

10: Dietrich Borne (Hrsg.):  
Bildungsstandards, Sönningh 2007

Olaf Köller

BILDUNGSSTANDARDS,  
EINHEITLICHE PRÜFUNGSANFORDERUNGEN  
UND QUALITÄTSSICHERUNG IN DER SEKUNDARSTUFE II

Mit ihren Beschlüssen von 2003 und 2004 hat die Kultusministerkonferenz (KMK) verbindliche länderübergreifende Bildungsstandards für die Grundschule und das Ende der Sekundarstufe I verabschiedet. Die Einführung der Standards und der damit verbundene politische Wille, deren Erreichung regelmäßig zu überprüfen, stellen einen Meilenstein in der Umsteuerung des allgemein bildenden Schulwesens dar. Dominierte bis in die 1990er Jahre die Inputsteuerung, bei der das Bildungssystem in erster Linie durch Lehrplan- und Strukturreformen gesteuert wurde, so ist als Konsequenz der *Dritten Internationalen Mathematik- und Naturwissenschaftsstudie* (TIMSS; vgl. Baumert/Lehmann u. a. 1997; Baumert/Bos/Lehmann 2000a, 2000b) und des *Programme for International Student Assessment* (PISA; Deutsches PISA-Konsortium 2001, 2002, 2004, 2005) die Outputsteuerung hinzugekommen, bei der eine systematische Überprüfung der Bildungserträge auf Seiten der Schülerinnen und Schüler Steuerungswissen auf der Systemebene generieren soll. Folgerichtig hat sich die KMK mit ihren Plöner Beschlüssen vom Juni 2006 auf eine Gesamtstrategie der Qualitätssicherung im Primarbereich und in der Sekundarstufe I verständigt. Qualitätsindikatoren des Bildungssystems sollen die von Schülerinnen und Schülern erreichten Kompetenzen sein, welche in festen Zeitabständen in der dritten Jahrgangsstufe und am Ende der Sekundarstufe I überprüft werden (zu Details s. KMK 2006).

Für die gymnasiale Oberstufe hat die KMK zwischen 2002 und 2004 eine Überarbeitung der Einheitlichen Prüfungsanforderungen (EPAs) vorgenommen. Die revidierten EPAs folgen wie die Bildungsstandards einem stärker funktionalistischen Bildungskonzept und legen besonderen Wert auf die Erreichung von Grundkompetenzen.

Im vorliegenden Beitrag soll argumentiert werden, dass Bildungsstandards und EPAs geeignete Grundlagen für die Steuerung im Bildungssystem darstellen, da sich aus ihnen wichtige Implikationen im Hinblick auf Unterrichtsinhalte ergeben, sie aber auch als Grundlage für die Entwicklung von Testinstrumenten zum Systemmonitoring dienen können. Es soll dabei anhand ausgewählter Fächer (Englisch und Biologie) verdeutlicht werden, dass sich Standards und EPAs chronologisch ergänzen. Anhand der Darstellung zentraler Befunde von *Large-scale Assessments* in gymnasialen Oberstufen soll anschließend abgeleitet werden, dass das vor kurzem von der KMK verabschie-

deten System der Qualitätssicherung im Bildungssystem auf die Sekundarstufe II ausgedehnt werden sollte. Verzichtet wird auf die erneute bildungstheoretische Verankerung der Bildungsstandards. Dies ist bezogen auf ihrer Bedeutung für vertiefte Allgemeinbildung in der so genannten Klieme-Expertise (Klieme u. a. 2003) geschehen. Die explizite Verortung in der fachspezifischen Bildungskonzeption ist bislang in erster Linie für die Mathematik (vgl. Blum/Drüke-Noe/Hartung/Köller 2006) und die Naturwissenschaften (im Überblick Tenorth 2004) überzeugend vorgenommen worden.

Begonnen werden soll allerdings mit der Vorgeschichte, welche überhaupt erst zur Etablierung der Bildungsstandards und einer Outputsteuerung geführt hat.

## 1. Die Vorgeschichte: Mediokres Abschneiden in internationalen Schulleistungsstudien

In Deutschland fehlte bis zu Beginn der 1990er Jahre eine systematische und regelmäßige Überprüfung von Erträgen schulischer Bildungsprozesse, wie dies etwa in den USA durch Programme wie NAEP (*National Assessment of Educational Progress*) üblich war und ist. Ein Hauptinteresse lag bis dahin in der Entwicklung und Erprobung von Modellen zur Optimierung der Arbeit in Einzelschulen und dem Entwurf didaktischer Modelle und deren Einführung in die Unterrichtspraxis (Input- und Prozess-Orientierung). Die Vergewisserung über das im Unterricht Erreichte (im Sinne einer Output- oder Outcome-Orientierung) trat demgegenüber in den Hintergrund. In dieser Situation waren Untersuchungen wie die internationale Lesestudie (*International Reading Literacy Study*, Elley, 1994) oder die dritte internationale Mathematik- und Naturwissenschaftsstudie (TIMSS; vgl. Baumert/Lehmann u. a. 1997; Baumert u. a. 2000a, 2000b), die auf einer breiten empirischen Basis die Beschreibung und Analyse der Erträge fachlichen Lernens in den Mittelpunkt rückten, unzeitgemäße Vorhaben. Umso beeindruckender waren die öffentlichen Reaktionen, welche insbesondere die Befunde von TIMSS auslösten. Die Bereitstellung von Basisinformationen über Ertragslagen deutscher Schulen stürzte das Bildungssystem auf Grund der eher mäßigen Leistungen unserer Schülerinnen und Schüler in die Krise. In Folge von TIMSS kam es mit den Konstanzer Beschlüssen der KMK von 1997 zur empirischen Wende in der Erziehungswissenschaft, die pädagogische Psychologie und die Psychometrie gewannen im Kontext Bildungsforschung an Bedeutung und weitere große Schulleistungsstudien auf regionaler, nationaler und internationaler Ebene wurden initiiert. Reflexionen über die Messbarkeit von Bildungserträgen traten in den Hintergrund zugunsten einer festen Überzeugung, dass fachliche Kompetenzen, sofern sie im Sinne von Leistungsstandards beschrieben werden, mess- und ü-

berprüfbar seien. Im Vordergrund stand jetzt die Frage, welche konkreten Leistungsniveaus Schülerinnen und Schülern erreichen (Outcome-Orientierung). Den ersten Höhepunkt dieser neuen Entwicklung stellte PISA 2000 (Deutsches PISA-Konsortium 2001, 2002) dar. Das wiederum unbefriedigende Abschneiden deutscher Jugendlicher löste neue Evaluationsmaßnahmen aus. In den Ländern wurden daraufhin Programme für flächendeckende Vergleichsarbeiten in verschiedenen Jahrgangsstufen und Fächern aufgelegt und auf Seiten der KMK wurde die Arbeit an den Bildungsstandards für die Grundschule und die Sekundarstufe I begonnen, parallel wurde die Revision der EPAs vorangetrieben. Diese auf den Weg gebrachten Arbeiten konnten natürlich nicht das mediokre Abschneiden deutscher 15jähriger in PISA 2003 (Deutsches PISA-Konsortium 2004, 2005, 2006) verhindern, da die Bildungsstandards erst Ende 2003 bzw. 2004 und die EPAs zwischen 2002 und 2004 verabschiedet wurden und ihre positiven Wirkungen eher langfristig sichtbar sein werden.

## 2. Bildungsstandards und Einheitliche Prüfungsanforderungen als Steuerungsinstrumente

Mit ihren Beschlüssen von 2003 und 2004 hat die KMK Bildungsstandards für die Grundschule und die Sekundarstufe I verabschiedet. Alle 16 Länder haben sich dazu verpflichtet Sorge zu tragen, dass im Fachunterricht Lerngelegenheiten geschaffen werden, um die in den Standards formulierten Kompetenzen zu erreichen (Implementation der Bildungsstandards). Die Tabelle 1 zeigt im Überblick die berücksichtigten Fächer und Abschlüsse, für die Standards vorliegen.

In ihrer Expertise zur Entwicklung von Bildungsstandards in Deutschland haben Klieme u. a. (2003) ausgeführt, wodurch sich Standards auszeichnen sollen. Danach benennen die Standards fachspezifisch die wesentlichen Ziele pädagogischer Arbeit, ausgedrückt als erwünschte Lernergebnisse bzw. Kompetenzen der Schülerinnen und Schüler zu bestimmten Zeitpunkten ihrer Bildungsbiographien. Konkret bedeutet dies, dass Standards als *Can do Statements* formuliert sind. Beispielsweise heißt es in den Standards für den Mittleren Abschluss im Fach Mathematik im Kompetenzbereich Modellieren: „den Bereich oder die Situation, die modelliert werden soll, in mathematische Begriffe, Strukturen und Relationen übersetzen“ (KMK 2004a, S. 12). Im Bereich der ersten Fremdsprache für den Mittleren Schulabschluss heißt es für den Kompetenzbereich Lesen u. a.:

- „Die Schülerinnen und Schüler können
- Korrespondenz lesen, die sich auf das eigene Interessengebiet bezieht, und die wesentliche Aussage erfassen,

– klar formulierte Anleitungen, unkomplizierte Anleitungen, Hinweise und Vorschriften verstehen (B1/B2) ...“ (KMK 2004b, S. 14).

Kompetenzen sind hier als gezeigtes Verhalten definiert. Folgt man einem stärker psychologisch geprägten Kompetenzkonzept (z. B. Weinert 2001), so wird man nicht das Verhalten selbst, sondern die bei der Schülerin/beim Schüler verfügbaren oder von ihm erlernbaren kognitiven Fähigkeiten und Fertigkeiten, die notwendig sind, um bestimmte Probleme bzw. Aufgaben lösen zu können, als Kompetenzen verstehen. In diesem Fall gewinnen Kompetenzen den Charakter psychologischer Konstrukte, welche mit Hilfe von Messinstrumenten operationalisiert werden können.

Tabelle 1: Überblick über die Fächer und Abschlüsse, für die Bildungsstandards vorliegen

	Ende der Sekundarstufe I	
	Grundschule Ende der 4. Jahrgangsstufe	Hauptschulabschluss Mittlerer Schulabschluss
Deutsch		
Mathematik		
Erste Fremdsprache		
- Englisch		
- Französisch		
Naturwissenschaften		
- Biologie		
- Chemie		
- Physik		

Mit der Fokussierung auf Schülerkompetenzen findet eine klare Abgrenzung von Lehrplänen statt, die stärker Unterrichtsinhalte präzisieren (Input-Orientierung). Hierin liegt die Innovationskraft der Bildungsstandards, indem sie auf Seiten der Lehrkräfte das Bewusstsein schärfen sollen, das professionelle Agieren im Unterricht nicht an Inhalten, sondern am Kompetenzaufbau zu orientieren. Damit wird auch die Funktion der Standards als Steuerungsinstrument für Unterricht sichtbar: Fokussierung auf Kompetenzen impliziert einen handlungsorientierten und kognitiv aktivierenden Unterricht, wie er von den Fachdidaktiken seit einiger Zeit gefordert wird (vgl. z. B. Baumert u. a. 2004).

### 2.1 Einheitliche Prüfungsanforderungen

In ihren Beschlüssen von 2002 hat die KMK die Einheitlichen Prüfungsanforderungen für das Abitur (EPAs) revidiert, um auch zukünftig die Qualität der

allgemeinen Hochschulreife als schulische Abschlussqualifikation zu sichern und die allgemeine Studierfähigkeit sowie der Übergang in eine berufliche Ausbildung zu gewährleisten. Die EPAs sollen ein zwischen den Ländern vergleichbares Orientierungswissen, fachliche Standards und Grundkompetenzen sichern, die über den Auftrag des jeweiligen Faches hinausgehen. Die EPAs stellen dabei explizit länderübergreifende Standards für die Abiturprüfung dar, indem sie konkrete Lern- und Prüfungsbereiche beschreiben und auch kursniveauspezifisch Vorgaben für die Bewertung von Abiturleistungen machen. „Als für die Studierfähigkeit grundlegende Kompetenzen werden die sprachliche Ausdrucksfähigkeit, insbesondere die schriftliche Darlegung eines knappen Gedankengangs, das verständige Lesen komplexer fremdsprachlicher Sachtexte sowie der sichere Umgang mit mathematischen Symbolen und Modellen hervorgehoben.“ (KMK 2002, S. 1). Darüber hinaus wird immer wieder die wissenschaftspropädeutische Funktion der gymnasialen Oberstufe betont, mit der Konsequenz, dass selbstständiges Lernen, wissenschaftsorientiertes Arbeiten sowie die Entwicklung der Kommunikationsfähigkeit und der Kooperationsbereitschaft als Kernziele genannt werden.

Beibehalten wird die Forderung, dass die gymnasiale Oberstufe den Zugang zu vertiefter Allgemeinbildung ermöglichen muss. Wiewohl die EPAs gerade mit dieser letzten Forderung über die stärker auf Basiskompetenzen fokussierten Bildungsstandards hinausgehen und zudem sehr viel umfangreichere inhaltliche Vorgaben für den Unterricht in der Oberstufe machen, soll im Folgenden anhand zweier Fächer (Englisch und Biologie) illustriert werden, dass Standards und EPAs ineinandergreifen, kompetenzorientiert sind und damit Basis einer Qualitätssicherung sein können.

### 2.2 Bildungsstandards und EPAs in Englisch

Im Bereich der ersten Fremdsprache (in der Regel Englisch) werden in den Bildungsstandards der KMK für die Sekundarstufe I vier zentrale Kompetenzbereiche genannt, die sich an dem *Gemeinsamen Europäischen Referenzrahmen für Sprachen* (GER; Europarat 2001) orientieren (vgl. Tabelle 2): (1) kommunikative Fertigkeiten, (2) Verfügung über sprachliche Mittel, (3) interkulturelle und (4) methodische Kompetenzen. Kommunikative Fertigkeiten umfassen die fünf Bereiche Leseverstehen, Hör- und Hör-/Sehverstehen, Sprechen (in den Formen an Gesprächen teilnehmen und zusammenhängendes Sprechen), Schreiben sowie Sprachmittlung. Kommunikative Fertigkeiten sind gebunden an die Verfügung über die sprachlichen Mittel Wortschatz, Grammatik, Aussprache und Intonation sowie Orthografie. Zur Beschreibung von Bildungsstandards für diesen Bereich werden GER-Skalen für linguistische Kompetenzen herangezogen.

Im Bereich der interkulturellen Kompetenzen werden vor allem Haltungen erfasst, die „ihren Ausdruck gleichermaßen im Denken, Fühlen und Handeln und ihre Verankerung in entsprechenden Lebenserfahrungen und ethischen

Prinzipien haben“ (vgl. etwa KMK 2004b, S. 16). Sie sind damit einerseits den kommunikativen Kompetenzen des GER (im Sinne soziolinguistischer Kompetenzen) zuzuordnen, andererseits werden durch sie Aspekte der allgemeinen Kompetenzen des GER zum Ausdruck gebracht. Schließlich greifen methodische Kompetenzen Teilaspekte allgemeiner sowie kommunikativer Kompetenzen (pragmatische Kompetenzen) auf.

Tabelle 2: Kompetenzbereiche in den Bildungsstandards für die erste Fremdsprache

<i>Funktionale Kommunikative Kompetenzen</i>	
<i>Kommunikative Fertigkeiten</i>	<i>Verfügung über die sprachlichen Mittel</i>
Leseverstehen Hör- und Hör-/Sehverstehen Sprechen An Gesprächen teilnehmen Zusammenhängendes Sprechen Schreiben Sprachmittlung	Wortschatz Grammatik Aussprache und Intonation Orthographie
<i>Interkulturelle Kompetenzen</i>	
Soziokulturelles Orientierungswissen Verständnisvoller Umgang mit kultureller Differenz Praktische Bewältigung interkultureller Begegnungssituationen	
<i>Methodische Kompetenzen</i>	
Textrezeption (Leseverstehen und Hörverstehen) Interaktion Textproduktion (Sprechen und Schreiben) Lernstrategien Präsentation und Mediennutzung Lernbewusstheit und Lernorganisation	

Im GER werden drei Grundniveaus – A, B und C – unterschieden, die in je zwei Teilniveaus aufgegliedert sind. Die Niveaus A1 und A2 stehen für elementare Sprachverwendung; die Niveaus B1 und B2 für selbstständige Sprachverwendung und die Niveaus C1 und C2 für eine kompetente Sprachverwendung. Die KMK-Standards gehen davon aus, dass mit Erreichen des Hauptschulabschlusses im Mittel Leistungen auf GER-Niveau A2, mit Erreichen des Mittleren Schulabschlusses im Mittel Leistungen auf GER-Niveau B1, teilweise auf B2, erbracht werden können.

Die EPAs orientieren sich ebenfalls am GER und definieren Kompetenzbereiche, die weitgehend identisch zu den Bildungsstandards sind. Wiederum stehen funktionale kommunikative Kompetenzen, interkulturelle Kompetenzen und Methodenkompetenzen im Vordergrund. In Abgrenzung von den Bildungsstandards (in denen die Bereiche zusammengefasst sind) wird auf Seiten der methodischen Kompetenzen zwischen (1) Umgang mit Texten und Medien und (2) den eigentlichen methodischen Kompetenzen unterschieden. Mit

Bezug auf den GER werden Leistungserwartungen auf den Niveaustufen B2 und C1 definiert.

*Cum grano salis* definieren die EPAs also für das Fach Englisch analoge Kompetenzen wie die Bildungsstandards, allerdings auf einem höheren Kompetenzniveau.

### 2.3 Standards und EPAs in Biologie

Im Fach Biologie werden in den Bildungsstandards – wie auch in Physik und Chemie – vier Kompetenzbereiche unterschieden: (1) Fachwissen, (2) Erkenntnisgewinnung, (3) Kommunikation und (4) Bewertung. Die Tabelle 3 präzisiert diese vier Bereiche ansatzweise:

Tabelle 3: Kompetenzbereiche in den Bildungsstandards für Biologie

<i>Kompetenzbereich</i>	<i>Beschreibung</i>
Fachwissen	Lebewesen, biologische Phänomene, Begriffe, Prinzipien, Fakten kennen und den Basiskonzepten zuordnen
Erkenntnisgewinnung	Beobachten, Vergleichen, Experimentieren, Modelle nutzen und Arbeitstechniken anwenden
Kommunikation	Informationen sach- und fachbezogen erschließen und austauschen
Bewertung	Biologische Sachverhalte in verschiedenen Kontexten erkennen und bewerten

Die Standards lehnen sich dabei eng an die internationale Diskussion um naturwissenschaftliche Grundbildung an (*Scientific Literacy*; vgl. Harms/ Mayer/Hamann/Bayrhuber/Kattmann 2004, AAAS, 1993). *Scientific Literacy* zeichnet sich danach durch folgende Eigenschaften aus:

- Vertrautheit mit der natürlichen Welt und Kenntnis ihrer Verschiedenheit und Einheit,
- Verständnis zentraler naturwissenschaftlicher Konzepte und Prinzipien,
- Kenntnis des Zusammenhangs zwischen Naturwissenschaften, Mathematik und Technik,
- Erkenntnistheoretische Vorstellungen von der konstruktiven Natur der Mathematik und der Naturwissenschaften sowie Kenntnis ihrer Stärken und Grenzen,
- Verständnis der Grundzüge naturwissenschaftlichen Denkens und
- Anwendung von naturwissenschaftlichem Wissen auf Sachverhalte des individuellen und sozialen Lernens.

Die Besinnung auf das Literacy-Konzept wird in der angelsächsischen Diskussion mit dem Argument neuer Qualifikationsanforderungen gerechtfertigt, die sich aus dem raschen Wandel von der Industrie- zur Wissensgesellschaft ergeben.

In ihrer Fassung vom Februar 2004 betonen die EPAs für die Biologie die Rolle des Faches beim Zugang zu vertiefter Allgemeinbildung sowie die wissenschaftspropädeutische Rolle (vgl. hierzu auch Harms u. a. 2004), ohne allerdings den Grundbildungsgedanken aufzugeben: „Die biologische Grundbildung der Schülerinnen und Schüler dient daher auch als Basis für vielfältige Diskussionen in gesellschaftswissenschaftlichen Fächern; sie ist ein unentbehrlicher Bestandteil der Allgemeinbildung und der Lebensplanung für Schülerinnen und Schüler“ (KMK 2004c, S. 2). Ganz analog zu den Standards werden vier Kernkompetenzen formuliert, (1) Fachwissen, (2) Methoden, (3) Kommunikation und (4) Reflexion, die in ihrer wissenschaftspropädeutischen Funktion präzisiert werden. Wie in den Standards wird beim Kompetenzkonzept auf Weinert (2001) Bezug genommen und die Kompetenzerwartungen werden als „*Can do Statements*“ formuliert, wobei eine Ausdifferenzierung nach Grund- und Leistungskursen stattfindet. Damit sind die Grundlagen für entsprechende Aufgaben gelegt und in den EPAs werden auf dieser Basis diverse Abituraufgaben präsentiert. Anstelle der Prüfungsaufgaben können selbstverständlich auch Testaufgaben abgeleitet werden.

Zusammenfassend gilt also für die Biologie wie auch für Englisch (s. o.), dass EPAs in ihrer Grundkonzeption wie Standards kompetenzorientiert formuliert sind, den Unterricht inhaltlich steuern und Basis für eine Qualitätssicherung im allgemein bildenden Schulsystem sein können.

### 3. Bildungsstandards im System der Qualitätssicherung

Bildungsstandards und EPAs können als Instrumente einer Output-Steuerung im Bildungssystem nur wirksam werden, wenn sie auf konkrete Aufgaben bzw. Testitems herunter gebrochen werden, die reliabel und valide die formulierten Kompetenzerwartungen erfassen. Solche Items müssen die Grundlage von nationalen Skalen zu den verschiedenen Kompetenzen sein. Gleichzeitig muss die fachdidaktische und kognitionspsychologische Analyse der Items, gekoppelt mit den curricularen und standardbezogenen Erwartungen die Basis für das so genannte „Standard Setting“ (vgl. Hambleton/Jaeger/Plake/Mills 2000) sein, bei dem Experten auf der Basis der inhaltlichen Analyse der Items und der Verteilung der Personenleistungen konsensuell zu einem Kompetenzmodell mit unterschiedlichen Niveaustufen kommen.

Aus einer psychometrischen Perspektive bedeutet dies, dass die Kompetenzen zunächst inhaltlich präzisiert und dann operationalisiert werden. Dies geschieht aktuell lediglich für die Bildungsstandards, nicht aber die EPAs. Für die Operationalisierung der Bildungsstandards im Fach Mathematik in der Sekundarstufe I geben Ehmke, Leiß, Blum und Prenzel (2006) einen guten Überblick. Die Mathematikaufgaben wurden von Lehrkräften, welche von Fachdidaktikern unterstützt wurden, entwickelt. Weitere Fachdidaktiker und

Bildungsforscher optimierten diese Aufgaben, die dann in einem großen Feldtest pilotiert wurden. Die Normierung der Testitems erfolgte schließlich im Rahmen der PISA-2006-Erhebung. Die rund 300 Aufgaben dieser Normierung wurden dabei einer repräsentativen Stichprobe von deutschen Neuntklässlern vorgegeben. Auf der Basis dieser Daten wird es möglich sein, nationale Skalen für unterschiedliche mathematische Kompetenzbereiche bereitzustellen und auf diesen Skalen Leistungsbereiche zu definieren, die Mindest-, Regel- und Idealstandards abbilden. Die letztendliche Festlegung, über welche Fähigkeiten Schülerinnen und Schüler verfügen müssen, um die Bildungsstandards zu erreichen, wird also erst auf der Basis der empirischen Befunde möglich sein.

Die weiteren Schritte in einem System der Qualitätssicherung hat die KMK in ihren „Plöner Beschlüssen“ vom Juni 2006 vorgelegt. Darin ist vorgesehen, dass das nationale Bildungsmonitoring zukünftig auf der Grundlage der Bildungsstandards durchgeführt werden soll. Auf der Basis von repräsentativen Stichproben soll in den Ländern festgestellt werden, welche Anteile der Schülerpopulation die Standards erreichen bzw. überschreiten. Die Befunde sollen dann eine relevante Grundlage für Steuerungsmaßnahmen im System sein.

Die Plöner Beschlüsse sehen im Grundschulbereich vor, dass die Überprüfung der Bildungsstandards im Fünf-Jahres-Rhythmus in der 3. Jahrgangsstufe geschehen soll, und zwar zeitlich gekoppelt an die Internationale Grundschulstudie (IGLU/PIRLS; vgl. Bos u. a. 2003), welche ebenfalls einem Fünf-Jahres-Rhythmus folgt. Letztmalig in IGLU/PIRLS 2006 wird der Ländervergleich auf der Basis internationaler Instrumente durchgeführt. Erstmals in 2011 mit der nächsten IGLU/PIRLS- und TIMSS-Erhebung wird es dann auf der Basis der Bildungsstandards zu einem Ländervergleich in den Fächern Deutsch und Mathematik kommen. Dies wiederholt sich dann passend zum IGLU/PIRLS-Rhythmus alle fünf Jahre.

Im Bereich der Sekundarstufe I sehen die Plöner-Beschlüsse vor, dass letztmalig 2006 der nationale Vergleich zwischen den Ländern auf der Basis der PISA-Instrumente erfolgt. Ab PISA 2009 wird dann der nationale Vergleich zwischen den Ländern auf Grundlage der Bildungsstandards durchgeführt, in 2009 für die Sprachen Deutsch, Englisch und Französisch, 2012 folgt der Vergleich in Mathematik und den Naturwissenschaften, 2015 folgen wieder die Sprachen, 2018 Mathematik und die Naturwissenschaften. Die Ländervergleiche in der Sekundarstufe I passen sich damit einem sechsjährigen Rhythmus an, getrennt für die Sprachen und Mathematik und die Naturwissenschaften. Berücksichtigt werden die 8. Jahrgangsstufen (Bildungsgänge, die zum Hauptschulabschluss führen) und die 9. Jahrgangsstufen (Bildungsgänge, die zum Mittleren Schulabschluss führen).

Im Grundschul- und Sekundarbereich wird damit eine Strategie realisiert, welche das nationale Bildungsmonitoring explizit an die vorgegebenen Lernziele koppelt und Schulpolitik, Schuladministration, Lehrkräfte und Eltern über die Erträge in den verschiedenen Länder der Bundesrepublik Deutschland

regelmäßig informiert. In wie weit es in Folge der regelmäßigen Ländervergleiche zu einer Angleichung der Leistungsniveaus kommen wird, ist eine offene Frage.

Für die Sekundarstufe II fehlt solch ein System der Qualitätssicherung. Auf der Basis ausgewählter Befunde aus Schulleistungsstudien soll daher im Folgenden argumentiert werden, dass der Aufbau eines entsprechenden Systems in der gymnasialen Oberstufe notwendig ist, um schulform- und länderübergreifende Leistungsstandards bei den Abiturleistungen sicherzustellen.

#### 4. Befunde aus zwei Schulleistungsstudien in der Sekundarstufe II

Im Gegensatz zur Primarstufe und Sekundarstufe I beteiligt sich Deutschland nicht regelmäßig an internationalen Schulleistungsvergleichen in der gymnasialen Oberstufe, auch liegt bislang keine Strategie der Qualitätssicherung bzw. -überprüfung vor. Dies ist insofern überraschend, als unter den Gegebenheiten zulassungsbeschränkter Studiengänge der Vergleichbarkeit der Abiturleistungen (der Abiturdurchschnittsnote) eine hohe Bedeutung zukommt. Verteilungsgerechtigkeit innerhalb aber auch zwischen den Ländern kann nur dann hergestellt werden, wenn hinter identischen Abiturdurchschnittsnoten auch vergleichbare Kompetenzniveaus stehen. Zwar sollen – wie oben bereits ausgeführt – die EPAs die Vergleichbarkeit gewährleisten, dass dies aber kaum gelingt, sollen die im Folgenden präsentierten Befunde aus zwei großen Schulleistungsstudien in der gymnasialen Oberstufe illustrieren.

##### 4.1 Mathematikleistungen in beruflichen und allgemein bildenden Gymnasien Baden-Württembergs: Die TOSCA-Studie

In Baden-Württemberg wurden seit den 1970er Jahren die beruflichen Gymnasien als alternative Wege zur Hochschulreife systematisch ausgebaut. Technische, wirtschaftswissenschaftliche, agrarwissenschaftliche, sozialpädagogische, ernährungswissenschaftliche und biotechnologische Zweige vergeben die allgemeine Hochschulreife, die zur Aufnahme jedes beliebigen Studienfaches berechtigt. Berufliche Gymnasien rekrutieren typischerweise eine Schülerschaft, die in der Sekundarstufe I einen nicht-gymnasialen Bildungsgang (in der Regel die Realschule) besucht hat. Somit bietet sich dieser Schülerschaft, die nach der Primarstufe nicht den Übergang in das allgemein bildende Gymnasium erreicht hat, quasi eine zweite Chance für das Erreichen des Abiturs. Mittlerweile erwerben rund 30 Prozent aller Abiturienten ihre Hochschulzugangsberechtigung an einem beruflichen Gymnasium. In wie weit die im Abitur erreichten Kompetenzniveaus an beruflichen und allgemein bildenden

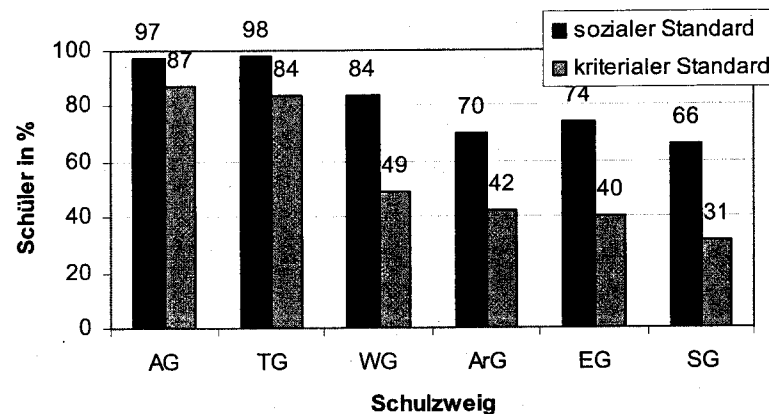
Gymnasien vergleichbar sind, haben Köller, Watermann, Trautwein und Lüdtke (2004) in der Studie *Transformation des Sekundarschulsystems und akademische Karrieren (TOSCA)* für die Fächer Mathematik und Englisch systematisch untersucht. Berücksichtigt wurden am Ende des Schuljahres 2001/2002 die Leistungen von  $N = 2854$  Schülerinnen und Schülern allgemein bildender und  $N = 1876$  Schülerinnen und Schülern beruflicher Gymnasien in der 13. Jahrgangsstufe. Mit Ausnahme biotechnologischer Gymnasien wurden alle Zweige beruflicher Gymnasien berücksichtigt. Im Folgenden findet bei der Berichterstattung der Ergebnisse aus Platzgründen eine Beschränkung auf die mathematischen Kompetenzen statt.

Die Mathematikleistungen wurden mit einem Test zur voruniversitären Mathematik erhoben, der bereits in TIMSS (vgl. Baumert u. a. 2000b) eingesetzt worden war. In einer Befragung an beruflichen und allgemein bildenden Gymnasien zeigte sich gleichermaßen, dass für die große Mehrzahl der Testaufgaben bis zum Abitur hinreichend viele Lerngelegenheiten bestanden hatten (zu Details s. Watermann/Nagy/Köller 2004). Hinsichtlich der erreichten Testleistungen wurden soziale und kriteriale Lernziele definiert. Ausgehend von der Annahme, dass Mathematiklehrkräfte in der gymnasiale Oberstufe ein geteiltes professionelles Verständnis davon besitzen, welche Kompetenzen hinter einer ausreichenden Leistung (5 Punkte) in einem Mathematik-Grundkurs haben, wurde die mittlere Testleistung aller Grundkursschüler, die im letzte Zeugnis 5 Punkte erreicht hatten, als (sozialer) Mindeststandard definiert. Kriteriale Standards im Sinne von Regelstandards wurden auf der Basis des in TIMSS verwendeten Kompetenzstufenmodells definiert (zu Details s. Watermann u. a. 2004). In diesem Modell beschreibt die zweite Kompetenzstufe ein Leistungsniveau, bei dem Kompetenzen, die typischerweise in der Sekundarstufe I erworben werden, sicher beherrscht werden. Die Kompetenzstufe III beschreibt die sichere Anwendung von in der Oberstufe erworbenen Lerninhalten. Letztere wurde im Sinne eines Regelstandards für Leistungskursschüler definiert, die Kompetenzstufe II als Regelstandard für die Grundkursschüler. Die Abbildung 1 zeigt den Prozentsatz der Schülerinnen und Schüler an den verschiedenen Zweigen, welche den sozialen Mindeststandard bzw. kriterialen Regelstandard erreicht haben. Deutlich erkennbar ist, dass die Erreichung der unterschiedlichen Standards für die allgemein bildenden und technischen Gymnasien unproblematisch ist, ein ganz anderes Bild zeigt sich an den übrigen Zweigen beruflicher Gymnasien, an denen erhebliche Schüleranteile die definierten Standards nicht erreichen. Die Leistungsunterschiede zwischen den Schulzweigen bildeten sich allerdings nicht in den entsprechenden Mathematiknoten ab, die sich kaum unterschieden (zu den Details Watermann u. a. 2004).

Zweifellos rückt mit diesen Befunden die Frage nach der Vergleichbarkeit der Abschlüsse und nach der Verteilungsgerechtigkeit beim Zugang zu zulassungsbeschränkten Studienfächern in den Vordergrund, implizieren doch die

Ergebnisse, dass identische Noten an beruflichen Gymnasien für andere Kompetenzniveaus stehen als an allgemein bildenden Gymnasien.

Abbildung 1: Prozentsatz der Schülerinnen und Schüler der TOSCA-Studie nach Schulform und Erreichen der sozialen und kriterialen Standards im Fach Mathematik am Ende der gymnasialen Oberstufe (vgl. Watermann u. a. 2004).



*Anmerkungen.* AG: allgemein bildendes Gymnasium; TG: technisches Gymnasium; WG: Wirtschaftsgymnasium; ArG: agrarwissenschaftliches Gymnasium; EG: ernährungswissenschaftliches Gymnasium; SG: sozialpädagogisches Gymnasium

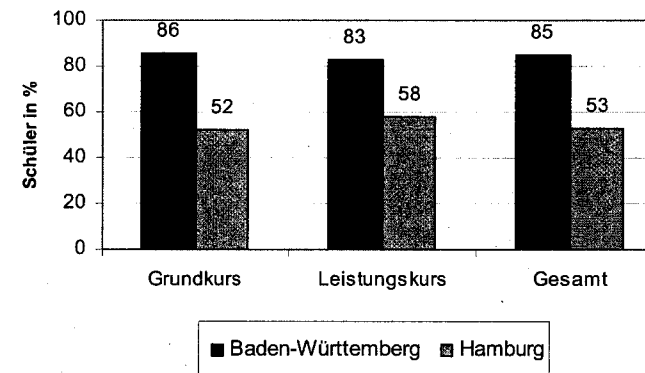
Im Hinblick auf die studienpropädeutische Funktion der gymnasialen Oberstufen haben Köller u. a. (2004) als Folge dieser Befunde angeregt, über systematische Zusatzangebote an beruflichen Gymnasien nachzudenken, welche die Kompetenzunterschiede zwischen den Zweigen reduzieren helfen. Ob solche Zusatzangebote fruchten, bliebe dann Gegenstand eines systematischen Systems der Qualitätsüberprüfung.

#### 4.2 Mathematikleistungen in allgemein bildenden Gymnasien Hamburgs und Baden-Württembergs: Ergebnisse der LAU- und TOSCA-Untersuchung

Der im vorherigen Abschnitt beschriebene Test zur voruniversitären Mathematik wurde auch in der Lernausgangslagenuntersuchung LAU 13 eingesetzt. In LAU 13 wurde der gesamte Hamburger Abiturjahrgang am Ende der 13. Jahrgangsstufe im Schuljahr 2004/2005 in verschiedenen Fächern hinsichtlich der erreichten Fachleistungen untersucht (s. hierzu ausführlich Trautwein/Köller/Lüdtke/Lehmann, 2007). Durch den Einsatz identischer Tests zur voruniversitären Mathematik können die Leistungen der Abiturienten aus

TOSCA mit denen aus LAU 13 verglichen werden. Die leitende Fragestellung der Analysen betraf die Vergleichbarkeit der Abschlüsse zwischen unterschiedlichen Ländern. Betrachtet man das Abschneiden der Gymnasiasten beider Länder in PISA 2003 (Deutsches PISA-Konsortium 2004), so weist Baden-Württemberg in dem mathematischen Kompetenzen einen Vorsprung von 29 Punkten auf (599 vs. 570 Punkten). Dieser Leistungsunterschied korrespondiert in etwa mit dem Wissenszuwachs, der in einem Schuljahr erreicht wird. Ausgehend von der Annahme, dass Vorwissen ein wichtiger Prädiktor für die am Ende der Sekundarstufe II erreichten Leistungen ist, konnte dementsprechend aus PISA 2003 abgeleitet werden, dass die untersuchten Schülerinnen und Schüler aus TOSCA denen in LAU überlegen sein sollten. Die deutliche Bestätigung dieser Hypothese findet sich in Abbildung 2 wieder. Dort ist im Sinne der oben definierten kriterialen Standards für die voruniversitäre Mathematik abgetragen, wie hoch die Anteile in Grund- und Leistungskursen beider Länder sind, welche die Standards erreichen.

Abbildung 2: Prozentsatz der Schülerinnen und Schüler an allgemein bildenden Gymnasien Hamburgs und Baden-Württembergs nach Kursniveau und Erreichen der kriterialen Standards im Fach Mathematik am Ende der gymnasialen Oberstufe (vgl. Trautwein u. a., 2007).



In Baden-Württemberg liegen die Anteile durchgängig über 80 Prozent, in Hamburg erreicht dagegen nur gut die Hälfte der Schülerinnen und Schüler die kriterialen Standards. Wie oben beim Schulformenvergleich ergibt sich hier erneut das Problem der regionalen Vergleichbarkeit des Abiturs und damit verbunden die Frage der Verteilungsgerechtigkeit beim Zugang zu zulassungsbeschränkten Fächern, in denen nach der Föderalismusreform und der Novelle der Hochschulrahmengesetzgebung vom August 2004 Absolventen beider Länder um die verfügbaren Plätze konkurrieren, ohne dass die früheren

länderspezifischen Quoten beim Zugang zu den *Numerus clausus*-Fächern noch ihre Gültigkeit besäßen.

#### 4.3 Schlussfolgerungen für die Qualitätssicherung in der Sekundarstufe II

Die hier vorgestellten Befunde aus zwei großen Schulleistungsstudien in der Sekundarstufe II haben exemplarisch für die voruniversitäre Mathematik deutlich gemacht, dass die aus PISA bekannten schulform- und länderspezifischen Unterschiede bei 15jährigen auch am Ende der gymnasialen Oberstufe bestehen und sich damit erhebliche Probleme im Hinblick auf die Vergleichbarkeit des Abiturs ergeben. Was hier am Vergleich zwischen Hamburg und Baden-Württemberg sichtbar wurde, lässt sich sicherlich auf weitere Länder ausdehnen, ohne dass dafür aber bislang empirische Evidenz bestünde. Wünschenswert ist in diesem Sinne die Etablierung eines Systems der Qualitätssicherung bzw. -überprüfung, in dem analog zur Gesamtstrategie der KMK für die Primarstufe und Sekundarstufe I die Leistungen deutscher Abiturienten in regelmäßigen Abständen in den Kernfächern Mathematik, Deutsch, Englisch und den Naturwissenschaften festgestellt werden. Die Grundlagen dafür sind mit den revidierten EPAs geschaffen worden (s. o.). Wiewohl damit keine automatische Angleichung der Leistungen in den Ländern zu erwarten ist, können solche Leistungsfeststellungen den Startpunkt für Maßnahmen der Qualitätssteigerung bilden, welche langfristig helfen können, Länderunterschiede zu reduzieren.

Ein System der Qualitätssicherung in gymnasialen Oberstufen kann zudem auf Seiten der Universitäten das Bewusstsein schärfen, dass die alleinige Berücksichtigung der Abiturdurchschnittsnote bei der Auswahl von Studienanfängern zu Verteilungsgerechtigkeiten führen kann und zudem keine hinreichend verlässlichen Informationen über die Studieneignung der Bewerberinnen und Bewerber liefert. In wie weit dies zu einer Etablierung von zusätzlichen Tests zur Feststellung der Studierfähigkeit führen wird, bleibt offen. Wünschenswert wäre solchen Tests aber auf jeden Fall (s. hierzu Formazin/Wilhelm/Köller 2006).

### Literaturverzeichnis

- American Association for the Advancement of Science (AAAS) (Ed.): Benchmarks for science literacy. Project 2061. New York 1993.
- Baumert, J./Bos, W./Lehmann, R. (Hrsg.): Dritte Internationale Mathematik- und Naturwissenschaftsstudie: Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Bd. 1: Mathematisch-naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit. Opladen 2000a.
- Dritte Internationale Mathematik- und Naturwissenschaftsstudie: Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Bd. 2: Mathemati-

- sche und physikalische Kompetenzen am Ende der gymnasialen Oberstufe. Opladen 2000b.
- Baumert, J./Kunter, M./Brunner, M./Krauss, S./Blum, W./Neubrand, M.: Mathematikunterricht aus Sicht der PISA-Schülerinnen und -Schüler und ihrer Lehrkräfte. In: Deutsches PISA-Konsortium (Hrsg.): PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs. Münster 2004, S. 314-354.
- Baumert, J./Lehmann, R. H. u. a.: TIMSS: Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Opladen 1997.
- Blum, W./Drüke-Noe/K./Hartung, R./Köller, O. (Hrsg.): Bildungsstandards Mathematik konkret. Sekundarstufe I: Aufgabenbeispiele, Unterrichtsideen und Fortbildungsmöglichkeiten. Berlin (im Druck).
- Bos, W./Lankes, E.-M./Prenzel, M./Schwippert, K./Walther, G./Valtin, R. (Hrsg.): Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich. Münster 2003.
- Deutsches PISA-Konsortium (Hrsg.): PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen 2001.
- PISA 2000. Die Länder der Bundesrepublik Deutschland im Vergleich. Opladen 2002.
- PISA 2003. Der Bildungsstand der Jugendlichen in Deutschland – Ergebnisse des zweiten internationalen Vergleichs. Münster 2004.
- PISA 2004. Die Länder der Bundesrepublik Deutschland im Vergleich. Münster 2005.
- Ehmke, T./Leiß, D./Blum, W./Prenzel, M.: Entwicklung von Testverfahren für die Bildungsstandards Mathematik. Unterrichtswissenschaft, 34, 2006, S. 220-238.
- Elley, W. B.: The IEA study of reading literacy: Achievement and instruction in thirty-two school systems. Oxford 1994.
- Europarat: Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren beurteilen. Berlin 2001.
- Formazin, M./Wilhelm, O./Köller, O.: Willkür vermeiden! Über sachlich gebotene Methoden der Beurteilung von Studienbewerbern. Forschung & Lehre, 12, 2006, S. 672-674.
- Hambleton, R. K./Jaeger, R. M./Plake, B. S./Mills, C.: Setting performance standards on complex educational assessments. Applied Psychological Measurement, 24, 2000, S. 355-366.
- Harms, U./Mayer, J./Hammann, M./Bayrhuber, H./Kattmann, U.: Kerncurriculum und Standards für den Biologieunterricht in der gymnasialen Oberstufe. In: H.-E. Tenorth (Hrsg.), Kerncurriculum Oberstufe II. Biologie, Chemie, Physik, Geschichte. Weinheim 2004, S. 22-84.
- Klieme, E./Avenarius, H./Blum, W./Döbrich, P./Gruber, H./Prenzel, M./Reiss, K./Riquarts, K./Rost, J./Tenorth, H.-E./Vollmer, H. J.: Zur Entwicklung nationaler Bildungsstandards: Eine Expertise. Berlin 2003.
- KMK: Vereinbarung über Einheitliche Prüfungsanforderungen in der Abiturprüfung. (Beschluss der Kultusministerkonferenz vom 01. 06. 1979 i. d. F. vom 24. 05. 2002). Bonn 2002.
- Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss. München 2004a.
- Bildungsstandards für die erste Fremdsprache (Englisch/Französisch) für den Mittleren Schulabschluss. München 2004b.



- Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie (Beschluss der Kultusministerkonferenz vom 01. 12. 1989 i. d. F. vom 05. 02. 2004). Bonn 2004c.
- Köller, O./Watermann, R./Trautwein, U./Lüdtko, O.: Wege zur Hochschulreife in Baden-Württemberg. TOSCA - Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien. Opladen 2004.
- Tenorth, H.-E. (Hrsg.): Kerncurriculum Oberstufe II. Biologie, Chemie, Physik, Geschichte, Politik. Weinheim 2004.
- Trautwein, U./Köller, O./Lehmann, R. H./Lüdtko, O. (Hrsg.): Schulleistungen von Abiturienten. Regionale, schulformbezogene und soziale Disparitäten. Münster 2007.
- Watermann, R./Nagy, G./Köller, O.: Mathematikleistungen in allgemein bildenden und beruflichen Gymnasien. In: O. Köller/R. Watermann/U. Trautwein/O. Lüdtko (Hrsg.): Wege zur Hochschulreife in Baden-Württemberg. TOSCA - Eine Untersuchung an allgemein bildenden und beruflichen Gymnasien. Opladen 2004, S. 205-283.
- Weinert, F. E.: Vergleichende Leistungsmessung in Schulen – Eine umstrittene Selbstverständlichkeit. In: F. E. Weinert (Hrsg.): Leistungsmessungen in Schulen. Weinheim 2001, S. 17-31.

Helmut Heid

## WAS VERMAG DIE STANDARDISIERUNG WÜNSCHENSWERTER LERNOUTPUTS ZUR QUALITÄTSVERBESSERUNG DES BILDUNGSWESENS BEIZUTRAGEN?

Die bildungspolitische Forderung, prüfbare Bildungsstandards zu entwickeln, war eine der ersten, weithin anerkannten und nun auch praktizierten Reaktionen auf die Ergebnisse der PISA-Vergleichsstudien, die aus deutscher Sicht als außerordentlich unbefriedigend bewertet werden (Klieme u. a. 2003, S. 6f.). Die Festlegung und die Überprüfung von Bildungs- oder Leistungsstandards werden inzwischen zu den wichtigsten Maßnahmen gerechnet, die die Verbesserung und Sicherung der Qualität von Unterricht, Schule sowie des nationalen Bildungswesens bezwecken. Im Folgenden soll geprüft werden, inwieweit die Standardisierung wünschenswerter Lernoutputs zur Qualitätsverbesserung des Bildungswesens beizutragen vermag.

### 1. Was versteht man unter „Standards“?

Das Wort „Standard“ wird verwendet, um die Kennzeichnung einer handlungsabhängigen oder von Handeln beeinflussbaren (dazu Klauer 1980, S. 64ff.) Sachverhaltsbeschaffenheit<sup>1</sup> unter Bezugnahme auf ein dafür unentbehrliches und von Entscheidungen abhängiges Beurteilungskriterium als

- mustergültig auszuzeichnen oder vorzuschreiben und
- allgemein-verbindlich festzulegen (vgl. Klieme u. a. 2003, S. 8).

Nicht schon die Kennzeichnung der zu normierenden Sachverhaltsbeschaffenheit bzw. Verhaltensausprägung konstituiert den Standard, sie beschreibt vielmehr nur den Sachgehalt des als Standard Postulierten oder Normierten. Erst die entscheidungsabhängige Normierung der (wertfrei) gekennzeichneten Verhaltensausprägung macht diese (Kennzeichnung der) Verhaltensausprägung zum Standard.

In der Debatte um die Qualitätsverbesserung bzw. -sicherung des Bildungswesens geht es im besonderen darum, Kompetenzen als Mindest-, Re-

<sup>1</sup> Zu diesen (bewusst „neutral“ bezeichneten) Sachverhalten rechne ich auch menschliche Verhaltensweisen, und zu den Sachverhaltsbeschaffenheiten Verhaltensausprägungen.