

SASCHA SCHANZE UND PETER NENTWIG

Standards im Naturwissenschaftlichen Unterricht – ein internationaler Vergleich

Standards in science education – an international comparison

Zusammenfassung

Die Entscheidung zur Einführung nationaler Bildungsstandards in Deutschland, Österreich oder der Schweiz gründet sich auf Erfahrungen, die in anderen Nationen mit vergleichbaren Bildungsreformen gemacht wurden. Die in der deutschen Bildungslandschaft vorhandenen Strukturen erfordern adaptive Prozesse, die sich durch alle Ebenen hindurch ziehen. Dazu gibt es übergreifende Fragestellungen, die aus der Perspektive von Nationen mit unterschiedlichen Erfahrungen betrachtet, hilfreich für diese Prozesse sein können. Im Rahmen eines Symposiums wurden daher Erfahrungen von naturwissenschaftlichen Fachdidaktikern aus 14 Nationen zu folgenden Fragen zusammengetragen: *Wie sind Standards für den naturwissenschaftlichen Unterricht definiert? Wie kann das Erreichen der Standards für den naturwissenschaftlichen Unterricht überprüft werden? Welche Funktion hat die Überprüfung? Welche Nebenwirkungen werden von der Standardüberprüfung befürchtet bzw. beobachtet? und Können Bildungsstandards für die Naturwissenschaften ohne den Zwang zur zentralen Überprüfung implementiert werden?* Dieser Beitrag fasst Ergebnisse des Symposiums zusammen.

Schlüsselwörter: Bildungsstandards, Naturwissenschaftlicher Unterricht, Kompetenzen, Bildungsmonitoring

Abstract

The decision to introduce national standards to the education systems of Germany, Austria or Switzerland was based on the experience made in other countries with similar reforms. The existing structures require adaptive processes throughout the implementation on all levels of the education system in Germany. For these processes certain questions are crucial, which to answer from other nations' experiences might prove helpful. Science educators from 14 countries therefore dealt with the following questions at an international symposium: *How are standards in science education defined? How can the achievement of science education standards be assessed? What function does the assessment have? What side-effects of testing standards are feared or observed? Can standards for science education be implemented without centralized assessment?* This article summarizes results of the symposium.

Keywords: standards, science education, competencies, educational monitoring

1 Zur Situation in Deutschland

Seit dem Schuljahr 2005/06 sind sie für alle Schulen in Deutschland verbindlich: Die Bildungsstandards für den mittleren Schulabschluss in den Fächern Biologie, Chemie und Physik. Sie beschreiben Kompetenzen, die Schülerinnen und Schüler am Ende der 10. Klassenstufe in den naturwissenschaftlichen Fächern erworben haben sollen. Die

Bundesländer haben sich dazu verpflichtet, die Standards zu implementieren, insbesondere Kerncurricula zu entwickeln und Lehrkräfte aus- und fortzubilden. Dieser Prozess ist noch nicht abgeschlossen und schon wird der nächste Schritt getan: Die Kultusministerkonferenz beschließt am 17. und 18. Oktober 2007 die Weiterentwicklung der Einheitlichen Prüfungsanforderungen für die Abiturprüfung (EPA) zu bundesweiten Bildungsstandards

für die Abiturprüfung¹. Wie schon für den Mittleren Schulabschluss werden zunächst die Fächer Deutsch, Mathematik und erste Fremdsprache (Englisch/Französisch) und ferner die naturwissenschaftlichen Fächer (Biologie, Chemie, Physik) berücksichtigt. Diese Standards sollen möglichst schon für die Schülerinnen und Schüler gelten, die im Jahre 2010/2011 in die Qualifikationsphase eintreten. Konzepte für die Überprüfung der Standards werden derzeit vom IQB dem Institut für Qualitätsentwicklung im Bildungswesen in Berlin erstellt.

Als ein Grund für die Einführung der Standards sind nach wie vor die enttäuschenden Ergebnisse von Schülerinnen und Schülern in den Vergleichsstudien TIMMS und PISA zu sehen. Ein besseres Abschneiden anderer Länder wird unter anderem auf bereits eingeführte Bildungsstandards und systematisch durchgeführte überregionale Testverfahren zur Überprüfung der Standards zurückgeführt. Die Expertise von Klieme et al. (2003) hat Realisierungen in verschiedenen Ländern analysiert und unter Berücksichtigung der dort bereits gemachten Erfahrungen, der Situation in der hiesigen Bildungslandschaft und den Erkenntnissen aus der Forschung in Pädagogik, Psychologie und Fachdidaktik Ziele und Kriterien für Bildungsstandards definiert. In nicht einmal einem Jahr wurden von Fachdidaktikern und Fachdidaktikerinnen sowie Schulpraktikern und Schulpraktikerinnen die Standards entwickelt und von der KMK verabschiedet. Sie werden zusammen mit den überarbeiteten Einheitlichen Prüfungsanforderungen in der Abiturprüfung (EPA) als geeignete Grundlage für die Steuerung im Bildungswesen gesehen, „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“ (Köller, 2007, 13).

Doch ließen sich aufgrund der kurzen Entwicklungszeit Empfehlungen der Experten-Gruppe um Klieme nicht in vollem Umfang umsetzen. So wurde der Paradigmenwechsel von input-orientierten Lehrplänen zu den output-orientierten Standards, die zu erreichende Kompetenzen für verschiedene Stufen im Verlaufe der Schulkarriere eines Lerners beschreiben sowohl in der Bildung allgemein (siehe z.B. Böttcher, 2003; GEW, 2003; Gröben, 2003; Saldern & Paulsen, 2003; Böttcher, 2004; Messner, 2004; Plöger, 2004; GEW, 2004a, 2004b; Arnold, 2005; Demmer & Schweitzer, 2005; Lange, 2005; Regenbrecht, 2005; Spinner, 2005; Stadtfeld, 2005; Benner, 2007; Herzog, 2007; Labudde, 2007a) als auch im Bereich der naturwissenschaftlichen Bildung (siehe z.B. Kremer, 2004; Kullmann, 2004; Voelzke, 2004; Bayrhuber et al., 2005; Parchmann, 2005; Labudde, 2007c) kontrovers diskutiert.

Ein wesentlicher Kritikpunkt ist, dass bei Einführung von Standards das traditionelle Ziel deutschen Schulunterrichts, die gegenwarts- und zukunftsorientierte allgemeine Bildung als Befähigung zu vernünftiger Selbstbestimmung auf der Strecke bleibe und einem rein pragmatischen, testbaren Nutzungswissen geopfert werde (z.B. Jürgens, 2004; Zedler, 2007). In der Schweiz, in der man wie in Deutschland Bildungsstandards einführen will, wird die Skepsis geäußert, „ob Messungen von Schülerleistungen und die Systemsteuerung per Output tatsächlich zu einer Verbesserung von Schule und Unterricht führen werden“ (Herzog, 2007, 59). Dies sei zwar intendiert, aber nicht zwingend vorgegeben.

Auch die Frage nach der theoretischen Fundierung der Standards und der ihnen zu Grunde liegenden Kompetenzen sowie nach ihrer demokratischen Legitimierung wird widersprüchlich diskutiert (z.B. Saldern & Paulsen, 2003; Jürgens, 2004; Oelkers, 2004; Orth, 2005; Tenorth, 2005). Vor allem im Zusammenhang mit der Beschreibung und Fun-

¹ KMK-Pressemitteilung vom 18.10.07, einzusehen unter <http://www.kmk.org/aktuell/home.htm> (zuletzt eingesehen am 25.01.08)

dierung von Kompetenzmodellen ist in den verschiedenen Fachdisziplinen Forschungs- und Entwicklungsarbeit zu leisten. Hier beteiligen sich gleichermaßen Fachdidaktik und Pädagogische Psychologie, kämpfen aber oft mit der Komplexität der Vorhaben (Tenorth, 2008). Die Didaktiken verfügen in manchen Fächern noch über wenig Tradition in der domänenspezifischen Unterrichtsforschung. Die Pädagogische Psychologie zeichnet sich zwar durch ihre Expertise in der empirischen Lehr- und Lernforschung aus, beweist aber auch eine Distanz zu fachspezifischen Themen- und Problemstellungen (Tenorth, 2008, Klieme & Rakoczy, 2008).

Kritiker bemängeln weiter, in ihrer gegenwärtigen Formulierung konzentrierten sich die Standards auf relativ eng umschriebene Lernziele, ohne die Qualität der Lernprozesse, die strukturellen Voraussetzungen an den Schulen oder die Qualifikation der Lehrkräfte ausreichend zu berücksichtigen. „Ohne Standards für die Lehrerrolle, für Unterrichtsbedingungen, für Arbeits- und Sozialformen in der Schule verkommt Bildung zu außen-gesteuertem Training.“ (Brügelmann, 2003, 235). Dass Bildungsstandards auch solche Aspekte umfassen können, zeigen u.a. die National Science Education Standards der USA, die über die inhaltlichen Lernziele hinaus ein ganzes System aufeinander bezogener Standards beschreiben (National Research Council, 1996; Bybee, 2002). Während in manchen anderen Ländern (vgl. Jauhainen, 2005 für Finnland) vielfältige Unterstützungssysteme parallel zu den Bildungsstandards geschaffen wurden, damit jeder einzelne Schüler sein individuelles Lernoptimum erreichen kann, wird beklagt, dass „mit den deutschen Bildungsstandards (bisher noch) kein Förderkonzept verbunden“ sei (Ratzki, 2005, 51).

Kritik wird auch an der Platzierung der Standardüberprüfung geübt (GEW, 2004a). Die Testaufgaben sollen helfen, den gegenwärtigen Leistungsstand der Schüler zu erfassen und damit eine Grundlage für Maßnahmen zur Förderung bilden. Dazu sollten die Überprüfungen aber im laufenden Unterricht und nicht als Abschluss an Schaltstellen der

Schullaufbahn stattfinden. Huber (2004) z.B. sieht dadurch die Gefahr, dass die Überprüfung für den Mittleren Schulabschluss zur bloßen Selektion für die Oberstufe würde. Die weitreichendste Befürchtung besteht darin, dass nach Etablierung von Standards nur noch das unterrichtet werde, was auch geprüft wird. „Teaching to the test“ ist das Stichwort, hinter dem ein verengtes, lediglich auf den Abschlusstest ausgerichtetes Lehren droht (vgl. Bachmann, Dangl, Mayr, & Spenger, 2004; Neuweg, 2004; Altrichter & Posch, 2005). Dass diese Befürchtung nicht abwegig ist, zeigen manche Erfahrungsberichte, etwa aus den USA (z.B. Rosengren, 2004).

Letztlich wird diskutiert, ob das Ende der Mittelstufe ein glücklich gewählter Zeitpunkt für die Überprüfung der Standards sei (GEW, 2004a). Das Messen von Lernfortschritten sollte im Sinne einer Diagnose die Basis für unterstützende Maßnahmen sein. Somit wären sie besser im Verlauf der Schulzeit aufgehoben als eher am Ende der Schulkarriere. Darüber hinaus wird befürchtet, sie könnten als Selektionsinstrument für den Übergang in die gymnasiale Oberstufe genutzt bzw. missbraucht werden (z.B. Huber, 2004).

2 Die Suche jenseits des Tellerrandes

Die Entscheidung, in Deutschland Bildungsstandards einzuführen, basierte unter anderem auf den Erfahrungen anderer Nationen, in denen Standards bereits einige Zeit etabliert waren. In den USA, zum Beispiel, wurden Ende der Achtziger Jahre Standards eingeführt, so dass inzwischen Erfahrungen aus fast zwei Jahrzehnten vorliegen. Zum Zeitpunkt der Einführung von Standards in Deutschland gab es einige wenige Berichte über Erkenntnisse aus anderen Ländern (z.B. Bonnet, 2002; BMBF, 2003; Ackeren, 2005; Baumann, Fessler, & Wilimann, 2005; Horstkemper, 2005; Jauhainen, 2005), die jedoch wenig über Standards speziell für die naturwissenschaftliche Bildung ausagten. Daher wurde im Februar 2006 im Rahmen der Kooperation zwischen der Science Education Group an der University of York

(UYSEG) und dem Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN) mit Unterstützung der Deutschen Forschungsgemeinschaft (DFG) ein internationales Symposium veranstaltet, zu dem Experten der naturwissenschaftlichen Fachdidaktiken aus zehn europäischen und vier außereuropäischen Ländern eingeladen wurden, in denen Erfahrung mit Bildungsstandards für den naturwissenschaftlichen Bereich vorlagen. Dabei wurde auf ein Format zurück gegriffen, dass sich bereits in den vorangegangenen Symposien der Zusammenarbeit von IPN und UYSEG bewährt hatte (Bennett, Holman, Millar, & Waddington, 2005; Nentwig & Waddington, 2005): Alle Experten lieferten Berichte über die Situation ihrer Länder in Bezug auf Bildungsstandards im naturwissenschaftlichen Bereich, die rechtzeitig vor der Veranstaltung allen Teilnehmern auf einer internen Webseite zugänglich gemacht wurden. Diese Berichte dienten als Basis für die detaillierten Diskussionen während des Symposiums, die schließlich zu den überarbeiteten Fassungen für eine gemeinsame Buchpublikation führten (Waddington, Nentwig, & Schanze, 2007). Als Orientierung für das Symposium galten fünf Leitfragen:

Wie sind Standards für den naturwissenschaftlichen Unterricht definiert?

In Deutschland werden die Bildungsstandards für die naturwissenschaftlichen Fächer Biologie, Chemie und Physik über Kompetenzen definiert. Dieses Konzept unterscheidet sich von dem in anderen Ländern, in denen zum Teil der Begriff Standard nicht einmal verwendet wird, wenn auch eine Vorstellung davon existiert, was von Lernenden zu bestimmten Zeiten im Lernprozess zu erwarten sein sollte. Für ein besseres Verständnis der verschiedenen Konzeptionen ist es notwendig zu klären, wie der Terminus „Standard“ in den verschiedenen Ländern genutzt wird.

Wie kann das Erreichen der Standards für den naturwissenschaftlichen Unterricht überprüft werden?

Es bedarf eines Testinstrumentes, um zu bestimmen, ob Lernende zu einem bestimmten Zeitpunkt die erwarteten Kompetenzen auf dem erforderlichen Niveau erreicht haben. Für diese Zwecke wurden verschiedene Testarten entwickelt und in zahlreichen Ländern eingesetzt. Der Charakter eines Tests hängt stark von der Definition der Standards und ihren Intentionen ab.

Welche Funktion hat die Überprüfung?

Das Testergebnis kann zu unterschiedlichen Konsequenzen führen. So können die Ergebnisse zum Beispiel der Unterstützung Lernschwacher oder der Selektion einer Elite dienen. Sie können Folgen für die einzelne Lehrkraft oder die Schule haben und schließlich können Konsequenzen für die Steuerung eines Bildungssystems aus ihnen abgeleitet werden.

Welche Nebenwirkungen werden von der Standardüberprüfung befürchtet bzw. beobachtet?

Kritiker der Bildungsstandards sehen die Gefahr, dass die Festlegung von Standards die Weiterentwicklung des Curriculums in der Weise dominieren könne, dass nur noch das unterrichtet würde, was auch messbar sei. Dies impliziert Auswirkungen sowohl auf das Curriculum als auch auf die Unterrichtspraxis.

Können Bildungsstandards für die Naturwissenschaften ohne den Zwang zur zentralen Überprüfung implementiert werden?

Eine zentrale Überprüfung der Bildungsstandards war für die naturwissenschaftlichen Fächer in Deutschland zur Zeit des Symposiums nicht geplant. Das könnte man als Chance für eine besonnene Weiterentwicklung der Standards sehen, es könnte aber auch die Motivation für eine Implementation der Standards hemmen: Standards, die nicht überprüft werden, könnten nicht ernst genommen werden.

In den nächsten Abschnitten werden die Auskünfte zusammengefasst, die die Experten aus den 14 Ländern zu diesen Leitfragen gegeben haben.

3 Bildungsstandards in anderen Ländern

Die Gründe, überhaupt Standards in das Bildungssystem einzuführen, sind vielschichtig. Die Ergebnisse internationaler Vergleichsstudien waren nicht nur in Deutschland eine treibende Kraft. Auch die Schweiz und Österreich sind im Prozess einer Bildungsreform. Als weitere Beweggründe gelten:

- **Das Bereitstellen eines Systems der Rechenschaftspflicht (accountability)** – Lehrkräfte und Schulen sollen sich für die Ergebnisse ihrer Arbeit vor Schülern, Eltern und Vertretern des Bildungswesens verantworten.
- **Der Bedarf an Feedback** – Schüler haben, wie auch ihre Eltern, ein reges Interesse an einer Rückmeldung über den eigenen Leistungsstand und darüber, welche Möglichkeiten sich ihnen für den zukünftigen beruflichen Werdegang bieten. Darüber hinaus möchten auch die Lehrkräfte und Schulen eine Rückmeldung über ihr pädagogisches Handeln bekommen, und schließlich wollen auch die für die Bildung zuständigen Behörden eine Rückmeldung darüber haben, wie gut ihre Investitionen in die Bildung umgesetzt werden.
- **Das Schaffen gleicher Bildungschancen** – In Ländern, die entweder ein föderales System haben (wie die USA) oder ein differenziertes Schulsystem (wie in Deutschland) oder aber beides (wie die Schweiz) ist die Harmonisierung der Bildung eine große Motivation. Standards können hier zumindest im Kern einheitliche Bildungsziele schaffen.

- **Die Verbesserung des Unterrichts** – In vielen Ländern wird eine Verbesserung der Unterrichtspraxis im naturwissenschaftlichen Bereich für dringend erforderlich gehalten. Standards werden daher auch eingeführt, um einen Rahmen für diese Entwicklung bereit zu stellen.

Wie die weiteren Ausführungen zeigen werden, hängen viele Maßnahmen in den einzelnen Nationen auch davon ab, welche der Ziele im Lande vorrangig verfolgt werden.

3.1 Wie sind Standards für den naturwissenschaftlichen Unterricht definiert?

In den meisten englischsprachigen Ländern wird der Begriff ‚Standard‘ benutzt, um zu beschreiben, was als Ergebnis der Lehr- und Lernanstrengungen in der Schule erwartet wird. Das Wort ist in viele andere Sprachen übernommen worden, auch wenn es nicht ursprünglich Teil der Landessprache war. Mit dem gleichen Sinngehalt werden aber auch andere Begriffe benutzt. In Australien wurden ‚frameworks‘ entwickelt, in Schweden ‚goals‘ und in Taiwan ‚guidelines‘, die „als Standards gelten, auch wenn sie nicht Standards genannt werden“ (Chiu, 315)². Bryce macht für Schottland die begriffliche Vielfalt deutlich: „Im Bemühen um Klarheit haben wir mit verschiedenen Begriffen gespielt – Ziele, Ergebnisse, Kriterien, Kompetenzen oder Standards – ohne zu eindeutigen Verhaltensbeschreibungen zu kommen.“ (Bryce, 257) Gemeinsam scheint allen Ländern das Bemühen, festzulegen, was Schüler eines bestimmten Alters oder einer Klassenstufe im naturwissenschaftlichen Bereich wissen, verstehen und handelnd anwenden können sollen. Wie diese Festlegung beschrieben wird, unterscheidet sich von Land zu Land.

² Die Seitenzahlen beziehen sich auf Beiträge in Waddington, D., Nentwig, P., Schanze, S. (Eds.) (2007) *Making it comparable. Standards in Science Education*. Münster: Waxmann. Die Zitate stehen zum Teil in der freien Übersetzung dieser Autoren.

3.1.1 Standard oder Curriculum

Die deutschen Bildungsstandards sind einschließlich der dazu gehörigen Kompetenzen von den Kultusministern der Länder gemeinsam erlassen worden. Der Weg, auf dem die Standards erreicht werden sollen, bleibt zunächst offen. Basierend darauf entwickeln die Bundesländer eigene Kerncurricula, die zwar auf Wege zur Förderung der durch die Standards festgelegten Kompetenzen eingehen können, in der Regel aber in Bezug auf zu vermittelnde Inhalte nicht konkret werden, sondern den Schulen als Vorgabe für die Erstellung des eigenen Schulcurriculums dienen. Deutschland hat sich damit gegen die Tradition gewandt, wonach die Standards aus dem Curriculum abgeleitet und erst dann durch die zu erreichenden Kompetenzen beschrieben werden (Dubs, 2007). Die beiden Zugänge unterscheiden sich im Wesentlichen durch die unterschiedlichen Kontrollabsichten. Das ‚National Curriculum‘ Englands beschreibt beispielsweise recht ausführlich, was den Kindern gelehrt werden soll. Auch andere Länder folgen mehr oder weniger detailliert diesem Modell der Vorgabe durch Curricula, Richtlinien oder Lehrpläne, d.h. sie kontrollieren den vorgesehenen Input in die Lehr- und Lernprozesse. Was die Lehrbemühungen tatsächlich leisten, bleibt unsicher. Millar (2007) unterscheidet daher das beabsichtigte Curriculum, das im Unterricht tatsächlich umgesetzte und das schließlich erreichte. Die Unterschiede zwischen diesen Varianten werden nur nachprüfbar, wenn im nächsten Schritt Standards definiert und Verfahren zu ihrer Überprüfung eingeführt werden. In England und Wales gibt das Curriculum deshalb nicht nur vor, was unterrichtet werden soll. Es werden zur Output-Kontrolle auch Standards zur Überprüfung der erreichten Leistungsstufe beschrieben. Auch in einigen Staaten der USA ist eine solche Mischung der beiden Paradigmen – Input- und Output-Kontrolle – zu beobachten, wenn ein Curriculum

vorschreibt, was unterrichtet werden soll, gleichzeitig aber ein etabliertes System von Tests implizit die Ergebnisse vorschreibt.

3.1.2 Standards oder Ziele

Neben Standards werden oft auch Ziele als generelle Richtlinien für Schulen beschrieben. Dänemark beschreibt „in eher allgemeinen Formulierungen national verbindliche Ziele für das Fach“ wie z.B. für Biologie: „Schüler sollen Kenntnisse über lebende Organismen und die umgebende Natur, Umwelt und Gesundheit erwerben – mit besonderer Betonung der Zusammenhänge“ (Dolin, 75). Die National Science Education Standards der USA fassen Ziele allgemeiner wie „naturwissenschaftliche Prinzipien und Denkweisen angemessen bei der persönlichen Entscheidungsfindung anwenden“ (Lederman & Lederman, 350). In Finnland sind Ziele von den Schulen einzuhaltende Richtlinien. Sie beschreiben generelle oder fachspezifische Lernerwartungen wie z.B. „In den Klassen 7 – 9 lernen die Schüler in Chemie die physikalischen und chemischen Konzepte, die die Eigenschaften der Stoffe beschreiben, und sie lernen, diese Konzepte anzuwenden.“ (Lavonen, 107).

In den meisten Fällen werden neben den sehr generellen Zielen auch spezifischere Ziele definiert. Schweden macht hier mit „goals to aim for“ als Richtungsvorgaben für die Wissensentwicklung und Basis für die Unterrichtsplanung sowie den „goals to attain“ als Minimalzielen, die von allen Schülern in der fünften und neunten Klassenstufe erreicht werden sollen und über die Versetzung entscheiden, eine explizite Unterscheidung (Strömdahl, 265). Finnland beschreibt mit „final-assessment criteria“ ebenfalls „outcomes, what pupils should understand and be able to do“ (Lavonen, 106). Diese Kriterien können, wie auch die schwedischen „goals to attain“, als Standards im engeren Sinne angesehen werden.

3.1.3 Inhalt oder Performance

Oft werden Standards über Inhalte formuliert, die für eine bestimmte Klassenstufe zu unterrichten sind. Sie orientieren sich damit wie in England an einem eher curricularen, input-orientierten Verständnis von Standards. Auch in Schottland besteht "der Kern (des betreffenden Dokuments) aus 16 Seiten inhaltsbezogener Aussagen." (Millar, 83)

Die National Science Education Standards (NSES) in den USA formulieren die Inhalte einheitlich nach dem Muster: „Als Ergebnis ihrer Bemühungen ... sollen alle Schüler ... entwickeln“ (Lederman & Lederman, 353). Obwohl sie so eine Output-Orientierung suggerieren, unterscheiden sie sich nicht von anderen traditionellen Lehrplänen mit inhaltlichen Vorgaben. Bemerkenswert ist allerdings die Fokussierung auf *Nature of Science* und auf *Inquiry*: „handeln können sowohl als wissen“ (Lederman & Lederman, 351).

Standards als Leistungsvorgaben (*Performance Standards*) dagegen beziehen sich auf das Ergebnis der Lehr- und Lernprozesse. Sie legen fest, was von Lernenden zu einem bestimmten Zeitpunkt zu erwarten ist. Beeinflusst von den Entwicklungen in Deutschland haben Österreich und die Schweiz sich entschieden, *Performance Standards* einzuführen. Auch Finnlands „final-assessment criteria“ können als solche angesehen werden. Das australische New South Wales beschreibt „levels of achievement of learning“ (Hafner, 37), die aus dem Curriculum des Staates abgeleitet werden und zur Entwicklung von *Performance Standards* dienen. In England und Wales legen „level descriptions“ (Millar, 89) ebenfalls fest, welche Leistungen von Schülerinnen und Schülern erwartet werden.

Curricula legen oft Leistungserwartungen fest. Nur verbunden mit einer Überprüfung der tatsächlichen Leistungen können diese Erwartungen auch als *Performance Standards* angesehen werden. Für das National Curriculum in England und Wales, das mit einem gut etablierten Testverfahren begleitet wird, trifft dies zu.

3.1.4 Standards und Kompetenzen

Einige Nationen haben damit begonnen, Leistungserwartungen als Kompetenzen zu beschreiben. Allerdings wird dieser Terminus leicht unterschiedlich benutzt. So beschreibt Taiwan acht Komponenten im Bereich Scientific und Technological Literacy als Kernkompetenzen, wobei jede Komponente durch Teilkompetenzen („competence indicators“ (Chiu, 309)) weiter ausdifferenziert wird. Diese Teilkompetenzen legen allerdings einen Schwerpunkt auf kognitive Fähigkeiten (Wissen und Verstehen). Das Curriculum von Portugal beschreibt Standards als generelle und spezifische Kompetenzen. „Zu diesen Kompetenzen gehören Wissen (fachbezogenes, prozedurales oder methodologisches, epistemologisches) Bewertung, Kommunikation und Einstellung“ (Galvao, Reis, Freire, & Oliveira, 240). Dänemark ergänzt die traditionellen inhaltspezifischen Anforderungen für die Naturwissenschaften durch vier Kompetenzen:

- „Eine empirische Kompetenz: die Fähigkeit, zu beobachten, zu beschreiben, zu experimentieren, zu messen, etc.
- Eine Repräsentationskompetenz: die Fähigkeit, Sachverhalte auf verschiedene Weise darzustellen (mit Grafiken, Zeichnungen, Bildern, etc.) und zwischen diesen Repräsentationsformen zu wechseln.
- Eine Modellierungskompetenz: die Fähigkeit Komplexität zu reduzieren, Kausalitäten festzustellen, unterschiedliche Arten von Modellen herzustellen und zu benutzen, etc.
- Eine perspektivische Kompetenz: die Fähigkeit, die Naturwissenschaften perspektivisch einzuordnen, die Rolle der Naturwissenschaften in der Gesellschaft zu reflektieren, naturwissenschaftliches Wissen im Verhältnis zu anderen Wissensbeständen zu reflektieren, etc.“ (Dolin, 77).

In Schottland werden die Ziele/Standards/Kriterien (s.o.) unter drei Rubriken zusammengefasst, die man ebenfalls als Kompetenzen auffassen kann. Danach sollen die Schüler:

- „Wissen und Verstehen in Bezug auf den Inhalt einer Unterrichtseinheit demonstrieren,
- Probleme in Bezug auf den Inhalt einer Unterrichtseinheit lösen,
- Informationen zusammenstellen und analysieren, die durch Experimente gewonnen wurden.“ (Bryce, 256)

Werden Standards durch Kompetenzen formuliert, dann sind sie in der Regel wie die deutschen Standards in mehreren Dimensionen – meist zwei oder drei – strukturiert. In Frankreich werden als Kompetenzbereiche „Wissen, Fähigkeiten und Einstellungen“ (Malléus, 130) beschrieben, und die fachliche Domäne bildet die zweite Dimension. Ähnlich sind die Standards in Portugal angelegt. Das National Curriculum in England sieht implizit drei Dimensionen vor, auch wenn sie nicht ausdrücklich so als Modell formuliert sind. Neben ‚sort of performance‘ (Kompetenz) und ‚attainment target‘ (Inhaltsbereich) bilden die ‚levels of attainment‘ (Kompetenzstufen) eine dritte Dimension. In den meisten Fällen werden statische Modelle dargestellt. Entwicklungsmodelle beschreiben Annahmen über die Kompetenzentwicklung und legen Stufen während des Lernprozesses fest. Derartige Modelle können entweder an ein Kontinuum des Lernens angelehnt sein, indem unabhängig vom Alter Lernfortschritte betrachtet werden, oder an Klassenstufen, indem in den Blickpunkt genommen wird, was ein Schüler abhängig vom Alter wissen und können sollte. Das „key stages“ Konzept in England und Wales kann als ein Beispiel für das erste Prinzip angesehen werden, die an Inhalten orientierten Standards der amerikanischen NSES als Beispiel für das zweite. Bis heute sind alle Kompetenzmodelle normativ. Kein Modell wurde bisher empirisch hergeleitet oder abgesichert. Die Bestre-

bungen zur empirischen Absicherung der eigenen Modelle in der Schweiz und in Deutschland sind insofern bemerkenswert.

3.1.5 Minimal- oder Regelstandards

Mit der Formulierung von Regelstandards sind die Kultusminister in Deutschland in einem wesentlichen Punkt von der Klieme-Expertise abgewichen. Sie empfiehlt Minimalstandards, bei denen der Unterricht sicher zu stellen hat, dass alle Schüler, abgesehen von Lernenden mit ernsthaften Lernschwierigkeiten, die Standards erreichen. Solche Mindeststandards sind in New South Wales oder Schweden formuliert. Dort ist eine Tendenz zu beobachten, dass Lehrkräfte und Schulen für die Ergebnisse effektiven Lehrens und Lernens verantwortlich gemacht werden.

Regelstandards, wie sie in Deutschland, Österreich und der Schweiz vorgesehen sind, beschreiben die von einem durchschnittlichen Lernenden erwartete Leistung. Auch in Finnland definieren „Kriterien für die Abschlussprüfung das Maß an Kenntnissen und Fertigkeiten für das Erreichen der Note 8 auf einer Skala von 4 bis 10“ (Lavonen, 109). Einige Lernende werden diese Vorgabe erreichen, andere werden daran scheitern. Bei derartigen Systemen ist zu beobachten, dass die Verantwortung für ein erfolgreiches Abschneiden tendenziell bei den einzelnen Lernenden verbleibt. Welche Konsequenzen Minimalstandards für ein System zur Verbesserung der Unterrichtsqualität haben könnten, wird bei diesen Beobachtungen bereits offensichtlich.

3.1.6 Standards, wo und wann?

Länder mit einem zentral gesteuerten Bildungssystem wie Frankreich oder England tendieren dazu, Standards auch zentral einzuführen. Allerdings führen zentral organisierte Bildungssysteme nicht notwendigerweise

auch zu zentral verordneten Standards, wie die Beispiele aus Finnland oder Schweden zeigen. Zwar werden auch dort generelle Tendenzen in Form globaler Ziele formuliert, auf lokaler Ebene sind dann aber von den behördlichen Bildungseinrichtungen gemeinsam mit den Lehrkräften die Curricula zu planen und außerdem die Überprüfung der Standards zu organisieren.

Länder mit föderalem System wie Österreich, Deutschland, die Schweiz oder die USA haben Standards als Rahmen eingeführt, die in den einzelnen Staaten oder Provinzen in spezifische Standards oder Curricula umgesetzt werden. In den deutschen Bundesländern scheinen die Realisierungen der Kerncurricula eng an den Vorgaben zu gelingen. In den meisten Staaten der USA ist eine Ähnlichkeit mit dem ursprünglich vorgegebenen Rahmen nur noch in Teilen zu erkennen (s. Lederman & Lederman, 2007). Standards werden außerdem für unterschiedliche Zeitpunkte im Bildungsverlauf beschrieben. Nationen wie Frankreich, Portugal oder Deutschland definieren die Standards für den Abschluss der allgemeinen Schulpflicht (Ende 9. oder 10. Klasse). Andere Länder legen Standards für verschiedene Stufen fest. In Schweden sind es die fünfte und neunte, in Finnland die vierte, sechste und neunte, in der Schweiz die zweite, sechste und neunte, in Taiwan die zweite, vierte, sechste und neunte Klassenstufe. In New South Wales, Australien werden Erwartungen für die zu erreichende Lernqualität für alle zwei Jahre (2, 4, 6, 10, 12) und in einigen Staaten der USA werden Standards sogar für jedes Jahr (K-12) beschrieben. (s. Nentwig, 2006)

3.2 Wie kann das Erreichen der Standards für den naturwissenschaftlichen Unterricht überprüft werden?

Es ist sicher keine Frage mehr, ob Standards überprüft werden sollten oder nicht. Ohne eine Überprüfung der erreichten Leistung in Relation zu der festgeschrie-

benen Lernerwartung ist das Instrument für Lehrer, Schüler und Eltern wirkungslos. Natürlich müssen die Tests eine Aussage über das Erreichen der Standards ermöglichen. Aber nicht nur diese Passung ist eine notwendige Voraussetzung. Standards und Testinhalte sollten auch die Unterrichtspraxis widerspiegeln und umgekehrt. Die Angleichung der Überprüfungen an die Vorgaben und den Unterricht scheint eines der sensibelsten Bereiche einer Bildungsreform zu sein (s. Mamlok-Naaman, 2007). Besondere Schwierigkeiten haben Nationen, in denen es bisher kein etabliertes Testsystem gibt. Aus Dänemark werden Erfahrungen beschrieben, die sicher auch anderen Nationen nicht fremd sind. Erste Tests waren dort einfache computerbasierte Multiple-Choice-Fragen, die eher Faktenwissen überprüften (Dolin, 2007). Wie aufwendig ein Prozess für alle beteiligten Ebenen sein kann, der versucht, Standards, Test und Unterrichtspraxis in Einklang zu bringen, macht die Beschreibung aus Australien deutlich:

„Der Prozess schließt ein:

- das Einsetzen einer Kommission für jeden Kurs,
- die Entwicklung eines Leitfadens und von Spezifizierungen,
- die Auswertung der Tests in Bezug auf den Inhalt des Lehrplans und der Testvorgaben,
- die Entwicklung von Bewertungsrichtlinien,
- die Bestimmung von Korrektoren – es werden mehrere Tausend Korrektoren für die Auswertung aller Higher School Certificate Prüfungen benötigt,
- die Überprüfung der Bewertungsrichtlinien durch Pilotierung,
- die Gewährleistung der Genauigkeit und Konsistenz der Bewertungen während des gesamten Prozesses,
- die Einordnung der Schülerleistung in Leistungsbereiche,
- die Überprüfung dieser Einordnung durch Expertenkommissionen “ (Hafner, 41).

Besonders im Bereich der Naturwissenschaften erhält die Frage der Validität eines Testinstrumentes in Bezug auf die Unterrichtspraxis eine besondere Bedeutung, wenn es darum geht, auch den experimentell-praktischen Unterrichtsanteil abzudecken. Es erscheint intern auf der Klassenebene in diesem Zusammenhang einfacher, valide Instrumente zu erstellen, als auf einer extern gesteuerten Ebene, die eher Papier- und Bleistift-Verfahren vorsieht (s. Hafner, 2007). Doch auch in letzterem Fall sind vielfältige Formate für Testaufgaben möglich, die den jeweiligen kognitiven Anforderungen angepasst sein sollten, „multiple-choice oder einfache Antwortformate für die Wiedergabe von Faktenwissen; umfangreichere, offene Aufgaben, wenn es um das Lösen von Problemen oder die Anwendung von Wissen geht.“ (Bryce, 258)

Einige Länder nutzen pragmatisch Daten aus bestehenden groß angelegten Studien wie TIMMS oder PISA (Australien für achte bzw. neunte Klasse). Andere Nationen wie Portugal entwickeln Aufgaben, die an die PISA Aufgaben angelehnt sind (s. Galvao et al., 2007). Gelegentlich werden Test-Datenbanken mit qualitativ hochwertigen und inhaltlich validen Aufgaben zur Verfügung gestellt. „Die Lehrkräfte nutzen diese Tests gern für die Leistungsüberprüfung in ihren Klassen, da sie gut an die Ziele des Curriculums angepasst sind.“ (Lavonen, 112).

3.2.1 Zentral oder lokal organisiert

Wie oben angesprochen, führen viele Staaten Standards zentral ein. Oft werden sie dann auch zentral überprüft, wie in England oder in den einzelnen Staaten der USA. Auch Deutschland, Dänemark oder Taiwan haben sich dafür entschieden. International scheint der Trend in diese Richtung zu gehen.

Unter der Prämisse „devolution of decision power and responsibility to the local level“ (Lavonen, 102) hat Finnland dagegen einen anderen Weg gewählt. „Es gibt keine nationalen Abschlussprüfungen am Ende

der Schulpflichtzeit. Die Gesetzgebung verpflichtet die Schulverwaltung gemeinsam mit Lehrkräften auf lokaler Ebene dazu, selbst ihre Bildungsanstrengungen zu evaluieren und sich an externen Evaluationen zu beteiligen, wie sie z.B. stichprobenartig vom National Board of Education organisiert werden.“ (Lavonen, 111). In Anbetracht der anhaltenden exzellenten Ergebnisse von Finnland bei den PISA Studien kann mit Interesse beobachtet werden, ob dieses Modell auch für andere Nationen zukunftsweisend sein wird.

3.2.2 Externe oder interne Überprüfung

Bei externen Überprüfungen ist die Lehrkraft in der Regel an dem gesamten Testverfahren nicht beteiligt. Sie und die Lernenden scheinen das in Ländern wie England oder Frankreich zu akzeptieren, in denen seit Jahren zentral und extern Abschlussprüfungen gesteuert werden.

Ein Schwachpunkt dieses Vorgehens ist, dass Besonderheiten der Schulen, deren Ausstattung und die vorherrschenden Arbeitsbedingungen höchstens eingeschränkt berücksichtigt werden können. Staaten wie New South Wales bevorzugen wegen ihrer höheren Validität interne Überprüfungen, nicht zuletzt um in den naturwissenschaftlichen Fächern auch praktische Fertigkeiten im Test einen Platz einräumen zu können (s. Hafner, 2007).

Allerdings lassen sich bei internen Prüfverfahren die Ergebnisse mit denen anderer Schulen schlechter vergleichen. Um diese Reliabilität zu verbessern, treffen sich in England zum Beispiel Lehrkräfte in Arbeitskreisen, um mit der gemeinsamen Bewertung mitgebrachter Schülerarbeiten ihr Beurteilungsvermögen regelmäßig zu ‚kalibrieren‘ (s. Millar, 2007). In Finnland werden auf nationaler Ebene Kriterien vorgegeben, die dann intern umgesetzt werden. Um sowohl Validität als auch Reliabilität sicher zu stellen, setzen Länder wie Australien sowohl externe als auch interne Testverfahren ein.

In den Niederlanden werden in den naturwissenschaftlichen Fächern der nationale Test und die Durchschnittsnote aus dem schulinternen Portfolio, welches eine Reihe von Aktivitäten zur Überprüfung verschiedener Kompetenzen abdeckt, zu gleichen Anteilen gewertet.

3.2.3 Inhalt oder Kompetenzen

Sollte sich die Leistungsüberprüfung an Inhalten oder an Kompetenzen orientieren – oder an beiden? Diese Frage scheint in vielen Ländern noch nicht zufriedenstellend beantwortet. Für überregionale externe Tests werden wegen ihrer besseren Handhabbarkeit Papier-und-Bleistift Tests bevorzugt, vielfach auch im multiple-choice Format. Häufig sind solche Tests eher inhaltslastig, d.h. sie neigen zum Abfragen angelernten Wissens. Andererseits ist weltweit die Tendenz zu beobachten, dass Standards über Kompetenzen definiert werden. Dafür sind Wege zu finden, wie auch diese zu überprüfen sind. Die wiederkehrende Diskussion um die Qualität der PISA-Items macht deutlich, dass es schwieriger wird, Standards zu überprüfen, je mehr sie sich an Kompetenzen orientieren. In vielen Ländern gibt es derzeit Bemühungen um valide, zuverlässige und zugleich handhabbare Testinstrumente, mit denen über das Abfragen von Wissen hinaus naturwissenschaftliche Kompetenzen gemessen werden können.

3.3 Welche Funktion hat die Überprüfung?

Methode, Organisation und Ausführung der Überprüfung werden stark von der ihr zugeordneten Rolle beeinflusst. Generell wird zwischen *high-stakes* und *low-stakes assessment* unterschieden. *Low-stakes* Überprüfungen dienen im Sinne eines Feedbacks als Basis für die Analyse von Stärken und Schwächen Einzelner, einer Schule oder eines Systems und unterstützen eine Entwicklung

und Verbesserung. Schüler mit schlechten Leistungen bekommen Nachhilfe. Lehrkräfte, dessen Klassen unterdurchschnittliche Leistungen zeigen, bekommen die Möglichkeit einer Beratung.

High-stakes Überprüfungen dienen der Einleitung von Konsequenzen für die betreffende Zielgruppe. Sie sind oft mit Begutachtungen und Selektionsprozessen verbunden. Schüler mit ausreichenden Leistungen werden in die nächste Klassenstufe versetzt. Lehrkräfte mit unterdurchschnittlichen Leistungen können mit Sanktionen bis hin zur Entlassung rechnen. Schulen, die wiederholt schlecht in Vergleichsstudien abschließen, können unter Aufsicht gestellt oder gar geschlossen werden. Das englische Wort ‚accountability‘ wird oft in diesem Zusammenhang benutzt, wenn die Zielgruppe der Untersuchung für ihre Leistung zur Rechenschaft gezogen wird. In Australien wird die High-stakes Testung auch als ‚assessment of learning‘ bezeichnet im Gegensatz zum ‚assessment for learning‘ (s. Hafner, 2007). In jedem Fall ist es notwendig, die beabsichtigte Funktion der Überprüfung zu kommunizieren, oder besser noch, wie es Labudde (2007) fordert, alle beteiligten Ebenen am Prozess der Definition von Zielen und Funktion der Tests zu beteiligen.

3.3.1 Überprüfungen als Feedback für Verbesserungen

Gewöhnlich sehen die Autoritäten die Überprüfung der Standards als Mittel zur Leistungssteigerung und möchten sie eher als Herausforderung und weniger als Bedrohung verstanden wissen. „Tests versorgen die Regierung mit den erforderlichen Informationen zur Überwachung der Standards, zur Kontrolle spezieller Programme und zur Entscheidung über die Zuweisung von Ressourcen. Schulen und Lehrkräfte bekommen Informationen über die Leistung der Schule insgesamt, ihrer Klassen oder einzelner Schüler, mit denen sie ihre Ressourcen steuern und das Lernen im Unter-

richt unterstützen können“ (Hafner, 43). Im Idealfall wirkt dieses konstruktive Konzept im Sinne einer formativen Evaluation für die weitere Überarbeitung des Curriculums oder der Unterrichtspraxis (s. Lederman & Lederman, 2007).

Es ist nicht zu bestreiten, dass derartige Datenerhebungen auch immer für Vergleiche oder Wettbewerbe herangezogen werden können. Im Interesse fairer Vergleiche sollten dabei immer auch die Bedingungen berücksichtigt werden, unter denen in den Schulen gelernt wird. „Wenn die Lernkultur durch Standards verbessert werden soll – wie die Schulbehörden stets betonen – dann müssen auch ‚opportunity-to-learn Standards‘ eingeführt werden (Klassengröße, Anteil ausländischer Schüler, Ausstattung der Klassenräume, Lage der Schule, Selektionsgrad, etc.)“ (Weiglhofer, 67).

Überprüfung als Feedback für Verbesserungen kann auf allen Ebenen eingesetzt werden. In Schweden wirken diese Ebenen eng zusammen: „Nationale Tests werden als Hilfe für die Lehrkräfte verstanden, wenn sie die Absichten des Lehrplans umsetzen und den Lernerfolg ihrer Schüler evaluieren wollen ... Auf lokaler Ebene bieten lokale, von Lehrern durchgeführte und durch freiwillige nationale Diagnoseinstrumente und ‚Testbanken‘ unterstützte Tests gute Möglichkeiten zur Durchsetzung nationaler Bildungsziele“ (Strömdahl, 275).

In Finnland ist die Einstellung zu Standards überprüfenden Tests ähnlich. „Mit Hilfe dieser Informationen können Schulen, die unterdurchschnittlich abschneiden, die erforderlichen Schritte zur Abhilfe einleiten. Die Evaluationsergebnisse werden nicht benutzt, um Schulen mit schlechteren Leistungen abzustempeln und zu bestrafen“ (Lavonen, 111).

3.3.2 Überprüfung des Systems

Landesweite Tests dienen oft einem Überblick über die Leistung der Lernenden im ganzen Land oder einem Vergleich verschiede-

ner Regionen untereinander. In Deutschland übernimmt das IQB diese Aufgaben (Rupp & Vock, 2007). Eine Hauptaufgabe dieses *System Monitoring* ist die Überprüfung, ob das Bildungssystem die gesetzten Ziele der nationalen Standards erreichen kann. Ob dafür alle Lernenden z.B. einer vorgegebenen Klassenstufe oder nur eine repräsentative Stichprobe einer vorgegebenen Altersstufe befragt werden sollte, wird unterschiedlich diskutiert. „Wenn man wissen will, was einzelne Schüler oder eine Schule leisten, müssen alle Schüler getestet werden. Wenn es um Informationen für das Bildungssystem als solches geht, genügen Stichproben der Gesamtpopulation“ (Malléus, 140).

3.3.3 Überprüfung für Harmonisierung

In Ländern mit föderaler Struktur kann eine landesweite Überprüfung helfen, Unterschiede zwischen den einzelnen Staaten oder Regionen auszugleichen. Für die Schweiz ist diese Harmonisierung zwischen den 26 Kantonen eine treibende Kraft der Bildungsreform, die sich auch in dem Projekt zur Reform HarmonS (Harmonisierung obligatorische Schule) widerspiegelt.

3.3.4 Überprüfung der Einhaltung des Curriculums

In Ländern mit verbindlichen nationalen Curricula dienen Teststrategien mit formativem Charakter oft auch dem Bemühen, eine Kohärenz zwischen dem intendierten und dem erreichten Curriculum sicher zu stellen. So bestärkt in England das externe Testverfahren die Bemühungen der Lehrkräfte, das gesamte Curriculum abzudecken, da sie nicht wissen, welche Bereiche in dem jeweiligen Abschlussjahr abgeprüft werden (s. Millar, 2007).

3.3.5 Überprüfung als eine Basis für den Leistungsvergleich

„Ganz offensichtlich – und offiziell – dienen Abschlusszsuren am Gymnasium der Selektion.“ (Dolin, 79). So deutlich wie in Dänemark werden Funktionen im Sinne einer High-Stakes Testung selten formuliert. In Deutschland sollen die vom IQB begleiteten Tests keine Konsequenzen für den einzelnen Lernenden haben, jedoch werden in einzelnen Bundesländern eigene Leistungsüberprüfungen entwickelt, die auf den Standards basieren und zu einem Teil in die Benotung bei Abschlussprüfungen einfließen sollen (s. Schecker & Parchmann, 2007). Auch für die Niederlande wurde oben bereits ein solches System beschrieben. In New South Wales erachtet die Regierung externe Überprüfung sogar für unabdingbar, um ein unabhängiges und faires Urteil über die Leistung der Lernenden in der Oberstufe zu erhalten (s. Hafner, 2007). In Ländern wie Schweden oder Schottland, in denen Ergebnisse aus externen Überprüfungen nicht für die Beurteilung der individuellen Schülerleistung heran gezogen werden, dienen diese objektiven Ergebnisse den Lehrkräften jedoch für die interne Überprüfung der eigenen Einschätzungen.

Natürlich kann die erhobene Schülerleistung auch als ein Indikator für die Qualität der Lehre heran gezogen werden. Dies gilt zum Beispiel in den Niederlanden am Ende der Abschlussklasse. In einigen Fällen können diese Vergleiche ernsthafte Folgen haben, von der Verweigerung einer Beförderung bis hin zur Gehaltsreduzierung.

Auch für Schulen können die Schülerleistungen folgenreich sein. So ist es in England oder Holland aber inzwischen auch in einigen deutschen Bundesländern nicht unüblich, Ergebnisse aus Vergleichsstudien zu veröffentlichen, um damit unter anderem die Entscheidung der Eltern für die Schulwahl ihrer Kinder zu beeinflussen. Obwohl, wie oben aufgeführt, die extern erhobenen Daten nicht unbedingt valide sind, wenn sie z.B. unterschiedliche lokale, sozio-ökono-

mische Faktoren nicht mit abbilden, scheint es manchen Bildungspolitikern verlockend zu sein, diese Ergebnisse zu nutzen, um den Wettbewerb zwischen den Schulen zu steigern (Schecker & Parchmann, 2007). In optimistischen Szenarien führen schwache Ergebnisse zu unterstützenden Maßnahmen, wie in Schweden, Finnland oder den Niederlanden. Im schlimmsten Fall können – wie aus einigen Staaten der USA berichtet – wiederholt schlechte Ergebnisse dazu führen, dass die Schulleitung ihre Autorität verliert oder letztendlich die Schule geschlossen wird.

3.4 Welche Nebenwirkungen werden von der Standardüberprüfung befürchtet bzw. beobachtet?

Neben den intendierten Zielen der Standards und deren Überprüfungen werden aus den einzelnen Ländern negative aber auch positive (mögliche) Auswirkungen der Maßnahmen berichtet.

3.4.1 Teaching to the test

Unter diesem Slogan werden negative Auswirkungen verstanden, wenn das gesamte System eines Landes Erfolgskriterien nur auf die Test-Ergebnisse reduziert. Lehrkräfte sehen sich dann verpflichtet, ihren Unterricht überwiegend auf zufriedenstellende Testresultate ihrer Schüler auszurichten, statt sich von pädagogischen Prinzipien leiten zu lassen. Aus Ländern mit einer längeren Tradition von Bildungsstandards wird gelegentlich von diesbezüglichen Problemen berichtet. (z.B. Sacks, 1999; Volante, 2004)

Andere sehen in der Überprüfung der Standards eher eine Chance für die Verbesserung der Unterrichtskultur. „Teaching-to-the-test ist sogar beabsichtigt!“ (Labudde, 292). Es ist nur eine Frage, was abgeprüft werden soll. Wenn die Tests darauf ausgerichtet sind, z.B. das Planen, Durchführen, Auswerten und Interpretieren eines Experi-

ments als naturwissenschaftliche Teilkompetenz zu überprüfen, dann sind die Lehrkräfte aufgefordert, ihre allzu oft lehrerzentrierten, auf Wissensvermittlung ausgerichteten Unterrichtsaktivitäten anzupassen. Teaching-to-the-test muss also nicht zwangsläufig ein angeborener Fehler der Einführung von Bildungsstandards sein.

3.4.2 Externe Tests betonen schlichte Lernziele

Wie oben ausgeführt, ist es schwieriger, Standards zu überprüfen, je abstrakter sie formuliert sind und je mehr sie auf Kompetenzen fokussieren. Das Problem, komplexe Lernziele in Tests zu fassen wird z.B. aus der Erfahrung in den USA beschrieben: „Probleme gibt es nicht mit der Überprüfung der Standards an sich, sondern damit, dass die Testverfahren für den naturwissenschaftlichen Unterricht nicht mit den gegenwärtig gültigen Bildungszielen Schritt gehalten haben“ (Lederman & Lederman, 359). Dies gilt noch stärker für allgemeine Bildungsziele, die gar nicht mehr in den Blick einer fachspezifischen Überprüfung genommen werden (s. Labudde, 2007). Es besteht die Gefahr, dass Tests vor allem für das entwickelt werden, was bequem messbar ist (lower-level skills) und dass sich der Unterricht vor allem darauf konzentriert. Was also fehlt sind gute Testinstrumente, die komplexe Kompetenzen überprüfen können.

3.4.3 Standards können Innovationen hemmen

Ein erwünschter Effekt für die Einführung von Standards ist eine gewisse Vereinheitlichung und Standardisierung der Unterrichtspraxis, mit dem Ziel, dadurch insgesamt die Qualität zu steigern. Dem können aber auch originelle Methoden zum Opfer fallen (s. Bryce, 2007). Konstruktivistische Unterrichtsansätze betonen alternative Me-

thoden wie das Rollenspiel, Portfolios oder Projektarbeit. In England sind Schulen jedoch zurückhaltend, von der gewohnten Unterrichtspraxis abzuweichen, um nicht schlechtere Testergebnisse zu riskieren (s. Millar, 2007).

3.4.4 Flächendeckende Tests laden zu fragwürdigen Verwendungen der Ergebnisse ein

Überregionale Tests führen zu großen Datenmengen, die für viele gute Zwecke verwendet werden können. „Leider sind die Bildungsbehörden eher geneigt, diese Daten für die Erstellung von Ranglisten zu missbrauchen, als dass Lehrer sie für die Anpassung ihres Unterrichts an die Bedürfnisse der Schüler nutzen“ (Malléus, 139), wird aus Frankreich geklagt. Solche Ranglisten für Schulen scheinen in vielen Ländern geführt zu werden. Sie bleiben jedoch fragwürdig, wenn Hintergrundvariablen wie sozio-ökonomische Faktoren nicht mit in Betracht gezogen werden. Millar schlägt ein anderes Beurteilungssystem vor, in dem für eine Schule Testergebnisse auf Basis der bekannten Hintergrundvariablen vorausgesagt werden. „Wenn die tatsächlichen Resultate einer Schule besser sind als die Vorhersage, dann hat sie ‘added-value’ produziert; die Schüler haben besser abgeschnitten, als auf Grund der Eingangsdaten zu erwarten war. Andererseits kann eine Schule auch Ergebnisse hervorbringen, die unter den Erwartungen liegen. Sie kann die Testergebnisse für interne Planungen verwenden – und die Daten würden eine faire Grundlage für Schulvergleiche bieten.“ (Millar, 1995). Mit diesem Auswertungsverfahren könnten auch stärker allgemein pädagogische Fähigkeiten der Schule berücksichtigt werden.

3.5 Können Bildungsstandards für die Naturwissenschaften ohne den Zwang zur zentralen Überprüfung implementiert werden?

Es scheint weitestgehend akzeptiert, dass in irgendeiner Weise Standards, allgemeine Ziele oder nationale Curricula allen an den Bildungsprozessen Beteiligten vorgeben, welches Wissen oder welche Kompetenzen im Bereich Naturwissenschaften zu welchem Zeitpunkt von den Lernenden erwartet werden. Genauso scheint eine Überprüfung dessen akzeptiert zu sein. Ungeklärt ist, ob diese Überprüfung wirklich zentral und extern durchzuführen ist. Diese Frage ist besonders in den Ländern von Bedeutung, die gerade Standards einführen.

Die Antwort dürfte davon abhängen, welche Funktion den Standards zugeordnet ist. Sind die Absichten auf Vergleichbarkeit gerichtet, so scheinen zentrale Überprüfungen unverzichtbar. Ist Förderung das Hauptaugenmerk, dann wäre ein stärker lokales, schülerorientiertes Testverfahren zu empfehlen (s. Weiglhofer, 2007). Für diesen Fall wäre die Überprüfung durch die jeweilige Lehrkraft einem externen Test vorzuziehen, da diese die Lernleistung, die Lernprozesse und individuelle Förderung der Schüler kontrollieren kann. Allerdings wäre bei einem solchen Vorgehen der Vergleich zwischen Klassen oder gar Schulen schwierig und die Kohärenz mit den durch die Standards festgelegten intendierten Ergebnissen nicht sicher gestellt.

Es gibt jedoch Möglichkeiten, dem zu begegnen. So böte eine stärkere Kooperation zwischen Lehrkräften einer Schule oder verschiedener Schulen eine Überwindung der Subjektivität und eine Anpassung der Urteilsfähigkeit der Lehrkräfte (s. Millar, 2007). Die Lehrkräfte könnten mit einer Datenbank von gut überprüften Test-Items unterstützt werden. Diese können direkt eingesetzt werden oder den Lehrkräften als Musteraufgaben zur Entwicklung eigener Aufgaben dienen. In einigen Ländern gibt es bereits solche Datenbanken, die zum Teil auch repräsentative

Referenzwerte zu den Items liefern. Die Leistung der eigenen Klasse im Vergleich zu dem Referenzwert gibt der Lehrkraft so eine objektive Rückmeldung über den Leistungsstand der Klasse.

Neben einzelnen Items aus einer Datenbank könnten den Schulen auch komplette Tests mit den dazu gehörigen Referenzwerten angeboten werden. Diese könnten zentral entwickelt, organisiert und evaluiert werden, aber bei freiwilliger Teilnahme könnten einige der angesprochenen Nebenwirkungen vermieden werden. Diese Art der zentralen aber freiwilligen Überprüfung findet in vielen Ländern hohe Akzeptanz.

Auch wenn Kontrolle und Steuerung eines ganzen Systems im Vordergrund stehen, erfordert das nicht zwingend flächendeckende, zentrale Tests. „Es wäre vermutlich ökonomischer und effektiver, jedes Jahr Stichproben von Schülern aus Stichproben von Schulen zu testen, als alle gleichzeitig einer externen Prüfung zu unterziehen“ (Millar, 98).

4 Zusammenfassung

Standards werden in der Bildung aus verschiedenen Gründen gesetzt. Am häufigsten dienen sie dem Monitoring des Bildungssystems, einem Mechanismus für Verantwortung (accountability), der Harmonisierung regional unterschiedlicher Bildungsverhältnisse sowie einem Feedback für Schulen und Lehrkräften zur Verbesserung der Unterrichtspraxis.

Standards und Curricula sind eng miteinander verknüpft, es herrscht aber Uneinigkeit in der Reihenfolge der Entwicklung. In Australien und vielen anderen Ländern bestimmt das Curriculum die Standards. Deutschland und einige andere Nationen gehen den anderen Weg. Sie geben Standards vor und entwickeln Curricula, die es den Lernenden ermöglichen sollen, die Standards zu erreichen.

Während einige Standards über Inhalte definiert werden, ist ein Trend in Richtung kompetenzbasierter Standards zu beobachten.

Kompetenzen werden verschieden definiert und auch Kompetenzmodelle werden in unterschiedlicher Form angedacht.

Standards werden in der Regel für eine bestimmte Klassenstufe abschließend beschrieben, nur wenige Länder definieren Standards für mehrere oder gar für jede Klassenstufe. Es herrscht weitestgehend Einigkeit darüber, dass es eine Überprüfung der Standards geben muss. Damit diese Überprüfung bedeutungsvoll ist, ist die Passung von Standards, Curriculum und Testverfahren ein notwendiges aber nur mit hohem Aufwand zu erreichendes Kriterium. Trotz der bekannten Defizite sind Papier-und-Bleistift Tests die gängige Praxis. Die Defizite werden besonders dann deutlich, wenn die Standards kompetenz-basiert formuliert sind. Stärker an der Praxis orientierte Tests, die besonders für den naturwissenschaftlichen Unterricht eine höhere Validität bedeuten würden, werden aus Gründen der schwierigen Umsetzbarkeit vermieden.

Große Unterschiede sind in der Frage zu beobachten, ob Tests zentral oder lokal, extern oder intern organisiert werden sollen. Es scheint eine Tendenz in Richtung einer nationalen und externen Testung zu geben, eine Testform, die allerdings die meisten Nebeneffekte hervorruft. Teaching-to-the-test ist das am häufigsten genannte Problem.

Alles in allem bestehen keine Zweifel darüber, dass eine Kontrolle des Outputs von Bildungsbemühungen notwendig ist und dass das Setzen und Überprüfen von Standards eine sinnvolle Methode für diesen Zweck ist. Allerdings müssen dafür die erforderlichen Ressourcen bereit gestellt werden, und die Lehrkräfte müssen in die Reform mit eingebunden werden. Mit der Verkündung neuer Standards ist es nicht getan. Die Überwindung vertrauter Routinen und die Veränderung gut etablierter Methoden verlangt „umfassende, kontinuierliche professionelle Weiterbildung über einen längeren Zeitraum“ (Lederman & Lederman, 369). Bleibt diese Unterstützung aus, dann werden die Lehrkräfte auch diese Reform ohne weitere Veränderungen überstehen – und geduldig auf die nächste warten.

Literatur

- Ackeren, I. v. (2005). Vom Daten- zum Informationsreichtum? Erfahrungen mit standardisierten Vergleichstests in ausgewählten Nachbarländern. *Pädagogik* (5), 24-28.
- Altrichter, H. & Posch, P. (2005). Bildungsstandards als Teil eines Qualitätskonzepts. *forum schule* (7), 125-140.
- Arnold, R. (2005). Die PISA-Lüge. *Friedrich Jahresheft*, XXIII, 65-66.
- Bachmann, G., Dangl, O., Mayr, J. & Spenger, J. (2004). zukunft: schule? Stellungnahme aus der Sektion „Empirische Pädagogische Forschung“ (O.E.P.F.) zum „Reformkonzept der österreichischen Zukunftskommission“. *ÖFEB-Newsletter*, 1, 5-14.
- Baumann, R., Fessler, W. & Willmann, J. (2005). Lehrlings-, Fach- und Expertenniveau. Vom Umgang mit Standards in einem sozial belasteten Umfeld. *Friedrich Jahresheft*, XXIII, 39-41.
- Bayrhuber, H., Bögeholz, S., Graf, Hammann, M., Harms, U., Hößle, C., Krüger, D., Lude, A., Mayer, J., Riemeier, T., Sandmann, A., Schlüter, K., Unterbruner, Upmeier zu Belzen, S., Vogt, H., & Ziemek (Hrsg.). (2005). *Bildungsstandards Biologie. Tagungsband Sektion Biologiedidaktik im VDBiol*. Kiel: IPN.
- Benner, D. (2007). *Bildungsstandards – Instrumente zur Qualitätssicherung im Bildungswesen. Chancen und Grenzen – Beispiele und Perspektiven*. Paderborn: Ferdinand Schöningh.
- Bennett, J., Holman, J., Millar, R. & Waddington, D. (Eds.). (2005). *Making a difference. Evaluation as a tool for improving science education*. Münster: Waxmann.
- BMBF (Ed.). (2003). *Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Staaten*. Bonn: BMBF.
- Bonnet, G. (2002). Schulentwicklung in Frankreich. *Pädagogik* (4), 44-48.
- Böttcher, W. (2003). Besser werden durch Leistungsstandards? *Pädagogik* (4), 50-52.
- Böttcher, W. (2004). Bildungsstandards und Kerncurricula – Potenzielle, intendierte und nichtintendierte Effekte eines zentralen Reformprojektes. In J. Schlömerkemper (Hrsg.), *Bildung und Standards* (Vol. Die Deutsche Schule, 8. Beiheft, pp. 231- 244). Weinheim: Juventa.
- Brügelmann, H. (2003). In fünf Jahren ... Über Kerncurricula, Bildungsstandards und Leistungstests. *Neue Sammlung*, 43, 235-237.

- Bryce, T. (2007). Standards in science education: The situation in Scotland. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 255-262). Münster: Waxmann.
- Bybee, R. (2002). Scientific Literacy – Mythos oder Realität? In W. Gräber, P. Nentwig, T. Koballa & R. Evans (Eds.), *Scientific Literacy. Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung*. Opladen: leske+budrich.
- Chiu, M.-H. (2007). Standards for science education in Taiwan. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 303-346). Münster: Waxmann.
- Demmer, M. & Schweitzer, J. (2005). Es fährt ein Zug nach nirgendwo ... *Friedrich Jahresheft*, XXIII, 68-69.
- Dolin, J. (2007). Science education standards and science assessment in Denmark. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 71-82). Münster: Waxmann.
- Dubs, R. (2007). Education Standards: attempting a critical judgement of development – a reflection of the symposium. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 405-430). Münster: Waxmann.
- Galvao, C., Reis, P., Freire, A. & Oliveira, T. (2007). Science curriculum in Portugal: From the development to the evaluation of students' competences. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 237-254). Münster: Waxmann.
- GEW (2003). *Nationale Bildungsstandards – Wundermittel oder Teufelszeug*. Available: <http://www.gew.de/Binaries/Binary3666/bildungsstandards.pdf> [2005, 14.02.].
- GEW (2004a). „Zweite Chance für die KMK“ – Gute Bildungsstandards benötigen ein Konzept, Zeit, wissenschaftlichen Sachverstand und Akzeptanz. In J. Schlömerkemper (Ed.), *Bildung und Standards* (Vol. Die Deutsche Schule, 8. Beiheft, pp. 152-183). Weinheim: Juventa.
- GEW (2004b). *Hoffnungsträger IQB?!* Available: http://www.le-ser.ch/_library/Fichiers_PDF/harmos/harmos_prise_de_position_gew.pdf [2005, 12.02.].
- Gröben, A. v. d. (2003). Wird den Lehrern das Denken abgenommen? *Pädagogik*(3), 11-14.
- Hafner, R. (2007). Standards in science education in Australia. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 23-60). Münster: Waxmann.
- Herzog, W. (2007). Pro und Kontra Bildungsstandards. Die Perspektive eines Skeptikers. In P. Labudde (Hrsg.), *Bildungsstandards am Gymnasium – Korsett oder Katalysator?* (pp. 57-64). Bern: h.e.p verlag ag.
- Horstkemper, M. (2005). Standards. Vermessungspädagogik oder Antrieb zur Verbesserung der Bildungsqualität? *Pädagogik* (9), 6-9.
- Huber, L. (2004). Nationale Standards und Gymnasiale Oberstufe. In J. Schlömerkemper (Hrsg.), *Bildung und Standards* (Vol. Die Deutsche Schule, 8. Beiheft, pp. 145-151). Weinheim: Juventa.
- Jauhiainen, K. (2005). Pädagogische Standards für finnische Schulen. *Pädagogik* (9), 16-19.
- Jürgens, E. (2004). Pädagogische Implikationen der KMK-Entwürfe für Bildungsstandards. In J. Schlömerkemper (Hrsg.), *Bildung und Standards* (Vol. Die Deutsche Schule, 8. Beiheft, pp. 48-65). Weinheim: Juventa.
- Klieme et al., E. (2003). *Zur Entwicklung nationaler Bildungsstandards – Eine Expertise*. Berlin: BMBF, DIPF.
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Z. f. Päd.* (54)2, 222-237
- Köller, O. (2007). Bildungsmonitoring und Vergleichsarbeiten in Deutschland. In P. Labudde (Hrsg.), *Bildungsstandards am Gymnasium – Korsett oder Katalysator?* (pp. 65-73). Bern: h.e.p verlag ag.
- Kremer, A. (2004). *Stellungnahme zum Entwurf der KMK „Bildungsstandards Physik für den mittleren Schulabschluss (Jahrgangsstufe 10)“*. Available: <http://www.gew.de/Binaries/Binary6015/bista-kremer-Physik.pdf> [2005, 25.08.].
- Kullmann, H. (2004). *Standardisiert die Bildungsstandards!* Eine Stellungnahme zu den Entwürfen der nationalen Bildungsstandards für den Mittleren Schulabschluss in den Fächern Biologie, Chemie und Physik. Available: <http://www.gew.de/Binaries/Binary6016/bista-kullmannNaWi.pdf> [2005, 01.10.].
- Labudde, P. (Hrsg.) (2007a). *Bildungsstandards am Gymnasium – Korsett oder Katalysator?* Bern: h.e.p verlag ag.
- Labudde, P. (2007b). How to develop, implement and assess standards in science education? 12 challenges from a Swiss perspective. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 277-302). Münster: Waxmann.
- Labudde, P. (2007c). Naturwissenschaftliche Bildung: Quo vadis? In P. Labudde (Hrsg.), *Bildungsstandards am Gymnasium – Korsett oder Katalysator?* (pp. 283-291). Bern: h.e.p. verlag ag.

- Lange, B. (2005). Bildungsstandards und Unterrichtsplanung – Konsequenzen für didaktisches Denken. *Lehren und Lernen* (5), 3-10.
- Lavonen, J. (2007). National science education standards and assessment in Finland. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 101-126). Münster: Waxmann.
- Lederman, N. G. & Lederman, J. (2007). Standards for science education in the United States: Necessary Evil? In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 347-371). Münster: Waxmann.
- Malléus. (2007). Science education standards in France. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 127-146). Münster: Waxmann.
- Mamluk-Naaman, R. (2007). "Science and technology for all" – an Israeli curriculum based on new standards in science education. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 199-220). Münster: Waxmann.
- Messner, R. (2004). Was Bildung von Produktion unterscheidet. In J. Schlömerkemper (Hrsg.), *Bildung und Standards* (Vol. Die Deutsche Schule, 8. Beiheft, pp. 26-48). Weinheim: Juventa.
- Millar, R. (2007). How standards in science education are set and monitored in the English education system. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 83-100). Münster: Waxmann.
- National Research Council (Ed.). (1996). *National Science Education Standards*. Washington, DC: National Academy Press.
- Nentwig, P. & Waddington, D. (Eds.). (2005). *Making it relevant. Context based learning of science*. Münster: Waxmann.
- Neuweg, G. H. (2004). Bildungsstandards in Österreich. *Pädaktuell* (2), 4-13.
- Oelkers, J. (2004). *Bildungsstandards und Schulentwicklung: Ein Blick in Geschichte und Zukunft*. Available: http://www.paed.unizh.ch/ap/downloads/oelkers/Vortraege/127_MuensterStandards.pdf [2005, 25.07.].
- Orth, G. (2005). Eine Chance für die Qualität schulischer Arbeit. *Friedrich Jahresheft*, XXIII, 63-65.
- Parchmann, I. (2005). Grundlagen für ein Verständnis. Standards als Impuls für eine Veränderung von Chemieunterricht. *Friedrich Jahresheft*, XXIII, 93-95.
- Plöger, W. (2004). Bildungsstandards in bildungstheoretischer Sicht. In J. Schlömerkemper (Hrsg.), *Bildung und Standards* (Vol. Die Deutsche Schule, 8. Beiheft, pp. 11-25). Weinheim: Juventa.
- Ratzki, A. (2005). Finnland. Erfahrungen mit Bildungsstandards und individuellen Fördermaßnahmen. *Friedrich Jahresheft*, XXIII, 50-52.
- Regenbrecht, A. (2005). Sichern Bildungsstandards die Bildungsaufgabe der Schule? *PÄD Forum* (1), 16-22.
- Rosengren, T. (2004). *Unethically teaching to the test*. Available: <http://www.districtadministration.com/pageprint.cfm?p=622> [2005, 14.02.].
- Rupp, A. A. & Vock, M. (2007). National educational standards in Germany: Methodological challenges for developing and calibrating standards-based tests In D. Waddington & P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 173-198). Münster: Waxmann.
- Saldern, M. v. & Paulsen, A. (2003). *Die nationalen Bildungsstandards für den Mittleren Schulabschluss der Kultusministerkonferenz im Vergleich zu den Vorschlägen des Gutachtens „Zur Entwicklung nationaler Bildungsstandards“ „Klieme-Gutachten“ und den Erkenntnissen nach PISA*. Universität Lüneburg: Arbeitsstelle Schulentwicklung. Available: <http://www.gew.de/Binaries/Binary3781/Vergleich.pdf> [2005, 25.08.].
- Schecker, H. & Parchmann, I. (2007). Standards and competence models: The German situation. In D. Waddington & P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 147-165). Münster: Waxmann.
- Spinner, K. (2005). Der standardisierte Schüler. *Friedrich Jahresheft*, XXIII, 88-91.
- Stadtfeld, P. (2005). Nationale Bildungsstandards – Problem oder Chance für die Schulentwicklung? *Lehren und Lernen*(5), 33-37.
- Strömdahl, H. (2007). On standards in science education in the contemporary Swedish school system. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 263-276). Münster: Waxmann.
- Tenorth, H.-E. (2005). Auch eine Konvention bedarf der Rechtfertigung. Legitimationsprobleme bei Bildungsstandards. *Friedrich Jahresheft*, XXIII, 30-31.
- Tenorth, H.-E. (2008). Bildungsstandards außerhalb der „Kernfächer“. Herausforderungen für den Unterricht und die fachdidaktische Forschung. *Z. f. Päd.* (54)2, 159-162

- Voelzke, U. (2004). *Zusammenfassende Stellungnahme zu den Bildungsstandards in den Naturwissenschaften*. Available: <http://gew.de/Binaries/Binary6019/bista-voelzkeBiChPhNW.pdf> [2005, 14.02.2005].
- Waddington, D., Nentwig, P. & Schanze, S. (2007). *Making it comparable. Standards in Science Education*. Münster: Waxmann Verlag.
- Weiglhofer, H. (2007). Austria at the beginning of the way to standards in science. In D. Waddington, P. Nentwig & S. Schanze (Eds.), *Making it comparable. Standards in science education* (pp. 61-70). Münster: Waxmann.
- Zedler, P. (2007). Vernachlässigte Dimensionen der Qualitätsentwicklung und Qualitätssicherung von Unterricht und Schule, Erziehung und Bildung. In D. Benner (Ed.), *Bildungsstandards – Instrumente zur Qualitätssicherung im Bildungswesen. Chancen und Grenzen – Beispiele und Perspektiven* (pp. 61-82). Paderborn: Ferdinand Schöningh.

Kontakt

Prof. Dr. Sascha Schanze
Leibniz Universität Hannover
Naturwissenschaftliche Fakultät
IDN – Institut für Didaktik der Naturwissenschaften
Fachgebiet Chemiedidaktik
Bismarckstraße 2
30173 Hannover
schanze@chemiedidaktik.uni-hannover.de

Autoreninformation

Sascha Schanze ist Professor für Chemiedidaktik an der Leibniz Universität Hannover. Forschungsschwerpunkte sind das forschende Lernen im naturwissenschaftlichen Unterricht sowie die Integration neuer Medien als Lernwerkzeuge für den Chemieunterricht.

Peter Nentwig ist Mitarbeiter der Abteilung Chemiedidaktik am Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN). Arbeitsschwerpunkte sind PISA, Bildungsstandards für den naturwissenschaftlichen Unterricht und Kontext orientiertes Lehren und Lernen im naturwissenschaftlichen Unterricht.