Standards eine Orientierungskraft für den Unterricht, indem sie unmittelbar einsichtig und nachvollziehbar, illustriert an konkreten Anforderungen, demonstrieren, welche Entwicklungs- und Niveaustufen fachliche Kompetenzen haben. Solche

Kompetenzmodelle fordern auch dazu heraus, Veränderungen im Verlauf des Lernens bei Schülerinnen und Schülern differenziert wahrzunehmen. Auch für Leistungsschwächere wird festgehalten, was sie bereits wissen und können, und es wird sichtbar, in welcher Richtung und auf welchem Weg sie gefördert werden können. Damit wird ein Gegengewicht geschaffen zu dem allzu üblichen Denken in Defizit-Modellen und sozialen Vergleichen. Werden Tests oder Vergleichsarbeiten geschrieben, ermöglichen Kompetenzmodelle eine kriteriumsorientierte Rückmeldung an die Schulen, die konkret benennt, was ihre Schülerschaft kann und welche Leistungen (als Indikatoren für bestimmte Kompetenzstufen) sie noch nicht zu erbringen vermag.

In die selbe Richtung argumentiert auch eine andere Expertise, die führende Erziehungswissenschaftler und Psychologen der USA für den National Research Council erarbeitet haben (NRC 2001). Unter dem Titel „Knowing what students know – The science and design of educational assessment“ gehen die Autoren hart mit den in den USA bislang üblichen Standards und Leistungstests ins Gericht, weil diese lediglich Lernziele und Testmarken festlegen, aber die Kompetenzentwicklung nicht darstellen können (Man beachte, dass sich diese Kritik nicht auf neue, fachdidaktisch fundierte Standards wie etwa die in Kapitel 3 beschriebenen NCTM-Standards bezieht, sondern auf traditionelle Lernziellisten und *performance standards*, die in den USA immer noch den Regelfall ausmachen!).

„While the existing standards emphasize what students should learn, they do not describe how students learn in ways that are maximally useful for guiding instruction and assessment“ (NRC 2001, S. 241).

„A model of cognition and learning should serve as the cornerstone of the assessment design process. This model should be based on the best available understanding of how students represent knowledge and develop competence in the domain. The model of learning can serve as a unifying element – a nucleus that brings cohesion to curriculum, instruction, and assessment“ (a.a.O., S. 2).

“This model may be fine-grained and very elaborate or more coarsely grained, depending on the purpose of the assessment, but it should always be based on empirical studies of learners in a domain. Ideally, the model will also provide a developmental perspective, showing typical ways in which learners progress toward competence“ (a.a.O., S. 5).

Einen gemeinsamen Begriff von Kompetenzen, deren Abstufung und Entwicklungslogik zu entfalten und darin das Verhältnis von fachlichem und außerfachlichem Lernen zu klären wird daher in Zukunft eine der wichtigsten Herausforderungen der Arbeit an Bildungsstandards sein. Wie in Kapitel 6 festgestellt, ist dies eine sehr anspruchsvolle Aufgabe. Nicht zuletzt die PISA-Studie hat hierbei jedoch entscheidende Fortschritte gebracht. In den Bereichen Mathematik (vgl. Kapitel 3 sowie Abschnitt 6.2), Lesekompetenz (Artelt et al. 2001), in den Naturwissenschaften (Prenzel et

al. 2001; vgl. auch Anhang b) sowie auch in den Fremdsprachen (vgl. Anhang a) liegen Modelle für Komponenten und Stufen der jeweiligen Kompetenzen vor. Auch die Autoren der genannten amerikanischen Expertise gehen davon aus, dass das Ziel, aussagefähige Kompetenzmodelle zu entwickeln und in Testverfahren umzusetzen, viele interdisziplinäre, Praxis und Wissenschaft einschließende Entwicklungsarbeiten erfordert.

### **12.3 Aktuelle Arbeiten in der Kultusministerkonferenz**

Länder-übergreifend hat die KMK eine Reihe von Arbeitsgruppen eingesetzt, die seit dem Herbst 2002 an den geplanten nationalen Bildungsstandards und den neuen Abiturprüfungsanforderungen arbeiten. Sie sind hauptsächlich mit Fachexperten aus den Landesinstituten für Lehrerbildung und Curriculumentwicklung besetzt, die über große Erfahrungen in der Lehrplanarbeit verfügen. Daneben wirken Praktiker und Fachdidaktiker mit. Eine Lenkungsgruppe, die in Personalunion vom Vorsitzenden des Schulausschusses der KMK geleitet wird, hat den Arbeitsgruppen Leitlinien vorgegeben. Diese Lenkungsgruppe wird von Erziehungswissenschaftlern und Fachdidaktikern beraten, darunter einige Mitglieder der Arbeitsgruppe, welche die vorliegende Expertise verfasst hat.

Der Prozess, der jetzt in der KMK in Gang gesetzt wurde, korrespondiert in vielerlei Hinsicht mit den Empfehlungen, die in dieser Expertise gegeben werden. So orientieren sich die Arbeiten der KMK an modernen fachdidaktischen Konzepten, wenn sie Grundprinzipien des Lernens und Leitbilder der Fächer definieren. Sie lehnen sich zumeist an fachdidaktisch etablierte Kompetenzmodelle an, beispielsweise die Konzeptionen mathematischer Grundbildung in TIMSS und PISA (vgl. Abschnitt 6.2) oder den Gemeinsamen Europäischen Referenzrahmen im Bereich der Fremdsprachen (vgl. Anhang a). Diese Kompetenzen werden im Bereich der Sekundarstufe I für alle Bildungsgänge nach denselben Grunddimensionen und Prinzipien beschrieben.

Die Dokumente, die in der KMK bis zum Frühjahr 2004 vorliegen sollen, werden erklärtermaßen noch nicht alle Anforderungen erfüllen können, die an Bildungsstandards gestellt werden können. Ein solcher Anspruch wäre nicht nur unrealistisch, sondern auch kontraproduktiv. Die KMK wird die Standards fortentwickeln. Drei Fragen, die dabei nach Meinung der Expertengruppe überdacht werden sollten, sollen hier kurz diskutiert werden:

- Für welche Zwecke und in welchen Jahrgangsstufen werden Standards formuliert?
- Handelt es sich um Mindeststandards oder um Regelstandards, die eine durchschnittliche Leistungserwartung ausdrücken?
- Werden Komponenten und Stufen von Kompetenzen ausgewiesen?

Die KMK greift bei der Entwicklung von Bildungsstandards sinnvollerweise auf vorhandene Vereinbarungen zu den Zielen schulischer Bildungsgänge zurück. Sie ist für die bundesweite Vergleichbarkeit von Abschlüssen und Zertifikaten zuständig, hat zu diesem Zweck Einheitliche Prüfungsanforderungen für die Abiturprüfung (EPA) sowie Vereinbarungen über den mittleren Schulabschluss vorgelegt und definiert dementsprechend zunächst Standards für die Abschlussjahrgänge. Die vorliegenden Entwürfe einzelner Länder gehen bei der Wahl der Jahrgangsstufe unterschiedliche Wege. Möglicherweise wird man sich mittelfristig darauf verständigen, sowohl in Abschlussklassen als auch in davor liegenden Jahrgängen Vergleichskriterien zu definieren. Die Expertengruppe empfiehlt jedenfalls, intermediäre Standards z.B. in den Klassenstufen 3 für die Grundschule und 9 für den mittleren Schulabschluss zu definieren. Insbesondere sollten sich die Vergleichsarbeiten in den Ländern und das nationale Bildungsmonitoring nicht auf die Abschlussjahrgänge beziehen. Rückmeldungen in der Mitte der Bildungsgänge können es den Schulen leichter machen, auf die Befunde auch zu reagieren. Zu dem sind die Ergebnisse von Evaluation und Monitoring in diesem Fall weniger leicht mit Abschlussprüfungen verwechselbar oder als solche missbrauchbar. Hinzu kommt, dass eine standardbezogene Testdurchführung in Jahrgangsstufe 9 mit dem PISA-Programm kompatibel wäre. Das Grundbildungskonzept von PISA und die PISA-Tests sind in der Diskussion in Deutschland weitgehend als Maßstäbe der Bildungsqualität akzeptiert worden, sodass es sinnvoll erscheint, auch die nationalen Standards in Beziehung zu PISA zu setzen und an denselben Schülergruppen (also im 9. Jahrgang) zu untersuchen.

In der KMK hat man sich ferner darauf verständigt, die Bildungsstandards zunächst als „Regelstandards“ zu definieren und noch nicht – wie hier in Kapitel 2 empfohlen wurde – Mindestanforderungen auszuweisen. Für diese Entscheidung sprechen in der Tat einige Argumente: Bei den jetzt zu erarbeitenden ersten Entwürfen würde man sich schwer tun, eindeutige, für die Entwicklung des Bildungssystems herausfordernde, aber auch realistische Minimalstandards festzusetzen. Indem man zunächst ein mittleres, „normales“ Erwartungsniveau anpeilt, lässt man bei der Umsetzung Spielraum, der in den Schulen erprobt und in der Testentwicklung gleichsam ausgelotet werden kann. So wird verhindert, dass der Prozess der Implementierung von Standards gleich mit massiven Unter- oder Überforderungen einhergeht. Es sei daran erinnert, dass Lehrplanexperten, die nach ihren Erwartungen hinsichtlich der PISA-Aufgaben gefragt wurden, die Leistungen deutscher Schülerinnen und Schüler zum Teil massiv überschätzten und als curricularen Erwartungshorizont Aufgaben identifizierten, die beispielsweise in der Mathematik von weniger als 50 Prozent der Testteilnehmer richtig gelöst wurden (Klieme, Neubrand & Lüdtke 2001).

Durch die Konzentration auf ein mittleres Erwartungsniveau ist man zudem in dieser ersten Version von Bildungsstandards noch nicht genötigt, Kompetenz-Modelle mit abgestuften Niveaus zu spezifizieren. Allerdings sollte mittelfristig auf die Entwicklung von Kompetenzmodellen mit ausgewiesenen Dimensionen und Stufen nicht verzichtet werden (siehe Abschnitt 12.2). Es ist also zu hoffen, dass die nationalen Bildungsstandards in späteren Versionen systematisch auf Kom-

petenz- und Kompetenzstufen-Modellen beruhen und dann auch – über Bildungsgänge hinweg - Mindestanforderungen definieren können.

## **12.4 Fortsetzung der Arbeiten ab Frühjahr 2004**

Die Erarbeitung national einheitlicher Standards und Leistungskriterien stellt im föderalen System der Bundesrepublik Deutschland einen sehr drastischen Einschnitt dar. Es gibt in Deutschland – jedenfalls über Landesgrenzen hinweg – keine einschlägigen Erfahrungen. Das Expertenwissen in der Bildungsverwaltung und die Ressourcen für wissenschaftliche Begleitung und Beratung müssen erst aufgebaut werden. Die betroffenen Lehrer, Eltern und Schüler müssen sich mit Standards, Vergleichsarbeiten und ähnlichen Konzepten genauer auseinandersetzen. Die Öffentlichkeit soll eine Chance haben, den Veränderungsprozess kritisch zu kommentieren und zu beeinflussen. Aus all diesen Gründen versteht auch die KMK die ersten Bildungsstandards als Beginn eines Prozesses, in dem sich das Konzept und die Produkte weiterentwickeln können.

Wenn im Herbst 2003 bzw. Frühjahr 2004 Bildungsstandards in einer ersten Version vorliegen, werden sich den Akteuren im Bildungswesen unterschiedliche Aufgaben stellen, für die etwa zwei weitere Jahre veranschlagt werden sollten:

### *1. Rezeption und kritische Auseinandersetzung mit den Standards in den Schulen und Institutionen des Bildungssystems*

Ob Bildungsstandards tatsächlich, wie intendiert, zur Qualitätsentwicklung im Bildungswesen beitragen, entscheidet sich letztlich in den Schulen. Es gilt, die Akzeptanz bei Lehrkräften und Schulleitern zu gewinnen. Ziel und Zweck, gewissermaßen die Notwendigkeit der Einführung von Bildungsstandards in das deutsche Schulsystem müssen verständlich gemacht werden. Die Initiatoren dieser Bildungsreform müssen in den Schulen für die Überzeugung werben, dass Bildungsstandards keinen unnützen „Überbau“ darstellen, sondern den Schulen mehr Freiräume für ihre pädagogische Arbeit geben (vgl. Kapitel 4). Befürchtungen über Einengung und Kontrolle durch Standards müssen aufgenommen werden. Dabei ist es wichtig, Kritik an Form, Inhalt und Nutzung der Standards und Verbesserungsvorschläge in einem öffentlichen Diskurs aufzugreifen und in die Überarbeitung der ersten Entwürfe einfließen zu lassen.

Für die Akzeptanz der Bildungsstandards in den Schulen ist es ausschlaggebend, dass die Schulen in etwa wissen, wie die Standards in den Schulen aufgegriffen und umgesetzt werden sollen. Ministerien, Landesinstitute und andere Akteure müssen den Schulen relativ konkrete Verfahrenspläne unterbreiten können, die für die Schule die zukünftigen Anforderungen und Arbeits-

schritte beschreiben. Dabei ist auch zu berücksichtigen, dass die gesamte Entwicklung von Bildungsstandards als dynamisch fortschreitendes Unternehmen anzusehen ist. Es geht zunächst darum, die beim aktuellen Stand der Ausarbeitung von Bildungsstandards für die Schulen sinnvollen Optionen und Handlungsmöglichkeiten aufzugreifen.

Die Verfahrenspläne werden zunächst einmal vorschlagen, wie sich Schulen mit den Standards vertraut machen können. Die Schulen können ihre Unterrichtspraxis und ihre internen Absprachen zum Curriculum daraufhin überprüfen, ob sie den Schwerpunktsetzungen der Standards entsprechen. Gegebenenfalls kann eine Fachgruppe von Lehrern sich darum bemühen, ein entsprechendes Schul-Curriculum zu erstellen, das, mit den Standards als Richtschnur, auf kumulatives Lernen über mehrere Jahrgänge hin angelegt ist.

Schulentwicklungsberater und Schulaufsicht müssen bereit stehen, Schulen bei der Rezeption, der Umsetzung und der Überprüfung von Standards zu beraten.

In der Aus- und Weiterbildung von Lehrern sollten die Standards, das dahinter liegende Verständnis des jeweiligen Faches, die Kompetenzmodelle und Leistungserwartungen thematisiert werden.

*Last but not least:* Die Lehrplanarbeit der Länder muss sich auf die nationalen Standards einstellen. Dies bedeutet, wie in Kapitel 8 ausgeführt, dass enge Input-Steuerungen zurückgenommen und schrittweise in die Mitverantwortung der Schulen übergeben werden müssen. Die neue Form von Lehrplänen oder Rahmenrichtlinien kann sicherlich in den ersten Jahren nach Einführung der Standards nur in Ansätzen entwickelt werden.

## *2. Öffentliche Diskussion*

Eltern und Öffentlichkeit können sich mit den Bildungsstandards auseinandersetzen. Entsprechen die Zielbestimmungen und Schwerpunktsetzungen der Standards den gesellschaftlichen Ansprüchen an das Bildungssystem? Sind die zu vermittelnden Kompetenzen verständlich dargestellt? Sind die Leistungserwartungen klar genug, durch die Bildungsziele legitimiert und mehr oder weniger realistisch? Passt der Unterricht meines Kindes zu diesen Standards?

Solche und ähnliche Fragen werden Eltern und Öffentlichkeit diskutieren und dabei Vorschläge zur Nutzung der Standards, aber auch zu deren Revision vorlegen.

## *3. Fachdidaktische Vertiefung der Kompetenzmodelle*

Die fachdidaktische Fundierung der Bildungsstandards, insbesondere der Kompetenzmodelle, ist, wie in Abschnitt 12.2 erläutert, ein wesentliches Desiderat der weiteren Arbeit. Es könnten hierzu Expertisen eingeholt werden, die – über Schulstufen und Bildungsgänge hinweg – entsprechende Grundlagen skizzieren. Der Zusammenarbeit mit fachdidaktischen Gesellschaften, der Rezeption einschlägiger Arbeiten aus anderen Staaten und der vertiefenden empirischen Forschung – u.a.

im Zusammenhang mit der Entwicklung von standardbezogenen Tests – kommt besondere Bedeutung zu.

#### *4. Umsetzung in Aufgabenpools und Testverfahren*

Die KMK verbindet Bildungsstandards, wie erwähnt, unmittelbar mit der Zusammenstellung von Aufgabenpools und mit der Entwicklung von Orientierungs- oder Vergleichsarbeiten auf Landesebene. Bevor jedoch solche Arbeiten in größerem Maßstab geschrieben und somit die Erreichung der Standards geprüft wird, müssen entsprechende Tests entwickelt und nach den professionellen Regeln der pädagogischen Diagnostik erprobt werden (vgl. oben, Kapitel 7). In den Kapitel 7 und 11 wurde vorgeschlagen, hierzu eine wissenschaftliche Einrichtung zu gründen, die im Auftrag der KMK und ggf. zugleich auch des Bundes die notwendigen Arbeiten plant, in Auftrag gibt, koordiniert und dabei für Qualitätssicherung und wechselseitigen Austausch sorgt. Diese „Agentur“, die selbst eine eher kleine Einrichtung wäre, könnte ab 2003 aufgebaut werden und ab Frühjahr 2004 die ersten Testentwicklungen anstoßen. Eines der ersten Themen, die in der Agentur zu bearbeiten wären, ist die Frage nach den Jahrgängen, für die Tests entwickelt werden sollen (siehe Diskussion unter 12.3).

Auftragnehmer wären hierbei Forschungsinstitute, Forschergruppen an Hochschulen, Landesinstitute, aber möglicherweise auch kommerzielle Einrichtungen, die in Zusammenarbeit mit Praktikern Aufgaben entwickeln, diese in Abstimmung mit den Behörden empirisch erproben und anhand der Befunde fundierte Kompetenzmodelle weiterentwickeln. Hierzu ist ein Forschungsprozess erforderlich, der sicherlich zwei Jahre dauern würde. In einigen Fächern könnte man hierbei Aufgabenbeispiele und Kompetenzmodelle aus vorangegangenen Vorhaben der Länder nutzen. So werden in 2004 bereits die ersten Erprobungs- und Normierungsläufe für das Projekt VERA in Rheinland-Pfalz (Helmke & Hosenfeld 2003) und für die Grundschul-Orientierungsarbeiten in Bayern stattgefunden haben. Die Aufgabenpools dieser Studien, aber auch die dort verwendeten und überprüften Kompetenzmodelle (Stufen und Teildimensionen) können die Arbeit, die nun auf nationaler Ebene ansteht, wesentlich vereinfachen.

Im Bereich der Sekundarstufe I kann man für die Mathematik inzwischen nicht nur auf TIMSS und PISA, sondern auf Schulleistungsstudien oder Vergleichsarbeiten in fast allen Ländern zurückgreifen. Eine Auswertung solcher Aufgaben aus drei Ländern im Rahmen von PISA 2000 zeigte, dass sie eine gemeinsame Skala mathematischer Kompetenz bilden. PISA 2003, mit seinem Schwerpunkt in Mathematik, wird weiteres Aufgabenmaterial und differenziertere Kompetenzmodelle zur Verfügung stellen. Es dürfte in diesem Fach relativ einfach sein, die vorhandenen Aufgabenpools mit den nationalen Bildungsstandards in Beziehung zu setzen. Ebenfalls auf guter Grundlage kann man in den Naturwissenschaften sowie – basierend auf der Hamburger LAU-Studie (Lehmann, Peek & Gänsfuß 1997) und der für 2003/04 geplanten KMK-Studie zu Deutsch- und Englischleistungen (DESI-Konsortium 2001; Beck & Klieme 2003) – in den Sprachen an die Entwicklung von Aufgaben gehen. Die neu zusammengestellten, an den Standards

orientierten Aufgabensätze müssen aber sodann empirisch erprobt werden. In der Sekundarstufe I erscheint es sinnvoll, Erprobungseinsätze der neuen Tests mit den PISA-Erhebungen zu koppeln, die auch nach 2003 vermutlich alle drei Jahre stattfinden und dabei die Gruppe der 15-jährigen sowie der Neuntklässler einbeziehen wird. Der Aufbau der PISA-Studien mit einer breit angelegten Testerprobung unter „Ernstfallbedingungen“ und einem späteren Einsatz ist für eine solche Testentwicklung und -validierung sehr geeignet. Die Verbindung spart Untersuchungsaufwand, und zugleich ermöglicht sie eine äußerst differenzierte Auswertung und Aufbereitung der Befunde, da PISA einen Vergleich mit internationalen Tests ermöglicht und viele Rahmendaten zu den sozialen, unterrichtlichen und schulischen Rahmenbedingungen zur Verfügung stellt. PISA 2006 – mit seinem Probedurchlauf in 2005 – stünde erstmals für eine solche Arbeit zur Verfügung.

#### *5. Überarbeitung und Inkraftsetzung der Bildungsstandards*

Alle Erkenntnisse sollten in eine revidierte Form der Bildungsstandards einfließen, die etwa 2006 verfügbar wären, die gemeinsam beschlossen und sodann in den Ländern als verbindlich erklärt werden könnten. Diese Überarbeitung sollte aus Sicht der Expertengruppe auch bildungsgangübergreifende Mindestanforderungen beinhalten

### **12.5 Längerfristige Perspektiven**

An eine umfassende Nutzung von nationalen Bildungsstandards für die Evaluation in einzelnen Schulen sowie für das Bildungsmonitoring (siehe Kapitel 9) und somit für Output-orientierte Steuerung des Bildungssystems ist demnach erst etwa ab 2006 oder 2007 zu denken, und auch dann erst nur in wenigen Fächern der Primarstufe und der Sekundarstufe I. Die Entwicklung und Implementation von Bildungsstandards für andere Fächer und für die Sekundarstufe II (evt. auch für den berufsbildenden Bereich) ist ein längerfristiges Projekt, das weit über 2006/2007 hinausreicht.

Dieser Zeithorizont mag viele überraschen. Allerdings wäre die Bildungspolitik aus Sicht dieser Expertengruppe gut beraten, die Zeitvorgaben nicht willkürlich enger zu setzen, ebenso wie die Öffentlichkeit gut beraten wäre, keinen unrealistischen Handlungsdruck auf Bildungspolitik und -administration auszuüben.

- Die Vorlage der ersten Bildungsstandards und deren breite Diskussion ist bereits ein wichtiger bildungspolitischer Impuls.
- Bildungsstandards gewinnen nicht erst durch Testeinsätze Bedeutung. Als Referenzrahmen für professionelles Handeln in den Schulen, als Orientierung für Schulen und Unterstützungs-

systeme und als Leitidee für die bildungspolitische Öffentlichkeit sind sie auch ohne Evaluation und Monitoring wichtig. Hier sei vor allem an die Mathematik-Standards der NCTM erinnert, die ihren großen Einfluss allein der pädagogischen Vision und dem Kompetenzmodell verdanken, aber keinen Tests.

- Wer innerhalb einzelner Länder oder Regionen Schulevaluation betreiben will, kann – für bestimmte Fächer und Qualitätskriterien – bereits jetzt auf eine Fülle von Instrumenten zurückgreifen, die aus TIMSS und PISA, aus landesweiten Vergleichsarbeiten und Schulleistungstudien stammen. Sie sind zwar nicht explizit an den nationalen Bildungsstandards orientiert, können aber den Schulen valide Hinweise auf mögliche Stärken und Schwächen geben. Auch für das Bildungsmonitoring bieten die für 2003 und 2006 geplanten PISA-Zyklen sowie die den Sprachkompetenzen (Deutsch, Englisch) gewidmete DESI-Studie ausreichendes Material. Und für Grundschulen wird es schon im Frühjahr 2003 hoch interessante Befunde aus der Internationalen Grundschulstudie IGLU geben.
- Wiederholt haben wir in dieser Expertise darauf hingewiesen, dass Standards und darauf aufbauende Tests ihren pädagogischen Nutzen vor allem durch die Verankerung in Kompetenzmodellen erhalten. Solche Modelle zu entwickeln, braucht Zeit und fundierte wissenschaftliche Unterstützung.
- Deutschland verfügt derzeit in der Bildungsadministration und der Bildungswissenschaft nur in Ansätzen über die erforderlichen Ressourcen. Die Deutsche Forschungsgemeinschaft hat auf die internationalen Vergleichsstudien mit großen Anstrengungen zur Intensivierung der Bildungsforschung und zur Förderung des wissenschaftlichen Nachwuchses reagiert, die aber erst nach einigen Jahren greifen werden.
- Angesichts der Vielzahl von Fragen, die mit dem Übergang zu output-orientierten Steuerungsverfahren verbunden sind, ist eine sorgfältige Planung erforderlich. Gerade am Beispiel Großbritanniens und der USA kann man ersehen, dass eine pädagogisch unzureichend reflektierte Nutzung von Standards (etwa im Sinne eines einfachen Schul-Rankings, wie in Großbritannien zeitweise üblich, oder einer Fehl-Nutzung von Schulleistungstests für individuelle Bewertungen wie beim *high stakes assessment* in den USA) kontraproduktiv sein kann.
- Die Diskussion über Bildungsstandards und deren Entwicklung sowie der Erfolg der vorgeschlagenen Maßnahmen wird aus der Sicht unserer Arbeitsgruppe ganz wesentlich von der Akzeptanz des Gesamtansatzes, insbesondere durch die Lehrerschaft, abhängen. Nur wenn es gelingt, die praktizierenden Lehrerinnen und Lehrer zu den Hauptträgern dieser Reform zu machen, nur wenn sie diese Reform als im ureigenen professionellen Selbstinteresse liegend verstehen, wird es auf lange Sicht zu einer Erneuerung und qualitativen Veränderung der Kultur des Lehrens und Lernens kommen.

- Bevor Bildungsreformen Ergebnisse erzielen, die empirisch nachweisbar sind, vergehen mit Sicherheit nicht eine, sondern zwei oder drei Legislaturperioden. Dies haben uns Länder wie Schweden, Finnland und Kanada gezeigt. Auch für Großbritannien sind die Ende der 80er Jahre eingeleiteten Reformen erst mehr als zehn Jahre später in internationalen Vergleichen sichtbar geworden.

Der amerikanische Naturwissenschaftsdidaktiker Rodger Bybee hat jüngst (2002) anhand der Einführung neuer Literacy-Konzepte für den Naturwissenschaftsunterricht dargelegt, dass solche Reformen von der Absichtserklärung bis zur Entwicklung von Leitideen und Standards drei bis vier Jahre benötigen, aber für deren Implementation, d.h. für die Entwicklung und Einführung innovativer Unterrichtskonzepte und -materialien, die Fortbildung von Lehrern usw. noch einmal vier bis sechs weitere Jahre. Diese Zeitperspektive wird verständlich, wenn man mit Bybee einkalkuliert, welche Größenordnungen solche Reformen haben: Im Verlauf der Implementierung steigt die Zahl der Beteiligten, werden immer mehr Institutionen einbezogen, wächst die Zahl und der Detaillierungsgrad der zu entwickelnden Materialien, werden immer tiefere Schichten des Handelns erfasst (von zunächst allgemeinen Zielorientierungen bis hin zu individuellen Unterrichtspraktiken), und damit steigt nicht zuletzt die Schwierigkeit, Personen, Institutionen und Interessengruppen für die Mitarbeit zu gewinnen und gemeinsame Überzeugungen zu entwickeln oder solche zu erhalten. Die Bildungspolitik braucht daher für Projekte wie die Einführung von Bildungsstandards einen langen Atem.

## Anhang (a):

# Der Gemeinsame Europäische Referenzrahmen für Sprachen

### Ein Beispiel für die Entwicklung eines Kompetenzmodells und von Sprachkompetenzskalen

Zum Abschluss soll noch ein Ausblick auf die weitere Entwicklung von Kompetenzmodellen gegeben werden. Dazu greifen wir als Beispiel auf europäische Ansätze aus dem Bereich der Fremdsprachen und des Fremdsprachentestens zurück. Konkret soll hier der *Gemeinsame Europäische Referenzrahmen für Sprachen: Lernen, Lehren, Beurteilen* des Europarats vorgestellt werden (Council of Europe 2000; deutsche Fassung: Goethe-Institut u.a. 2001; im Folgenden als GER oder *Referenzrahmen* abgekürzt), der bereits seit längerem entwickelt und in verschiedenen Versionen bekannt geworden ist und der auch in der deutschen Diskussion bereits eine beachtliche, wenn nicht überragende Rolle spielt. Der Referenzrahmen zielt darauf ab, Transparenz und Vergleichbarkeit bei der Bestimmung, Entwicklung und Überprüfung von Fremdsprachenkompetenzen in Europa herzustellen und damit eine breite Basis für die Planung des Lehrens, des Lernens und vor allem des Evaluierens von Fremdsprachenfähigkeit zur Verfügung zu stellen. Er ist das Produkt einer kontinuierlichen, über 30-jährigen Grundlagenarbeit des Europarats zur Ausformulierung eines Kompetenzmodells. Dennoch muss man vorwegnehmend festhalten, dass der Gemeinsame Europäische Referenzrahmen zum jetzigen Zeitpunkt immer noch eine Art aktueller (theoretischer) Baustelle bzw. eine un-abgeschlossene Innovation darstellt, die sich sehr wohl aber auch als ein „Schatzkästlein“ (Quetz 2003) oder als ein „Steinbruch“ begreifen lässt, auf die sich verschiedene Nutzergruppen professionell beziehen, aus dem sie Ideen und Perspektiven gewinnen und je nach kons-tatiertem Bedarf Anregungen zur Formulierung von Standards ableiten können. Dabei präsentiert sich der Referenzrahmen als ein „flexibles Verzweigungsmodell“ (Goethe-Institut u. a. 2001, 40), dessen definierte Kompetenzbereiche jeweils weiter zu untergliedern bzw. zu sub-dimensionieren sind und dessen Skalenniveaus ebenfalls weiter ausdifferenziert und verfeinert werden können. Wie dies in wissenschaftlich begründeter Weise passieren kann, dazu gibt der GER ebenfalls Auskunft.

## 1. Wie sieht das Kompetenzmodell des Referenzrahmen aus?

Das Kompetenzmodell des Referenzrahmens wird als sprachliches Handlungsmodell definiert. Dieses Modell beschreibt, was es bedeutet eine (Fremd)Sprache zu können, was alles dazu gehört und wie der jeweilige Ausprägungsgrad von Sprachbeherrschung auf jeder Dimension, in jeder Teilkompetenz (verbal) am besten zu formulieren ist. Der Referenzrahmen selbst ist kein diagnostisches Instrument, aber er systematisiert Dimensionen und Stufen sprachlicher Handlungsfähigkeit, konkretisiert sie durch verhaltensnahe Deskriptoren und ist so eine gute Ausgangsbasis für die Entwicklung von Aufgaben, Tests und anderen Prüfungsverfahren, mit denen dann die Sprachkompetenzen einzelner Personen eingeschätzt werden können.

Die Dimensionen des GER sind im nationalen wie im transnationalen Diskurs breit verankert, theoretisch plausibel und konsensfähig. Auf der Basis dieser Dimensionen lassen sich Testinstrumente entwickeln, die beschreiben und erfassen können, was ein beliebiger Fremdsprachenlerner mit der jeweils gelernten Sprache machen kann (was er verstehen, ausdrücken oder mitteilen kann) – egal ob er nun ganz am Anfang des Lernprozesses steht oder ob er über die Schule hinaus lebenslang weiter lernt und damit in seinen Kenntnissen und in seinem Können auch sehr weit fortgeschritten sein kann. Mit Hilfe dieser Instrumente lassen sich also ebenso fremdsprachliche Anfänger wie auch außerordentlich kompetente Sprecher einer Fremdsprache (*near-nativeness*) zuverlässig einstufen. Damit erfüllt der GER eine wesentliche Forderung, die in dieser Expertise an Kompetenzmodelle für Standards gerichtet wird: er bildet Unterschiede und Entwicklungsstufen ab und ist so geeignet, Kompetenzniveaus innerhalb eines kumulativen Lernprozesses zu identifizieren. Der GER erlaubt es dabei, die der Sprachbeherrschung zugrunde liegenden Kompetenzen bzw. den jeweiligen Stand der Sprachfähigkeit in ihrer ganzen Breite, mit all ihren Facetten, zu erfassen, und zwar unabhängig davon, wie die Sprache im einzelnen gelernt und erworben worden ist: Er konzentriert sich darauf festzustellen, *ob und wie gut etwas gekonnt ist* und will dazu ein möglichst umfassendes Bild zeichnen. Auch dadurch eignet er sich hervorragend als Basis für Standards und für sog. Sprachstandsmessungen zu beliebigen Zeitpunkten in der Entwicklung bzw. im Leben eines Fremdsprachenlernenden (vgl. Vollmer 2003).

Der GER beruht auf einem umfassenden Verständnis von menschlicher Kommunikationsfähigkeit, zu der sowohl allgemeine (weniger eng mit der Sprache verknüpfte) Kompetenzen als auch und vor allem kommunikative Sprachkompetenzen beitragen; dementsprechend unterscheidet er diese beiden globalen Kompetenzbereiche und untergliedert sie jeweils wie folgt weiter:

## Allgemeine Kompetenzen

- 1) Deklaratives (Welt)Wissen,
- 2) Allgemeine praktische Fertigkeiten und prozedurales Wissen,
- 3) Persönlichkeitsbezogene Kompetenzen (wie Einstellungen, Motivationen, Wertvorstellungen, kognitiver Stil und viele weitere Persönlichkeitsfaktoren),
- 4) Lernfähigkeit (Sprach- und Kommunikationsbewusstsein, Lerntechniken).

## Sprachbasierte kommunikative Kompetenzen (im engeren Sinne)

1. Linguistische Kompetenzen (Verfügen über Sprachwissen und über Regeln zu dessen Anwendung),
2. Soziolinguistische Kompetenzen (Kenntnis und Beachtung der soziokulturellen Bedingungen der Sprachverwendung),
3. Pragmatische Kompetenzen (Funktionaler Gebrauch sprachlicher Ressourcen).

Jeder dieser drei Teilbereiche wird mehrfach unterteilt, so dass ein recht verzweigtes Modell von Kommunikativer Kompetenz vorliegt, das je nach Bedarf weiter spezifiziert werden kann.

- Innerhalb des *Linguistischen Kompetenzbereichs* wird zwischen lexikalischer, grammatischer, semantischer, phonologischer, orthographischer und orthoepischer Kompetenz unterschieden.
- *Pragmatische Kompetenzen* werden *nach Diskurskompetenz, Funktionaler Kompetenz und Schemakompetenz* unterteilt. Die Diskurskompetenz wiederum (an anderer Stelle auch *Textkompetenz* genannt) umschließt Dimensionen wie *Flexibilität* (in Bezug auf die Umstände der Kommunikationssituation), *Sprecherwechsel* (nur fürs Mündliche), *Themenentwicklung* sowie *Kohärenz* und *Kohäsion*; zum letzteren gehört dann auch die *Kenntnis und Beherrschung verschiedener Textsorten und Texttypen*.
- Innerhalb der *Soziolinguistischen Kompetenzen* werden als Komponenten die *Fähigkeit zur sprachlichen Kennzeichnung sozialer Beziehungen*, die *Kenntnis und Beachtung von Höflichkeitskonventionen* und *von Registerunterschieden*, die *Kenntnis von Redewendungen, Aussprüchen, Zitaten usw.* sowie die *Fähigkeit, sprachliche Variation zu erkennen*, genannt (vgl. Goethe-Institut u.a. 2001, bes. 118-120).

Übereinstimmend mit dem in Kapitel 6 ausgeführten Verständnis werden diese Kompetenzen als Dispositionen betrachtet, also als interne Repräsentationen, Mechanismen und Fähigkeiten, die

das beobachtbare Verhalten und die Performanz von Menschen bestimmen. „Dabei gilt, dass jedes Lernen diese internen Repräsentationen, Mechanismen und Fähigkeiten weiterentwickelt und verändert“ (ebd., S. 25).

Schließlich werden noch zwei weitere Kompetenzdimensionen benannt, die das Sprachhandeln wesentlich mitbestimmen (sollen): zum einen handelt es sich um *kommunikative Strategien*, die jegliches Sprachhandeln mental begleiten bzw. steuern und die insofern ein wichtiger Teil der Sprachfähigkeit sind (in Anlehnung an Bachman 1990, und Bachman & Palmer 1996), zum anderen ist von *Interkulturellen Fähigkeiten* die Rede.

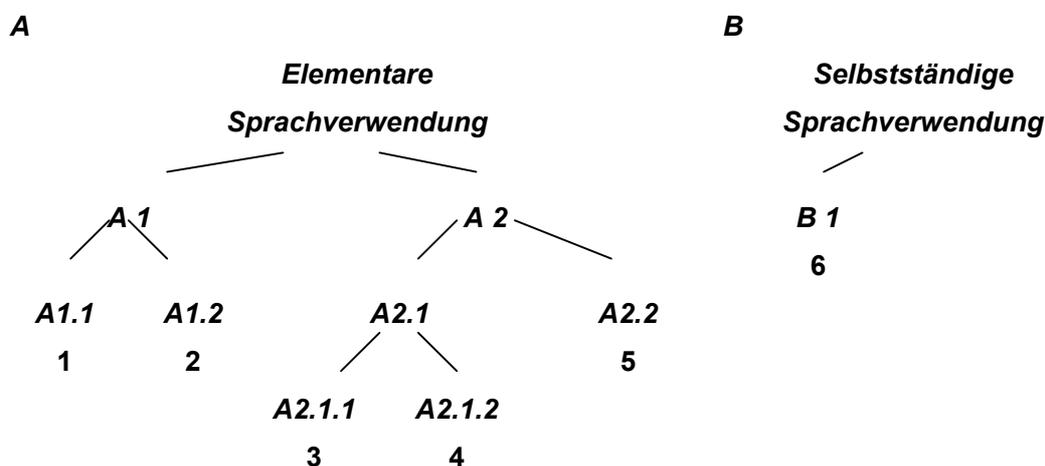
- Der Referenzrahmen geht jedoch über die Identifizierung und Skizzierung der zugrunde liegenden Sprachkompetenzen weit hinaus: er geht davon aus, dass diese nicht direkt beobachtbar sind, sondern dass sie sich in verschiedenen *kommunikativen Sprachaktivitäten* (rezeptiv, interaktiv, produktiv) dokumentieren; diese finden ihrerseits in bestimmten *Lebensbereichen* (Domänen) (z.B. öffentlich, privat, beruflich) bzw. innerhalb der Domänen in bestimmten *Situationen* statt und erhalten dadurch ihren Kontext. Schließlich realisieren sich die genannten Kompetenzen immer in der Bewältigung bestimmter Aufgaben (die ihrerseits bestimmte Strategien der Bearbeitung erfordern) und führen letztlich zum Verstehen, Bearbeiten oder Produzieren bestimmter Texte bzw. Textsorten oder Textfunktionen als den gesellschaftlich relevanten Ausdrucksformen von Sprachlichkeit. Auf der Basis dieser Einsicht wird deshalb *zusätzlich unterhalb der erwähnten drei kommunikativen Kompetenzbereiche* ein System kommunikativer Aktivitäten entfaltet, das im Wesentlichen die Performanzbereiche strukturiert und auflistet, in denen Kompetenz sich konkret ausdrückt und niederschlägt. Die sprachlichen Handlungsbereiche *Rezeption, Interaktion, Produktion* werden nach *Mündlich* und *Schriftlich* unterschieden. Innerhalb des mündlich-interaktiven Subbereichs etwa werden sodann weitere Handlungskontexte untergliedert, so z.B. Muttersprachliche Gesprächspartner verstehen, Konversation, Informelle Diskussion, Formelle Diskussion und Besprechungen, Zielorientierte Kooperation, Dienstleistungsgespräche (Transaktionen), Informationsaustausch, Interview usw. Wie leicht erkennbar liegt hier insgesamt eine Nähe zu den üblichen Bereichen kommunikativer Tätigkeit wie Hörverstehen, Leseverstehen, Mündliche Interaktion, Monologische Sprachproduktion und Schreiben vor, diese werden allerdings weiter durchstrukturiert und ausdifferenziert nach Handlungsanlass, Handlungskomplexität, Grad der geforderten Abstraktion oder Expliztheit des Ausdrucks u.a.m. All diese zusätzlichen Performanzkategorien und -parameter sowie deren Ausdifferenzierungen werden benötigt, um für die Beschreibung von Handlungsniveaus, für den Aufbau und die Formulierung von Handlungssequenzen ebenso wie für die Konstruktion von beispielhaften Testaufgaben genügend spezifizierte Kontextbedingungen benennen zu können. Kompetenzen sollen einer Person ja nur zugeschrieben werden, wenn sie diese in unterschiedlichen Kontexten zeigt.

## 2. Wie sehen die Skalen aus, nach denen Sprachkompetenzen beurteilt werden?

Der Referenzrahmen bietet nicht nur ein differenziertes Kompetenzmodell für den Bereich der Fremdsprachenfähigkeit an, sondern er hat auch eine Vielzahl von Beurteilungskriterien und -maßstäben entwickelt, die es erlauben, einen Fremdsprachenlerner nach der Bearbeitung und Lösung bestimmter Aufgaben im Hinblick auf einen bestimmten Kompetenzbereich relativ genau und sicher einzustufen, also eine zuverlässige Aussage darüber zu treffen, wie gut er etwas kann, d.h. auf welchem Niveau eine sprachliche Handlung jeweils vollzogen wurde und wo auf der Gesamtskala diese Leistung genau anzusiedeln ist.

Der Referenzrahmen unterscheidet sowohl für die globaleren Dimensionen als auch für die Unterdimensionen von Fremdsprachenkompetenz jeweils sechs Niveaustufen: A) *Elementare Sprachverwendung*, B) *Selbstständige Sprachverwendung*, C) *Kompetente Sprachverwendung*, jeweils mit zwei Unterausprägungen A1/A2, B1/B2, C1/C2. Jede Niveaustufe in jedem Kompetenzbereich ist durch verbale Deskriptoren umschrieben, die in mehr oder minder differenzierter, leicht verständlicher und nachvollziehbarer Weise beschreiben, was ein Fremdsprachenlerner jeweils kann oder können muss, wenn er einer bestimmten Niveaustufe zugeordnet wird. Diese Stufungen sind inzwischen noch weiter ausdifferenziert worden, entweder durch Verwendung von + oder - Symbolen oder aber durch weitere Dezimalzahlen wie A1.1, A2.1.2 usw. („flexibles Verzweigungsmodell“). Der Referenzrahmen eignet sich also durchaus (je nach Bedarf, nach Lernergruppe oder Bildungssystem) zur Verfeinerung der Kompetenzabstufungen.

Beispielsweise könnte man in einem Schulsystem, das von der Primarstufe bis zur Sekundarstufe I reicht, oder in einem System der Erwachsenenbildung, in denen es für notwendig erachtet wird, Fortschritte auf niedrigeren Stufen sichtbar zu machen, den Stamm 'Elementare Sprachverwendung' weiter verzweigen und ein System von sechs 'Meilensteinen' mit ausdifferenzierten Zwischenstufen auf dem Niveau A 2 (*Waystage*) entwickeln, auf dem es wahrscheinlich eine große Zahl von Lernenden gibt.



Diese Sprachkompetenzskalen mit ihren verschiedenen Deskriptoren (auf den jeweils sechs Niveaustufen, mit möglichen Unterstufen) sind in äußerst aufwendiger Weise mit verschiedenen Experten- und Nutzergruppen entwickelt, mehrfach revidiert und empirisch so gut wie möglich validiert worden (vgl. z.B. North 2000). Dennoch ließen sich eine Reihe von Problemen benennen, die (noch) nicht wirklich geklärt bzw. gelöst sind, z.B. welche Performanzaspekte innerhalb einer Deskriptorformulierung in welcher Gewichtung miteinander verknüpft sind oder ob es sich bei den Niveaustufen um jeweils gleiche Abstände handelt. Für manche Kompetenzbereiche liegen systematisch entwickelte und validierte Skalen vor, für andere noch nicht bzw. in noch nicht hinreichend abgesicherter Form. Zur Illustration seien hier zwei Skalen vorgestellt, eine eher globale zur *Mündlichen Produktion Allgemein* (Abbildung 6) sowie eine spezifischere zur *Flüssigkeit im Mündlichen* (Abbildung 7).

| <b>MÜNDLICHE PRODUKTION ALLGEMEIN</b> |                                                                                                                                                                                                                                  |
|---------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>C2</b>                             | Kann klar, flüssig und gut strukturiert sprechen und seinen Beitrag so logisch aufbauen, dass es den Zuhörern erleichtert wird, wichtige Punkte wahrzunehmen und zu behalten.                                                    |
| <b>C1</b>                             | Kann komplexe Sachverhalte klar und detailliert beschreiben und darstellen und dabei untergeordnete Themen integrieren, bestimmte Punkte genauer ausführen und alles mit einem angemessenen Schluss abrunden.                    |
| <b>B2</b>                             | Kann Sachverhalte klar und systematisch beschreiben und darstellen und dabei wichtige Punkte und relevante stützende Details angemessen hervorheben.                                                                             |
|                                       | Kann zu einer großen Bandbreite von Themen aus seinen/ihren Interessengebieten klare und detaillierte Beschreibungen und Darstellungen geben, Ideen ausführen und durch untergeordnete Punkte und relevante Beispiele abstützen. |
| <b>B1</b>                             | Kann relativ flüssig eine unkomplizierte, aber zusammenhängende Beschreibung zu Themen aus ihren/seinen Interessengebieten geben, wobei die einzelnen Punkte linear aneinandergereiht werden.                                    |
| <b>A2</b>                             | Kann eine einfache Beschreibung von Menschen, Lebens- oder Arbeitsbedingungen, Alltagsroutinen, Vorlieben oder Abneigungen usw. geben, und zwar in kurzen listenhaften Abfolgen aus einfachen Wendungen und Sätzen.              |
| <b>A1</b>                             | Kann sich mit einfachen, überwiegend isolierten Wendungen über Menschen und Orte äußern.                                                                                                                                         |

Abb. 6

| <b>FLÜSSIGKEIT (MÜNDLICH)</b> |                                                                                                                                                                                                                                |
|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>C2</b>                     | Kann sich auch in längeren Äußerungen natürlich, mühelos und ohne Zögern fließend ausdrücken. Macht nur Pausen, um einen präzisen Ausdruck für seine/ihre Gedanken zu finden oder ein geeignetes Beispiel oder eine Erklärung. |
| <b>C1</b>                     | Kann sich beinahe mühelos spontan und fließend ausdrücken; nur begrifflich schwierige Themen können den natürlichen Sprachfluss beeinträchtigen.                                                                               |

|           |                                                                                                                                                                                                                                                                                                                                 |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>B2</b> | Kann sich spontan verständigen und drückt sich auch in längeren und komplexeren Redebeiträgen oft mit bemerkenswerter Leichtigkeit und Flüssigkeit aus.                                                                                                                                                                         |
|           | Kann in recht gleichmäßigem Tempo sprechen. Auch wenn er/sie eventuell zögert, um nach Strukturen oder Wörtern zu suchen, entstehen kaum auffällig lange Pausen.<br>Kann sich so spontan und fließend verständigen, dass ein normales Gespräch mit einem Muttersprachler ohne Belastung für eine der beiden Seiten möglich ist. |
| <b>B1</b> | Kann sich relativ mühelos ausdrücken und trotz einiger Formulierungsprobleme, die zu Pausen oder in Sackgassen führen, ohne Hilfe erfolgreich weitersprechen.                                                                                                                                                                   |
|           | Kann sich ohne viel Stocken verständlich ausdrücken, obwohl er/sie deutliche Pausen macht, um die Äußerungen grammatisch und in der Wortwahl zu planen oder zu korrigieren, vor allem, wenn er/sie länger frei spricht.                                                                                                         |
| <b>A2</b> | Kann sich in kurzen Redebeiträgen verständlich machen, obwohl er/sie offensichtlich häufig stockt und neu ansetzen oder umformulieren muss.                                                                                                                                                                                     |
|           | Kann ohne große Mühe Redewendungen über ein vertrautes Thema zusammenstellen, um kurze Gespräche zu meistern, obwohl er/sie ganz offensichtlich häufig stockt und neu ansetzen muss.                                                                                                                                            |
| <b>A1</b> | Kann sehr kurze, isolierte und meist vorgefertigte Äußerungen benutzen, macht dabei aber viele Pausen, um nach Ausdrücken zu suchen, weniger vertraute Wörter zu artikulieren und Abbrüche in der Kommunikation zu reparieren.                                                                                                  |

Abb. 7

### 3. Welche Bedeutung hat der Referenzrahmen für die Beurteilung und Bewertung von Sprachkompetenz mit Hilfe entsprechender Tests?

Der Europarat umreißt die drei wichtigsten Verwendungszwecke des Referenzrahmens selbst wie folgt: „1. Er kann inhaltliche Spezifizierungen für Tests und Prüfungen liefern. 2. Er kann Kriterien bereitstellen, mit deren Hilfe man bei der Beurteilung einer bestimmten mündlichen oder schriftlichen Leistung feststellen kann, ob ein Lernziel erreicht wurde oder nicht, und er kann dies sowohl für die kontinuierliche Beurteilung durch Lehrende oder die Lerngruppe tun als auch für die Selbstbeurteilung. 3. Er kann helfen, die Kompetenzniveaus von bereits existierenden Tests und Prüfungen zu beschreiben und dadurch den Vergleich zwischen unterschiedlichen Qualifikationssystemen ermöglichen“ (Goethe-Institut u.a. 2001, 30).

In der Tat liegt mit dem Kompetenzansatz und dem Deskriptorensystem des Referenzrahmens, in Kopplung mit den entsprechenden Vorarbeiten der *Association of Language Testers in Europe (ALTE)* ein Raster vor, das in kürzester Zeit Einfluss genommen hat auf die Entwicklung neuerer Sprachtests in Europa, vor allem im außerschulischen Bereich. ALTE hatte zunächst ein fünf-stufiges System anvisiert, auf dem die Prüfungen der Mitgliedsinstitutionen „geeicht“ wurden, um eine Vergleichbarkeit konkurrierender und komplementärer Angebote in Europa zu sichern (ALTE 1998; vgl. auch Milanovic 2001). In Schweden dagegen wird bis auf den heutigen Tag zur Überprüfung der minimal definierten Standards für den Sprachbereich ein siebenstufiges Bewertungsraster angewendet (vgl. Skolverket 2001).

Die *CAN-DO-statements* der ALTE sind inzwischen durch genaue und detaillierte statistische Auswertung vieler Jahrgänge von Prüfungen in aller Welt erprobt und validiert worden. Auch andere Prüfungen wie etwa die der Weiterbildungs-Testsysteme (WBT 1998, 2001), des Deutschen Industrie- und Handelstages (DIHT 2001), des *Goethe-Instituts Inter Nationes* oder die neuen Hochschulzugangsprüfungen für Deutsch als Fremdsprache des TESTDAF-Instituts (vgl. Projektgruppe TestDaF 2000) sowie das vom Europarat geförderte DIALANG-System<sup>5</sup> nehmen explizit auf den Referenzrahmen Bezug, so dass man schon von einem sich festigenden „Verbundsystem“ sprechen kann, dessen einzelne Bausteine ineinander greifen und sich gegenseitig stützen (vgl. Quetz 2003). Auch das von der KMK in Auftrag gegebene Projekt „Deutsch-Englisch-Schülerleistungen International“ (DESI), das derzeit vorbereitet wird, greift bei seinen Testkonzepten für die 9. Jahrgangsstufe die Stufungen des Europarates auf (vgl. Beck & Klieme 2003; Nold 2003).

Im Bereich des staatlichen Bildungswesens hat der Referenzrahmen ebenfalls eine große Dynamik entfaltet: Es finden sich bereits lebhafte Diskussionen darüber, auf welchem Referenzniveau Grund- oder Leistungskurse abschließen sollten (Profil auf B2 bzw. C1?; vgl. die neuen *Einheitlich Prüfungsanforderungen in der Abiturprüfung Englisch* (EPA) vom Mai 2002; KMK 2002a) oder auf welchem Niveau die vorgesehenen Vergleichsarbeiten am Ende von Klasse 10 des Gymnasiums anzusiedeln sind. In den Ländern Baden-Württemberg und Rheinland-Pfalz liegen inzwischen neue „Lehrpläne“ bzw. Rahmencurricula vor, die explizit auf den Referenzrahmen Bezug nehmen und sich an ihm orientieren. Auch die Entwicklung eines zwar offen angelegten, aber konsensfähigen und länderübergreifenden *Portfolios für Sprachen* wird sich notwendigerweise auf die Vorarbeiten des Gemeinsamen Europäischen Referenzrahmens beziehen, geht es hier doch um die Erstellung eines Dokumentations- und Evaluationswerkzeugs, das der kontinuierlichen Weiterentwicklung des Lehrens und Lernens von Sprachen (einschließlich der Selbstbewertung) gerade an den Schnittstellen von der Grundschule zum Sekundarschulwesen sowie von dort zum Berufsleben und zum nicht-institutionalisierten Weiterlernen dient. Die Arbeit an einem solchen gemeinsamen Portfolio für alle Länder könnte u.a. dazu beitragen, einheitliche Beurteilungskriterien für die Erfassung von Lernerträgen und Leistungsständen zwischen Grundschule und weiterführenden Schulen bzw. zwischen Sekundarschulen und Berufswelt zu vereinbaren und auf diese Weise Kompetenzen und Kompetenzniveaus aller aufeinander aufbauen zu lassen. Dies würde auch eine Anpassung bzw. Erweiterung der Deskriptoren des Referenzrahmens für curriculare Zwecke sowie die Erstellung und Erprobung möglicher Beispielaufgaben bedeuten und damit eine enge Kooperation der Länder erfordern (vgl. zum Europäischen Portfolio der Sprachen allgemein Babylonia 1999, 2000; Schneider 2001; zur Diskussion in Deutschland u.a.

---

<sup>5</sup> Es handelt sich hierbei um ein internetbasiertes System zur Selbstbeurteilung, die durch einen Test überprüft wird, dessen Rückmeldung mit einer Lernberatung gekoppelt ist. Das System ist für 14 Sprachen vorgesehen und soll auch Lernenden im beruflichen Bereich offen stehen, die nicht mehr in institutionalisiertes Lernen eingebunden sind. Auch DIALANG benutzt wie gesagt die sechs Referenzniveaus des GER als Stufungsinstrument (DIALANG 2002).

Landesinstitut für Schule und Weiterbildung 2000; Thüringer Kultusministerium 2002; KMK 2002b; Thümann 2003).

Obwohl also der Europarat in seinem Referenzrahmen Begriffe wie „Standard“ oder „Kerncurriculum“ überhaupt nicht verwendet, werden sein Kompetenzmodell und seine Stufungen, Kategorien wie Beschreibungsansätze bereits in diesem Sinne verwendet. Der Referenzrahmen kann als domänenspezifisches Beispiel dafür gelten, wie auf der Grundlage eines ausdifferenzierten Kompetenzmodells bestimmte Erwartungen nach Qualifikationen und Qualifikationsniveaus (mit entsprechenden Systemen der *Ausbildung* und der *Förderung* auf diese Zielsetzungen hin) formuliert werden könnten, die dann ihrerseits durch breite Konsensbildung zur *Setzung von verbindlichen Bildungs- oder Leistungsstandards* (mit entsprechenden *Forderungen* und einer externen *Überprüfung auf Erreichung* bzw. *Einhaltung*) führen können.

## Anhang (b):

### “Common Framework of Science Outcomes” (Kanada)

Der *Common Framework of Science Outcomes* ist ein Ergebnis der überregionalen Zusammenarbeit der kanadischen Bildungsminister (Council of Ministers of Education). Er geht zurück auf die *Victoria-Declaration* von 1993, in der die künftigen Entwicklungslinien im Bildungswesen hinsichtlich des lebenslangen Lernens und gesellschaftlicher Anforderungen vorgezeichnet wurden. Es wurde die Absicht formuliert, aufgrund gemeinsam geteilter Bildungsziele die Harmonisierung bei der Erreichung dieser Ziele zu sichern. Ein Schritt auf diesem Weg war das *Pan-Canadian Protocol for Collaboration on School Curriculum* von 1995. Das Protokoll bestätigt zwar die Zuständigkeit der Provinzen für die Steuerung des Bildungssystems, aber gleichzeitig erkennt es an, dass eine Kooperation zwischen den provinziellen Zuständigkeitsbereichen dazu beitragen kann, die Qualität von Bildung im gesamten Land zu verbessern. Der *Common Framework of Science Outcomes* ist das erste gemeinsame, durch das Protokoll initiierte Projekt. Ein erklärtes Ziel des *Frameworks* ist es, die Harmonisierung von naturwissenschaftlichen Curricula zu befördern. Der Adressat dieses *Framework* sind in erster Linie diejenigen, die sich mit der Entwicklung von Curricula beschäftigen.

Wir drucken den *Common Framework* hier in Auszügen ab, weil wir in ihm ein gutes Beispiel dafür sehen, wie in einem föderalen System über Länder/Provinzen hinweg auf der Basis fachdidaktischer Überlegungen (hier: auf der Grundlage einer differenzierten Konzeption von *scientific literacy*) Standards entwickelt werden. Interessant ist auch der Aufbau des *Framework*, in dem man deutlich die Komponenten (1) Bildungsziele, (2) Kompetenzmodell und (3) abgestufte Kompetenzanforderungen identifizieren kann, die im Sinne der hier vertretenen Konzeption zur Formulierung von Bildungsstandards gehören. Die Kompetenzdimensionen ziehen sich – ähnlich wie bei den in Kapitel 3 dargestellten Mathematik-Standards der NCTM – über die Bildungslaufbahn durch und werden für einzelne Jahrgänge konkret entfaltet, woraus sich eine systematische, auf kumulative Kompetenzentwicklung angelegte Konzeption des schulischen Lehrens und Lernens ergibt.

|                                                                                                             |
|-------------------------------------------------------------------------------------------------------------|
| <b>Präambel: Gesellschaftliche Zielsetzung und Stellung des Faches im Prozess von Bildung und Erziehung</b> |
|-------------------------------------------------------------------------------------------------------------|

**a) A vision for scientific literacy in Canada**

The framework is guided by the vision that all Canadian students, regardless of gender or cultural background, will have an opportunity to develop scientific literacy. Scientific literacy is an evolving combination of the science-related attitudes, skills, and knowledge students need to develop inquiry, problem-solving, and decision-making abilities, to become lifelong learners, and to maintain a sense of wonder about the world around them.

Diverse learning experiences based on the framework will provide students with many opportunities to explore, analyse, evaluate, synthesize, appreciate, and understand the interrelationships among science, technology, society, and the environment that will affect their personal lives, their careers, and their future.

**b) The scientific literacy needs of Canadian students and society**

Canadian society is experiencing rapid and fundamental economic, social, and cultural changes that affect the way we live. Canadians are also becoming aware of an increasing global interdependence and the need for a sustainable environment, economy, and society. The emergence of a highly competitive and integrated international economy, rapid technological innovation, and a growing knowledge base will continue to have a profound impact on our lives. Advancements in science and technology play an increasingly significant role in everyday life. Science education will be a key element in developing scientific literacy and in building a strong future for Canada's young people. Consistent with views expressed in a variety of national and international science education documents, the following goals for Canadian science education have been established for the purposes of this framework. Specifically, science education aims to:

- encourage students at all grade levels to develop a critical sense of wonder and curiosity about scientific and technological endeavours
- enable students to use science and technology to acquire new knowledge and solve problems, so that they may improve the quality of their own lives and the lives of others
- prepare students to critically address science-related societal, economic, ethical, and environmental issues

- provide students with a foundation in science that creates opportunities for them to pursue progressively higher levels of study, prepares them for science-related occupations, and engages them in science-related hobbies appropriate to their interests and abilities
- develop in students of varying aptitudes and interests a knowledge of the wide variety of careers related to science, technology, and the environment

*Science education must be the basis for informed participation in a technological society, a part of a continuing process of education, a preparation for the world of work, and a means for students' personal development.*

*Science Council of Canada (1984)*

## **1. Bildungsziele**

### **Foundation statements for scientific literacy in Canada**

*Scientific literacy should remain the abstract image that leads science education reform. Eisenhart, M. et al. (1996)*

In light of the vision for scientific literacy and the need to develop scientific literacy in Canada, four foundation statements were established for this framework. Curriculum developers should note that these foundation statements delineate the four critical aspects of students' scientific literacy. They reflect the wholeness and interconnectedness of learning and should be considered as interrelated and mutually supportive. The learning outcomes in this framework are stated in relation to these foundation statements.

#### **Foundation 1:**

##### **Science, technology, society, and the environment (STSE)**

Students will develop an understanding of the nature of science and technology, of the relationships between science and technology, and of the social and environmental contexts of science and technology.

#### **Foundation 2:**

##### **Skills**

Students will develop the skills required for scientific and technological inquiry, for solving problems, for communicating scientific ideas and results, for working collaboratively, and for making informed decisions.

### **Foundation 3:**

#### **Knowledge**

Students will construct knowledge and understandings of concepts in life science, physical science, and Earth and space science, and apply these understandings to interpret, integrate, and extend their knowledge.

### **Foundation 4:**

#### **Attitudes**

Students will be encouraged to develop attitudes that support the responsible acquisition and application of scientific and technological knowledge to the mutual benefit of self, society, and the environment.

### **Exkurs: Didaktische Voraussetzungen zum Erreichen der Bildungsziele**

#### **Development of the four foundation statements**

Bevor die *Foundations* im einzelnen beschrieben werden - bevor also im Sinne dieser Expertise ein Kompetenzmodell entfaltet wird -, stellt der Common Framework dar, welche Grundkonzeption des naturwissenschaftlichen Lernens und Lehrens zugrunde gelegt wird. Auch wenn der Framework dezidiert keine Festlegungen zur didaktisch-methodischen Ausgestaltung des Unterrichts enthält, gründet er doch – wie die in Kapitel 3 dargestellten Mathematik-Standards der NCTM - in einer „Vision“ des naturwissenschaftlichen Unterrichts. Hier wird deutlich, dass diese Standards auf verständnisorientiertes Lernen abzielen. Der Unterricht soll Lernaktivitäten beinhalten, die entdeckendes Lernen und Problemlösen ermöglichen.

Curriculum developers should note that the following considerations about student learning and the teaching of science were taken into account during the development of the framework.

Student learning is affected by personal and cultural preconceptions and prior knowledge. Students learn most effectively when their study of science is rooted in concrete learning experiences, related to a particular context or situation, and applied to their world where appropriate. Science activities, therefore, occur within a socio-cultural context, are interpreted within that context, and are designed to extend and challenge existing views.

The ideas and understandings that students develop are progressively extended and reconstructed as students grow in their experiences and in their ability to conceptualize. Learning involves the process of linking newly constructed understandings with prior knowledge and adding new contexts and experiences to current understandings.

Learning is enhanced when students identify and solve problems. Through such learning, students develop attitudes, skills, and a knowledge base that allow them to explore increasingly complex ideas and problems, especially if these are placed in a meaningful context.

Students learn to understand the world by developing personal conceptions, constructing mental images, and sharing these with others using everyday language, in diverse situations that respect a wide variety of learners.

*[I]t is important for students to learn that they can understand and deal with the world by means of their own observations and constructed explanations, that all such explanatory frameworks have their limitations, and that science offers frameworks for explanations and control which, while also limited in scope, have been shown to possess particular explanatory power and which have thus become accepted by the scientific community and by society as a whole.*

*Science Council of Canada (1984)*

*"(...) Presenting a body of knowledge to students (whether it is in telling them more or showing them better) will not suffice in order for students to understand, memorize and internalize that knowledge. Every student must individually and personally construct each bit of understanding, using tools at her or his disposal, namely her or his own ideas and thought processes."*

*De Vecchi, G. & Giordan, A. (1990)*

### **Teaching of science**

This framework of outcomes is designed to support the development in students of the attitudes, skills, and knowledge needed for developing problem-solving and decision-making abilities, for becoming lifelong learners, and for maintaining a sense of wonder about the world around them in short, to develop scientific literacy.

Development of scientific literacy is supported by instructional environments that engage students in active inquiry, problem solving and decision making. Diverse learning experiences involve designing activities so they are set in meaningful contexts. It is through these contexts that students discover the significance of science to their lives and come to appreciate the interrelated nature of science, technology, society, and the environment.

To facilitate instructional planning, examples of instructional contexts (called "illustrative examples") are provided in the section that presents learning outcomes by grade. The selection of particular contexts and their development will likely vary with the local situation, and reflect factors such as the prior learning of the students, the dynamics of the classroom, the nature of the local environment, and available learning resources.

Although the particular contexts may vary, the overall scope and focus will normally include the following broad areas of emphasis:

- a science inquiry emphasis, in which students address questions about the nature of things, involving broad exploration as well as focussed investigations
- a problem-solving emphasis, in which students seek answers to practical problems requiring the application of their science knowledge in new ways
- a decision-making emphasis, in which students identify questions or issues and pursue science knowledge that will inform the question or issue

Each of these three areas of emphasis provides a potential starting point for engaging in an area of study. These studies may involve a variety of learning approaches for exploring new ideas, for developing specific investigations, and for applying the ideas that are learned. Specific ways of encouraging students to explore, develop and apply ideas are modelled in the illustrative examples.

To achieve the vision of scientific literacy, students must increasingly become engaged in the planning, development, and evaluation of their own learning activities. In the process, they should have the opportunity to work collaboratively with other students, to initiate investigations, to communicate their findings, and to complete projects that demonstrate their learning.

## **2. Allgemeines Kompetenzmodell**

### **Description of the foundation statements**

Die *descriptions of the foundation statements* geben einen Überblick über die Kompetenzdimensionen, die im *framework* abgedeckt werden. Dazu gehören folgende vier *foundations*: foundation 1: science, technology, society and the environment; foundation 2: skills; foundation 3: knowledge; foundation 4: attitudes. Im Folgenden wird nur ein *foundation statement* (foundation 2: skills) exemplarisch wiedergegeben.

#### **Foundation 2:**

##### **Skills**

Students will develop the skills required for scientific and technological inquiry, for solving problems, for communicating scientific ideas and results, for working collaboratively, and for making informed decisions.

Students use a variety of skills in the process of answering questions, solving problems, and making decisions. While these skills are not unique to science, they play an important role in the

development of scientific understandings and in the application of science and technology to new situations. The listing of the skills is not intended to imply a linear sequence or to identify a single set of skills required in each science investigation. Every investigation and application of science has unique features that determine the particular mix and sequence of skills involved. Skills are identified for each grade grouping and at each grade level. Most of the basic skills are given considerable attention in the early years, while specific skills are developed and refined in the senior years.

Four broad areas of skills are outlined in the framework. Each group of skills is developed from kindergarten to grade 12, with increasing scope and complexity of application.

### **Initiating and planning**

These are the skills of questioning, identifying problems, and developing preliminary ideas and plans.

### **Performing and recording**

These are the skills of carrying out a plan of action, which involves gathering evidence by observation and, in most cases, manipulating materials and equipment.

### **Analysing and interpreting**

These are the skills of examining information and evidence, of processing and presenting data so that it can be interpreted, and of interpreting, evaluating, and applying the results.

(siehe unter 3)

### **Communication and teamwork**

In science, as in other areas, communication skills are essential at every stage where ideas are being developed, tested, interpreted, debated, and agreed upon. Teamwork skills are also important, since the development and application of science ideas is a collaborative process both in society and in the classroom.

*There can be no greater contribution or more essential element to long-term environmental strategies leading to sustainable development that respects the environment... than the education of future generations in matters relating to the environment.*

*UNESCO (1988)*

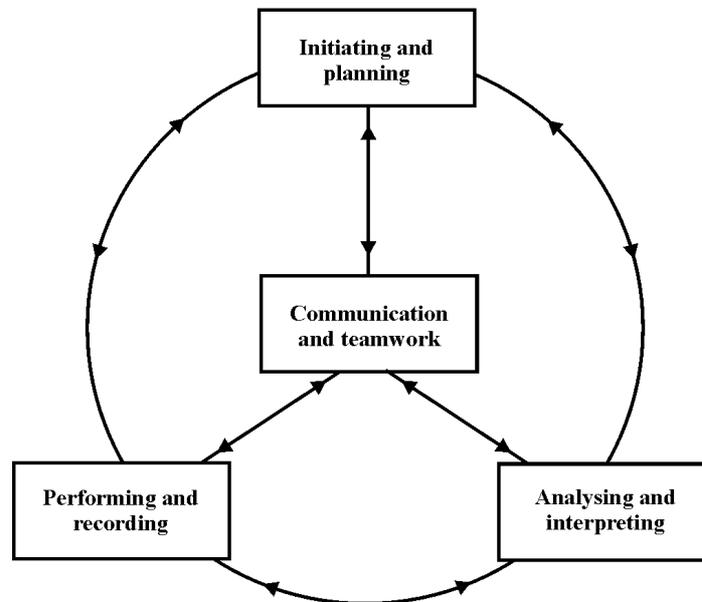
*Science is a creative process which attempts to discover and understand, thereby generating knowledge.... Science is often viewed as both a product and a process.*

*Hart, E.P. (1987)*

Scientific knowledge is necessary but is not in itself sufficient for understanding the relationships among science, technology, society, and the environment. To understand these relationships, it is

also essential to understand the values inherent to science, technology, a particular society, and its environment.

### Interactions among the four areas of skills



As students advance from grade to grade, the understandings about STSE interrelationships are developed and applied in increasingly demanding contexts. In the early years, considerable attention is given to students acquiring an operational understanding of these interrelationships. In the later years, these understandings are more conceptual in nature. Growth in STSE understandings may involve each of the following elements:

- complexity of understanding from simple, concrete ideas to abstract ideas; from limited knowledge of science to more in-depth and broader knowledge of science and the world
- applications in context from contexts that are local and personal to those that are societal and global
- consideration of variables and perspectives from one or two that are simple to many that are complex
- critical judgement from simple right or wrong assessments to complex evaluations
- decision making from decisions based on limited knowledge, made with teacher guidance, to decisions based on extensive research, involving personal judgement and made independently, without guidance

For individual students, the development of STSE understandings may be earlier or later than the times identified in the framework, depending in large part on their stage of cognitive and social development.

### **3. Ausprägung von Kompetenzen auf unterschiedlichen Niveaustufen**

**a) als general learning outcomes**

**b) als specific outcomes**

Im Folgenden wird für die oben beschriebene *foundation 2* (skills) die Kompetenzanforderungen für die Fähigkeiten des Analysierens und des Interpretierens exemplarisch dargestellt. Ausgewählt wurden die Jahrgangsstufen 6 und 9, um die Graduierung der Kompetenzanforderungen darstellen zu können. Die Kompetenzanforderungen teilen sich in allgemeine und in spezielle Ergebniserwartungen (general and specific learning outcomes) von Bildungsprozessen auf. Die ersteren beziehen sich auf Bildungsabschnitte, wohingegen die letzteren genau festlegen, was am Ende einer Jahrgangsstufe erreicht werden sollte.

#### **a) General learning outcomes by the end of grade 6**

##### **Analysing and interpreting**

*It is expected that students will...*

206

interpret findings from investigations using appropriate methods

#### **b) Specific Learning Outcomes (Grade 6)**

##### **Analysing and interpreting**

206-1

classify according to several attributes and create a chart or diagram that shows the method of classifying (e.g., classify organisms found in pond water using criteria they have developed themselves)

206-9

identify new questions or problems that arise from what was learned (e.g., identify questions such as “How can students from different parts of the country and around the world communicate effectively about animals and plants?”)

### **General learning outcomes by the end of grade 9**

#### **Analysing and interpreting**

*It is expected that students will...*

210

analyse qualitative and quantitative data and develop and assess possible explanations

### **b) Specific Learning Outcomes (Grade 9)**

#### **Analysing and interpreting**

210-4

predict the value of a variable by interpolating or extrapolating from graphical data (e.g., predict the time of ovulation from a graph of daily body temperatures)

210-6

interpret patterns and trends in data, and infer and explain relationships among the variables (e.g., suggest an explanation for trends in the optimum reproductive years of women)

210-8

apply given criteria for evaluating evidence and sources of information (e.g., consider the date of publication, the relevance, and the perspective of the author of an information source on reproductive technologies)

# Literaturverzeichnis

- Aldrich, R. (2000). Educational Standards in Historical Perspective. In: Goldstein, H. & Heath, A. (Eds.). Educational Standards. Proceedings of the British Academy, No. 102. Oxford: Oxford University Press. S. 39-67.
- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In: J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K. J. Tillmann & M. Weiß (Hrsg.), Pisa 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich. S. 69-140.
- Association of Language Testers in Europe (ALTE) (1998). ALTE Handbuch europäischer Sprachprüfungen und Prüfungsverfahren. Cambridge: The University of Cambridge Local Examinations Syndicate (UCLES).
- Babylonia (1999). Themenheft 1999/Nr. 1: European Language Portfolio. Hrsg. v. Rolf Schärer. Zu beziehen über <http://babylonia.romsem.unibas.ch>.
- Babylonia (2000). Themenheft 2000/Nr. 4: European Language Portfolio II. Hrsg. v. Rolf Schärer. Zu beziehen über <http://babylonia.romsem.unibas.ch>.
- Bachman, L. & Palmer, A. S. (1996). Language testing in practice. Oxford: Oxford University Press.
- Bachman, L. (1990). Fundamental Considerations in Language Testing. Oxford: Oxford University Press.
- Baumert, J. (2002b) Deutschland im internationalen Bildungsvergleich. In: N. Killius, J. Kluge & L. Reisch (Hrsg.). Die Zukunft der Bildung. Frankfurt a.M.
- Baumert, J., Artelt, C., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Schümer, G., Stanat, P., Tillmann, K.-J. & Weiß, M. (Hrsg.) (2002a). Pisa 2000 – Die Länder der Bundesrepublik Deutschland im Vergleich: Zusammenfassung zentraler Befunde. Opladen: Leske + Budrich.
- Baumert, J., Bos, W. & Lehmann, R. (Hrsg.) (2000a). TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 1: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit. Opladen: Leske + Budrich.
- Baumert, J., Bos, W. & Lehmann, R. (Hrsg.) (2000b). TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 2: Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe. Opladen: Leske + Budrich.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J. & Weiß, M. (Hrsg.). (2001). PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.
- Baumert, J., Lehmann, R., Lehrke, M. & al. (1997). TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich: deskriptive Befunde. Opladen: Leske + Budrich.
- Bausch, K.-R., Christ, H., Königs, F. G. & Krumm, H.-J. (Hrsg.) (2003). Der Gemeinsame europäische Referenzrahmen für Sprachen in der Diskussion. Tübingen: Narr.

- Beck, B. & Klieme, E. (2003, in Vorbereitung). DESI – Eine Längsschnittstudie zur Untersuchung des Sprachunterrichts in deutschen Schulen. In: Evaluation im Brennpunkt – Thema Fremdsprachen lernen und lehren. Landau: Verlag Empirische Pädagogik.
- Benner, Dietrich (2002). Die Struktur der Allgemeinbildung im Kerncurriculum moderner Bildungssysteme. In: Zeitschrift für Pädagogik 48 (2002), S. 68-88.
- Biehl, J., Hopmann, S. & Ohlhaber, F. (1996). Wie wirken Lehrpläne – Modelle, Strategien, Widersprüche. In: Pädagogik 48 (1996) 5, S. 33-37.
- Biehl, J., Ohlhaber, F. & Riquarts, K. (1999). Sekundäre Lehrplanbindungen: Vergleichende Untersuchungen zur Entstehung und Verwendung von Lehrplanentscheidungen. Endbericht zum DFG-Projekt. Kiel: IPN.
- Biehl, J., Ohlhaber, F., Riquarts, K. & Hopmann, S. (1998). Lehrplanarbeit in der Bundesrepublik Deutschland. In: R. Künzli & S. Hopmann (Hrsg.) Lehrpläne. Wie sie entwickelt werden und was von ihnen erwartet wird. Chur & Zürich, S. 277-296.
- Böttcher, W. & Kalb, P. E. (Hrsg.) (2002). Kerncurriculum. Was Kinder in der Grundschule lernen sollen. Weinheim und Basel: Beltz Verlag.
- Böttcher, W. (2002). Kann eine ökonomische Schule auch eine pädagogische sein? München: Juventa.
- Comparison of Education Systems in Selected Countries: Understanding the International Variation of PISA Results (2003; in Vorbereitung)
- Council of Europe. (2000). A Common European Framework of Reference for Languages: Learning, teaching, assessment. Cambridge: Cambridge University Press.
- DESI-Konsortium (2001). Deutsch-Englisch-Schülerleistungen International. Projektplan und Angebot an die Kultusministerkonferenz. Unveröffentlichtes Manuskript. Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.
- Deutscher Industrie- und Handelstag (Hrsg.) (2001). Arbeitsplatz Europa: Sprachkompetenz wird messbar. Berlin: DIHT.
- DIALANG (2001). Tests in 14 europäischen Sprachen mit Online-Diagnose [<http://www.dialang.org>].
- Europarat (Hrsg.). Europäisches Sprachenportfolio (ESP). Strasbourg: Europarat.
- Goethe-Institut/Inter-Nationes u.a. (Hrsg.) (2001). Gemeinsamer Europäischer Referenzrahmen für Sprachen: Lernen, Lehren, Beurteilen. München. [<http://www.goethe.de/referenzrahmen>].
- Goldstein, H. & Heath, A. (Eds.) (2000). Educational Standards. Proceedings of the British Academy, No. 102. Oxford: Oxford University Press.
- Häcker, H., Leutner, D. & Amelang, M. (Hrsg.) (1998). Standards für pädagogisches und psychologisches Testen. Supplementum 1/1998 der „Diagnostica“ und der „Zeitschrift für Differentielle und Diagnostische Psychologie“. Göttingen: Hogrefe.
- Hager, W., Patry, J.-L. & Brezing, H. (2000). Evaluation psychologischer Interventionsmaßnahmen. Standards und Kriterien: Ein Handbuch zur Qualitätssicherung. Bern: Huber.
- Helmke, A. & Jäger, R. S. (2003). Vergleichsarbeiten (VERA): eine Standortbestimmung zur Sicherung schulischer Kompetenzen. In: Schulverwaltung Hessen, Rheinland-Pfalz, Saarland (in Vorbereitung).

- Helmke, A. & Jäger, R. S. (Hrsg.). (2002). Die Studie MARKUS – Mathematik-Gesamterhebung Rheinland-Pfalz: Kompetenzen, Unterrichtsmerkmale, Schulkontext. Landau: Verlag Empirische Pädagogik.
- Hentig, H. v. (1996). Bildung. Ein Essay. München: Hanser.
- Hiebert, J. (1999). The relationship between research and the NCTM Standards. *Journal for Research in Mathematics Education*, 30, S. 3-19.
- International Technology Education Association (2000). Standards for Technological Literacy: Content for the Study of Technology. Reston, VA: ITEA.
- Kane, M. (1994). Validating the Performance Standards Associated With Passing Scores. *Review of Educational Research*, 64(3), S. 425-461.
- Klieme, E. (2000). Fachleistungen im voruniversitären Mathematik- und Physikunterricht. In: J. Baumert, W. Bos & R. Lehmann (Hrsg.), TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Band 2: Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe. Opladen: Leske + Budrich.
- Klieme, E., Baumert, J. & Schwippert, K. (2000). Schulbezogene Evaluation und Schulleistungsvergleiche – Eine Studie im Anschluss an TIMSS. In: H.G. Rolff, W. Bos, K. Klemm, K. Pfeiffer & R. Schulz-Zander (Hrsg.). *Jahrbuch der Schulentwicklung*, Band 11. München: Juventa.
- Klieme, E., Neubrand, M. & Lüdtke, O. (2001). Mathematische Grundbildung: Testkonzeption und Ergebnisse. In: Deutsches PISA-Konsortium (Hrsg.), PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich. S. 139 bis 190.
- Kluge, F. (1999). *Etymologisches Wörterbuch der deutschen Sprache*. Berlin, New York: Gruyter.
- Knoche, N., Lind, D., Blum, W., Cohors-Fresenborg, E., Flade, L., Löding, W., Möller, G., Neubrand, M. & Wynands, A. (Deutsche PISA-Expertengruppe Mathematik) (2000). Die PISA-2000-Studie, einige Ergebnisse und Analysen. In: L. Hefendehl-Hebeker, G. Walther & B. Wollring (Hrsg.). *Journal für Mathematik-Didaktik*. 23 (2002) Heft 3/4. S. 159-202.
- Kultusministerkonferenz (2002a). Einheitlich Prüfungsanforderungen in der Abiturprüfung Englisch. Bonn: KMK (Beschluss der KMK vom 1. Dezember 1989 in der Fassung von 24. Mai 2002).
- Kultusministerkonferenz (2002b). Zwischenbericht der länderoffenen Arbeitsgruppe „Europäisches Portfolio der Sprachen“ (vom 17.12.2002). Bonn: KMK.
- Landesinstitut für Schule und Weiterbildung (Hrsg.) (2000). *Europäisches Portfolio der Sprachen*. Soest: LSW.
- Lehmann, R. H., Gänsfuß, R. & Peek, R. (1999). Ergebnisse der Erhebung von Aspekten der Lernausgangslage und der Lernentwicklung von Schülerinnen und Schülern an Hamburger Schulen – Klassenstufe 7. Bericht über die Untersuchung im September 1998. Hamburg.
- Lehmann, R. H., Peek, R. & Gänsfuß, R. (1997). Aspekte der Lernausgangslage von Schülerinnen und Schülern der fünften Klassen an Hamburger Schulen. Bericht über die Untersuchung im September 1996. Hamburg: Behörde für Schule, Jugend und Berufsausbildung., Amt für Schule.
- Linnakylä, in Vorbereitung

- Martin, W. G. (in Vorbereitung). Are the NCTM Standards Working? Perspectives on Student Learning as a Point of Impact in Research on Standards.
- McLaughlin, M. & Shepard, L. A. (1995). Improving Education through Standards-Based Reform: A Report by the National Academy of Education Panel on Standards-Based Education Reform. Stanford, CA: National Academy of Education.
- McLeod, D. B., Stake, R. E., Schappelle, B., Mellissinos, M. & Gierl, M. J. (1996). Setting the Standards: NCTM's role in the reform of mathematics education. In: S. A. Raizen & E. D. Britton (Eds.). Bold ventures: U.S. innovations in science and mathematics education. Vol 3: Cases in mathematics education (pp. 13-132). Dordrecht, The Netherlands: Kluwer.
- Milanovic, M. (2001). The Association of Language Testers in Europe (ALTE) – Working to wards a Framework of European Language Examinations. Fremdsprachen Lehren und Lernen. 30, S. 28-45.
- National Council of Teachers of Mathematics (1980). An agenda for action: Recommendations for school mathematics. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (1991). Professional teaching standards for school mathematics. Reston, VA: Author.
- National Council of Teachers of Mathematics (1995). Assessment standards for school mathematics. Reston, VA: NCTM.
- National Council of Teachers of Mathematics (2000). Principles and standards for school mathematics. Reston, VA: NCTM.
- National Research Council (1996). National Science Education Standards. Washington, DC: National Academy Press.
- National Research Council (1999). Testing, Teaching, and Learning. Washington, DC: NRC.
- National Research Council (2001). Knowing what students know – The science and design of educational assessment. Washington, DC: NRC.
- Neubrand, M., Klieme, E., Lüdtke, O. & Neubrand, J. (2002). Kompetenzstufen und Schwierigkeitsmodelle für den PISA-Test zur mathematischen Grundbildung. In Unterrichtswissenschaft 2, S. 100-119.
- Nold, G. & Schröder, K. (in Vorbereitung). Zum Testen kommunikativer Kompetenz im Englischen bei Schülern der 9. Jahrgangsstufe. Frankfurt a.M.: DIPF.
- North, B. (2000). The Development of a Common Framework Scale of Language Proficiency. Bern/Frankfurt/New York: Lang. (Theoretical Studies in Second Language Acquisition, 8).
- OECD (1991). Schulen und Qualität: ein internationaler OECD-Bericht. Frankfurt a.M.: Lang. (Englisches Original 1989).
- OECD (2001). Bildung auf einen Blick. OECD-Indikatoren 2001. Zentrum für Forschung und Innovation im Bildungswesen. Paris.
- Oser, F. & Oelkers, J. (Hrsg.) (2001). Die Wirksamkeit der Lehrerbildungssysteme. Von der Allrounderbildung zur Ausbildung professioneller Standards. Chur/Zürich: Rüegger.

- Prenzel, M. (2000). Steigerung der Effizienz des mathematisch-naturwissenschaftlichen Unterrichts: Ein Modellprogramm von Bund und Ländern. *Unterrichtswissenschaften*, 28 (2). S. 103 bis 126.
- Prenzel, M., Rost, J., Senkbeil, M., Häußler, P. & Klopp, A. (2001). Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In: J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.). *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich. S. 192-250.
- Projektgruppe TestDaF (2000). *TestDaF: Konzeption, Stand der Entwicklung, Perspektiven*. *Zeitschrift für Fremdsprachenforschung*, 11(1), S. 63-82.
- Quetz, J. (2003). Der Gemeinsame europäische Referenzrahmen: Ein Schatzkästlein mit Perlen, aber auch mit Kreuzen und Ketten. In: K.-R. Bausch u.a. (Hrsg.), S. 145-155.
- Ravitch, D. (1995). *National Standards in American Education*. Washington, D.C.: Brookings Institution Press.
- Rost, J. (1996). *Lehrbuch Testtheorie Testkonstruktion*. Bern: Huber.
- Schneider, G. & North, B. (2001). Fremdsprachen können – was heißt das? Skalen zur Beschreibung, Beurteilung und Selbsteinschätzung der fremdsprachlichen Kommunikationsfähigkeit. Chur/Zürich: Rüegger.
- Schneider, G. (2001). Kompetenzbeschreibungen für das „Europäische Sprachenportfolio“. *Fremdsprachen Lehren und Lernen*. 30, S. 193-214.
- Skolverket – National Agency for Education (Hrsg.) (2001). *Languages. Syllabuses, Grading Criteria and Comments: Gy 2000*. Stockholm: Fritzes.
- Stanat, P. u.a. (Hrsg.) (2002). *Rückmeldung der PISA 2000-Ergebnisse an die beteiligten Schulen*. Berlin.
- Swanson, Ch. B. & Stevenson, D. L. (2002). Standards-based reform in practice: Evidence on state policy and classroom instruction from the NAEP state assessment. *Educational Evaluation and Policy Analysis*, 24(1), 1-27.
- Tate, W. F. (in Vorbereitung). What is a standard?
- Tenorth, H. E. (Hrsg.) (2001). *Kerncurriculum Oberstufe. Mathematik – Deutsch – Englisch. Experten im Auftrag der Ständigen Konferenz der Kultusminister*. Weinheim und Basel: Beltz Verlag.
- Tenorth, H.-E. (1994): „Alle alles zu lehren“. *Möglichkeiten und Perspektiven allgemeiner Bildung*. Darmstadt: Wiss. Buchges. (WB-Forum; 82).
- Terhart, E. (Hrsg.) (2000). *Perspektiven der Lehrerbildung in Deutschland. Abschlussbericht der von der Kultusministerkonferenz eingesetzten Kommission*. Weinheim und Basel: Beltz Verlag.
- TESTDAF. Information über die aktuelle Fassung und über Bezugsquellen auf der Website <http://www.testdaf.de>
- Thüringer Kultusministerium (2002). *Europäisches Sprachenportfolio*. Erfurt: KM Thüringen.
- Thürmann, E. (Hrsg.) (2003). *Online Arbeitsbereich zum Europäischen Portfolio der Sprachen*. Soest: Landesinstitut für Schule und Weiterbildung. In: [www.learnline.nrw.de/angebote/portfolio](http://www.learnline.nrw.de/angebote/portfolio).

- Vollmer, H. J. (2003). Ein gemeinsamer europäischer Referenzrahmen für Sprachen: Nicht mehr, nicht weniger! In: K.-R. Bausch, H. Christ, F. G. Königs & H.-J. Krumm (Hrsg.). Der Gemeinsame europäische Referenzrahmen für Sprachen in der Diskussion. Tübingen: Narr. S. 192 bis 206.
- Vollstädt, W., Tillmann, K.-J., Rauin, U., Höhmann, K. & Tebrügge, A. (1999). Lehrpläne im Schulalltag. Eine empirische Studie zur Akzeptanz und Wirkung von Lehrplänen in der Sekundarstufe I. Opladen: Leske und Budrich.
- Weinert, F. E. (1999). Concepts of competence (Contribution within the OECD project Definition and selection of competencies: Theoretical and conceptual foundations (DeSeCo)). Neuchâtel: DeSeCo.
- Weinert, F. E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), Leistungsmessungen in Schulen. Weinheim und Basel: Beltz Verlag, S. 17-31.
- Weiterbildungs-Testsysteme (WBT) (Hrsg.) (1998). The European Language Certificates, Certificate in English: Learning Objectives and Test Format. Frankfurt a.M.: WBT.
- Weiterbildungs-Testsysteme (WBT) (Hrsg.) (2001). START English: Learning Objectives and Test Format. Frankfurt a.M.: WBT.
- William, D. (1996). Meanings and Consequences in Standard Setting. *Assessment in Education*, 3(3), S. 287-307.