

Peter Labudde (Herausgeber)

## **Bildungsstandards am Gymnasium**

**Korsett oder Katalysator?**



**PHBern**  
Pädagogische Hochschule



Bildung  
Medien  
Kommunikation

[www.hep-verlag.ch](http://www.hep-verlag.ch)  
der bildungsverlag

Peter Labudde (Hrsg.)  
**Bildungsstandards am Gymnasium – Korsett oder Katalysator?**  
ISBN 978-3-03905-339-1

Satz: Suzanne Castelberg, Biel  
Umschlagbild: Andrea Stuker, Bern

Bibliografische Information der Deutschen Bibliothek  
Die Deutsche Bibliothek verzeichnet diese Publikation in der  
Deutschen Nationalbibliografie; detaillierte bibliografische Angaben  
sind im Internet unter <http://dnb.ddb.de> abrufbar.

1. Auflage 2007  
Alle Rechte vorbehalten  
Copyright © 2007 h.e.p. verlag ag, Bern

h.e.p. verlag ag  
Bildung.Medien.Kommunikation  
Brunngasse 36  
CH-3011 Bern

[www.hep-verlag.ch](http://www.hep-verlag.ch)

## Erfahrungen mit Standards in England und den Niederlanden

Isabell van Ackeren

Die öffentliche Verantwortung für die Qualität der Schule begründet das länderübergreifende Interesse, Bildungsziele als Ergebniserwartungen zu erreichen und sicherzustellen. Schulen attestieren eine öffentlich erbrachte Leistung, die anschließenden Systeme – z. B. im Falle gymnasialer Bildungsgänge die Universitäten – müssen sich auf die Vergleichbarkeit der Zertifikate verlassen können. Dies gilt umso mehr in Staaten mit einem heterogenen föderalen System und gestärkter einzelschulischer Autonomie. Zur Qualitätsentwicklung und Vergleichbarkeitssicherung wurden deshalb in vielen Ländern – oft in Ergänzung zentraler Abschlussprüfungen (z. B. in Form des Zentralabiturs) – laufbahnbegleitende Lernstandserhebungen implementiert, die wissenschaftlich fundiert die Möglichkeit der relativen Positionsbestimmung von Schulen bezüglich einer Vergleichsgruppe bieten. Über diesen sozialen Vergleichsmaßstab hinaus stellt sich zugleich die Frage nach den kritischen Bezugsnormen: Es geht nicht allein um die Frage, was Schüler und Schülerinnen tatsächlich können, sondern was sie – normativ gesetzt – können sollen, wobei sich beide Bezugsmaßstäbe dann wechselseitig beeinflussen, wenn Erwartungen hoch, aber realistisch sein sollen.

Die im deutschsprachigen Raum so bezeichneten «Bildungsstandards» lassen sich vor diesem Hintergrund länderübergreifend als verbindliche Festlegungen charakterisieren, die solche zentralen Kompetenzbereiche beschreiben, die Schü-

ler und Schülerinnen bis zu einer bestimmten Jahrgangsstufe erworben haben sollen. Ein zentraler Unterschied zu Lehrplänen ist die mit Standards verbundene Kontrollabsicht, die empirische externe Überprüfung der tatsächlichen Zielerreichung. Damit verknüpft ist die Idee, Lernprozesse vom Ende her zu denken und die Aufmerksamkeit von Lehrenden und Lernenden auf langfristig angelegtes, kumulatives Lernen zu lenken und – z. B. mit Blick auf das Gymnasium – schulförmenspezifische aber auch -übergreifende Kernaufgaben zu reflektieren.

Vor diesem Hintergrund können Ländervergleiche helfen, allgemeine Entwicklungstrends zu identifizieren und aus Erfahrungen zu lernen, die sowohl Chancen als auch Gefahren von Reformelementen in ihrem Zusammenwirken offenbaren. Exemplarisch werden aufgrund ihrer unterschiedlichen Steuerungsphilosophien England und die Niederlande gewählt. Die folgende Darstellung von Standards als Elemente grundlegender Steuerungsphilosophien stützt sich auf Literatur- und Materialrecherchen sowie in den Untersuchungsländern durchgeführte Expertenbefragungen (van Ackeren 2003; Arbeitsgruppe Internationale Vergleichsstudie 2003; weitere Länderstudien finden sich u. a. bei Klieme et al. 2003 sowie Huber et al. 2006). Die länderspezifische Situations- und Problembeschreibung erfolgt jeweils nach einem gemeinsamen Raster:

- von der Skizzierung der Steuerungsphilosophie,
- über eine Betrachtung der Ausgestaltung von Standards,
- bis hin zu beschriebenen Effekten.

Daran schließt sich der Vergleich an, der in einer vorsichtigen Skizzierung einer sinnvollen Steuerungspolitik mündet.

## England

England steht für die Einbindung von Standards in ein zentralisiertes Prüfungswesen innerhalb einer marktorientierten Schulentwicklung.

### Steuerungsphilosophie

In England wurde Ende der 80er-Jahre ein Wettbewerb zwischen Schulen initiiert, um deren Leistungsfähigkeit zu entwickeln. Das 1988 eingeführte *National Curriculum* gibt verbindliche Kerninhalte und darauf abgestimmte «Leistungsziele» (*Attainment Targets*) vor, deren Erreichung durch viele obligatorische und optionale Tests geprüft wird. Die Ergebnisse werden regelmäßig publiziert, auch in Form von Leistungsrankings. Die grundsätzliche Schulwahlfreiheit soll von Eltern dazu genutzt werden, sich bei der Schulwahl an diesen Daten zu orientieren. Gute Schulen – so die Theorie – ziehen mehr Schüler und Schülerinnen an und erhalten aufgrund der Pro-Kopf-Finanzierung mehr finanzielle Ressourcen. Schlechte Schulen sollen sich verbessern (van Ackeren 2003 und 2006).

Die deutliche Leistungssteigerung bei internationalen Leistungsstudien im Verlauf der 90er-Jahre (van Ackeren 2002) wurde vielfach auf die neue Steuerungsphilosophie zurückgeführt. Mittlerweile geht man eher von einer Schockwirkung dieses reinen *Pressure-Ansatzes* (als ständiges Unterdrucksetzen der Schulen) aus, jedoch ohne langfristige Effekte erzielt zu haben. So gehört England z. B. weiterhin zu den Ländern, in denen Lernerträge besonders eng an die soziale Herkunft gekoppelt sind (PISA-Auswertungen; weiterhin Tymss 2004).

Mittlerweile wird zunehmend stärker darauf gesetzt, externe Evaluationen als Entwicklungsinstrument zu begreifen und Schulen in benachteiligten Regionen mittels zusätzlicher Ressourcen zu unterstützen (*Support-Ansatz*). Zudem will die Schulinspektion OFSTED (*Office for Standards in Education*) zukünftig stärker darauf setzen, Selbstevaluation in den Schulen anzuregen und mit der externen Überprüfung zu verknüpfen, wengleich dies insbesondere auch eine kostensenkende Maßnahme ist. (OFSTED 2007).

### Standards: Begriffsverständnis und Ausgestaltung

Die Einführung eines *National Curriculum* (NC) bedeutete eine erhebliche Einengung der schulischen Freiräume, wengleich die traditionellen zentralen Examina vermutlich immer schon den Unterricht normiert haben. Das NC schreibt etwa 80% der Lerninhalte für die 5- bis 16-Jährigen vor. Gleichzeitig werden die Mindestanforderungen an eher konkret formulierte Leistungsziele (*Attainment-Targets*) benannt, sodass inhaltliche und leistungsbezogene Standards als *Content* und *Performance Standards* gemeinsam aufgeführt werden (NC 2007)<sup>1</sup>. Die Standards beziehen sich auf vier Stufen, die mehrere Jahrgangsstufen umfassen (*Key Stages, KS*). Für die KS sind wiederum im Schwierigkeitsgrad steigende Leistungsniveaus auf einer 8-Punkte-Skala definiert (*Level Descriptions*), die die Leistungen im Verhältnis zu den Zielvorgaben kontinuierlich abbilden sollen. Am Ende von KS 1 (7 Jahre) sollte ein Anspruchsniveau von Level 2, zum Abschluss von KS 2 (11 Jahre) Level 4 usw. erreicht sein. In den Kernfächern *Englisch*, *Mathematik* und *Science* (Letzteres erst ab KS 2) wird dies über von der *Qualifications and Curriculum Authority* (QCA 2007a) entwickelte, schulexterne Tests abgefragt. Dabei werden auch fachübergreifende Kompetenzen berücksichtigt.

Die in den NC-Tests eingesetzten Aufgaben, die die Standards operationalisieren, werden in einem aufwendigen *Standard-Setting-Verfahren* auf der Grundlage von Expertenratings und statistisch ausgewerteter Pretestdaten entwickelt (Quinlan/Scharaschkin 1999; Tymms 2004).

Darüber hinaus werden die «Standard»-Erwartungen als Merkmale guter Schulen und guten Unterrichts von OFSTED beschrieben, die sich als *Opportunity-to-learn-Standards* bezeichnen lassen und in ihrer Erreichung bei den Schulvisitationen überprüft werden (Thomas/Peng 2004).

<sup>1</sup> Dort finden sich auch Arbeitsmaterialien sowie Schülerarbeiten auf den unterschiedlichen curricularen Niveaus.

## Beschriebene Effekte

Die Wirkungen der NC-Standards sind an das Testsystem geknüpft. Dabei wird einerseits von nicht-intendierten *Short-Term-Strategien* berichtet, mit denen Schulen gute Ranking-Plätze zu erzielen suchen (Gleeson/Husbands 2001):

- Dazu gehört ein verstärktes *Teaching-to-the-test*, indem der Unterricht auf das fokussiert wird, was erwartet und abgefragt wird. Jenkins (2000) konnte z. B. für den *Science-Unterricht* zeigen, dass signifikant weniger Zeit auf praktische Aktivitäten, z. B. im Labor, verwendet wird und sich die Lehrkräfte im Vergleich zurzeit vor der Einführung des NC darin eingeschränkt fühlen, einen differenzierenden und motivierenden Unterricht zu gestalten.
- Weiterhin wird der permanente Unterrichtsausschluss störender Schüler und Schülerinnen beschrieben, deren Verhalten das Lernen einer Klasse und schließlich das Ranking beeinflussen könnte.
- Die Konzentration auf Schüler und Schülerinnen mittlerer Leistungsfähigkeit – zum Vorteil besserer Rankingsplätze, jedoch zum Nachteil leistungsstarker und leistungsschwacher Schüler und Schülerinnen.
- Schließlich wird berichtet, dass Examenskandidaten nicht zur Abschlussprüfung zugelassen werden, wenn es unmöglich scheint, «*high grades*» zu erreichen.

Andererseits konnte in Untersuchungen mit Fallstudiencharakter gezeigt werden, dass es Schulen gibt, die rückgemeldete Daten dazu verwenden, Stärken und Schwächen in den Jahrgangsstufen fachspezifisch im Zeitverlauf zu identifizieren, diese in Ziele für Fachgebiete auf Schul-, Klassen- und Gruppenebene (Anteile erfolgreich lernender Jungen, Mädchen, Migrantenkinder ...) sowie in die curriculare Planung und in die Personalentwicklung einzubauen. Zudem werden die Resultate in das Schulentwicklungsprogramm aufgenommen, vielleicht auch deshalb, weil dies vom Schulinspektorat verlangt und abgefragt wird (Ashby/Sainsbury 2001; Kirkup et al. 2005). Konkrete Beispiele der Nutzbarmachung von Testergebnissen finden sich auf der QCA-Website (QCA 2007b).

Schulen sind offensichtlich besonders produktiv im Umgang mit Daten, wenn die – durch die unterschiedlichen lokalen Schulbehörden verständlich und benutzerfreundlich aufbereiteten – Kennzahlen geeignet sind, die eigene Praxis konkret zu hinterfragen und wenn die Schulen entsprechende Interpretationshilfen und Fortbildungen erhalten hatten (Saunders/Rudd 1999).

## Niederlande

Die Niederlande stehen für den Versuch, Qualität und Vergleichbarkeit in einem auf Schulautonomie setzenden Land u. a. über wirkungsbezogene Standards zu sichern.

### Steuerungsphilosophie

Es gibt keine staatlichen Lehrpläne als umfängliche Stoffpläne und somit kaum inhaltsbezogene Input- und wenig Prozesssteuerung. Auf der anderen Seite sollen Qualität und Vergleichbarkeit auf nationaler Ebene durch bestimmte Verpflichtungen garantiert werden: Die Erstellung von Schulprogrammen ist obligatorisch und wird vom Schulinspektorat überprüft. Zudem werden 50% der Abschlussprüfungsaufgaben in allen Schulformen zentral gestellt und ausgewertet; die anderen 50% liefert die Schule, um dem Schulautonomiegedanken gerecht zu werden. Darüber hinaus gibt es eine obligatorische, hinsichtlich des Testmaterials jedoch einzelschulisch frei wählbare Lernstandsüberprüfung zum Ende der Primarstufe im Alter von 12 Jahren (zumeist wird der Test des *Centraal Instituut voor Toetsontwikkeling* (CITO) eingesetzt), die u. a. der Zuweisung der Schüler und Schülerinnen auf die Schulformen des gegliederten Systems dient, sowie eine Vielzahl an Lernzuwachstests für schulinterne Evaluationsarbeit in den unterschiedlichen Schulstufen und nationale Stichprobenerhebungen (van Weeren 2007; CITO 2007). Schließlich gibt es ein Schulinspektionssystem.

Seit 1997 werden die bis dahin von der Schulaufsicht vertraulich behandelten Leistungs- und Inspektionsdaten, deren Publikation von der Presse unter Berufung auf das Recht der Informationsfreiheit eingeklagt wurde, veröffentlicht. Mittlerweile versuchen die Behörden, ein differenziertes Leistungsbild zu vermitteln (Inspectie van het Onderwijs 2007). Gleichwohl haben die benannten Elemente zusammen mit der traditionellen Schulwahlfreiheit – ohne dies zu intendieren – bildungsmarktähnliche Elemente geschaffen und den Entwicklungsdruck auf Schulen erhöht.

### Standards: Begriffsverständnis und Ausgestaltung

Bis 1993 gab es keine einheitlichen Zielvorgaben für verschiedene Fächer. Um Qualität und Vergleichbarkeit zunächst in der gemeinsamen Grundbildung zu sichern, wurden Kernziele (*kerndoelen*) für die Primarstufe eingeführt, um ein Basisniveau für die Sekundarstufe zu sichern. Mittlerweile gibt es auch *kerndoelen* für die ersten beiden Jahre der Sek. I, die schulformübergreifend angelegt sind und deren Erreichung etwa 75% der Unterrichtszeit ausmachen soll (Eurybase 2006; Kennisnet Ict op school 2007). Dies bedeutet z. B. für den gymnasialen Bildungsgang (*VWO*) eine Neuorientierung des eigenen Bildungsverständnisses, indem nun stärker die Anschlussfähigkeit zu anderen Bildungsgängen berücksichtigt werden muss.

Die Kernziele beschreiben als normative Festlegung, was Schüler und Schülerinnen wissen und können müssen, ohne den Schulen Vorgaben zur Erreichung der Ziele zu machen; es gibt jedoch umfangreiche Unterstützungsangebote (Schulbegleitedienste). Die Erfahrungen mit Kernzielen zeigten, dass diese schwierig auf die tägliche Unterrichtspraxis zu übertragen sind und Angaben zu Kompetenzniveaus fehlen. Vor diesem Erfahrungshintergrund wurden Zwischenziele für Kernfächer entwickelt, die in ihrer Zusammenschau sogenannte «Lehrlinien» abbilden, um eine integrative Sicht auf Lehr- und Lernprozesse zu vermitteln.

Die Erreichung der Kernziele wird im Rahmen des CITO-Test sowie durch nationale, stichprobenartig angelegte Lernstandserhebungen (*peilingsonderzoek*) in unterschiedlichen Domänen (einschließlich Weltkunde, Leibes- und Verkehrserziehung u. Ä.) im Rahmen mehrjähriger Untersuchungspläne gemessen.

Neben der Aufgabenentwicklung erfolgt auch die Bestimmung der Leistungsniveaus über Expertenbefragungen, die die Testitems, die auch praktische Aufgaben umfassen, den drei Niveaus – ausreichend, minimal und fortgeschritten – zuordnen, wobei auf der Grundlage von Pretests bereits bekannt ist, wie viel Prozent der Schüler und Schülerinnen die Items richtig lösen (van Weeren 2007). Es handelt sich um eine Mischung aus normativen Standarderwartungen und faktischen Leistungen.

Die Datenanalyse der Stichprobenerhebungen konzentriert sich z. B. auf Zusammenhänge zwischen Leistung und ethnischer bzw. sozialer Herkunft und dient in erster Linie der Generierung bildungspolitischen Steuerungswissens. Darüber hinaus gibt es Broschüren für die Schulverwaltung und Schulen zur Vermittlung von Orientierungswissen.

In der Sekundarstufe dienen die teilzentralen Abschlussprüfungen der Standardüberprüfung.

### Beschriebene Effekte

Es gibt «bislang keine aussagekräftigen Studien, in denen den Rückmeldungen aus den nationalen Lernstandserhebungen bestimmte Effekte zugesprochen werden» (van Weeren 2007). Visscher (2001) geht davon aus, dass die global gehaltenen Indikatoren den Schulen keine konkreten Verbesserungsmöglichkeiten aufzeigen. Van den Akker und Letschert (2002) ergänzen, dass der Glaube an eine umfassende *Top-down-Reform* zunehmend geringer werde und es vielmehr einer Initiierung und nachhaltigen Gestaltung von Schulentwicklungsprozessen von innen heraus bedürfe. Diese Einschätzung wird durch einen aktuellen Bericht des Schulinspektorats untermauert, wonach jede vierte Grundschule potenzielle Kandidaten und Kandidatinnen des niederländischen Hauptschulbildungsgangs nicht am Grundschultest teilnehmen lässt und jede dritte Schule die zur landesweiten Publikation verwendeten Ergebnisse dieser Schülergruppe nicht mitteilt. In großen Städten mit erheblicher Konkurrenz zwischen den Schulen sind dies sogar 63 % der Schulen.

Aus der Lehrer- und Schülerperspektive besonders handlungsrelevant erscheint hingegen das *leerlingvolgsysteem* (Schülerbegleitsystem), das sich ebenfalls an den Kernzielen orientiert, und mit dem schülerindividuelle Leistungsentwicklungen diagnostiziert und abgestimmte Materialien bei identifizierten Schwächen eingesetzt werden können – und tatsächlich genutzt werden (Roeleveld et al. 2002).

### Vergleichendes Fazit

Der Vergleich zeigt in aller Kürze, dass sich beide Länder nicht allein auf die Formulierung von – an kritischen und sozialen Bezugsnormen orientierten – Ergebniserwartungen verlassen, sondern die Auseinandersetzung mit ihnen bzw. ihre Erreichung über Lernstandsmessungen und die Beschäftigung mit den Resultaten im Kontext von Schulinspektionen in den Blick nehmen. Darüber sollen zum einen zuverlässige Informationen über die Vergleichbarkeit und Qualität von Lernerträgen als Ausgangspunkt möglicher schulsystemischer Steuerungsstrategien erzeugt werden. Zum anderen sollen Standards in Schulen transportiert werden, um sie auf einzelschulischer Ebene handlungsleitend werden zu lassen. Dies ist vor dem Hintergrund wichtig, dass Lehrkräfte vor allem für die Gestaltung von Lehr-Lernprozessen und kaum für die Formulierung von Zielen im Sinne von Standards und schon gar nicht für ihre angemessene Überprüfung, die testpsychologischen Gütekriterien genügt, ausgebildet sind.

Die englische Steuerungsphilosophie ist dabei eine ausgesprochen kontrollierende, die von der Erwartung getragen ist, dass Standards ihre Orientierungsfunktion erst über die sanktionsbegleitete Einforderung öffentlicher Rechenschaftslegung erhalten. In den traditionell sehr autonomen Niederlanden versucht man, Outputorientierung und Schulautonomie zu verknüpfen, indem Tests stichprobenartig angelegt sind und Abschlussprüfungen teilzentral gestaltet sind. Diagnostische, nicht publizierte Tests ergänzen das Programm und dienen der individuellen Lernförderung im Klassenzimmer. Gleichwohl deuten sich an den Stellen, wo Schulen über öffentliche Vergleiche unter Druck gesetzt werden, ähnliche *Short-Term-Strategien* wie in England an.

Insgesamt erscheint die Kopplung an eine Vielzahl an *High-Stakes-Tests*, die mit Selektionsentscheidungen und Ergebnispublikationen verknüpft sind, problematisch: Die methodische Engführung des Unterrichts sowie die Konzentration auf fachliche Kernbereiche sind befürchtete, teils in Erfahrungsberichten beschriebene, allerdings auch kaum systematisch analysierte Effekte.

Insbesondere macht der Blick auf die exemplarisch gewählten Länder deutlich, dass der produktive Umgang mit Standards und Tests, der – wie Fallbeispiele zeigen – möglich ist, angeleitet und unterstützt werden muss. Aber offensichtlich können solche Reforminitiativen nicht auf direktem Wege eine «Evaluationskultur» im Umgang mit extern generierten Daten initiieren, um darüber Qualität zu entwickeln. Vielmehr scheint umgekehrt zu gelten, dass eine bereits vorhandene Evaluationskultur (z. B. effektive interne Kommunikationsstrukturen, ein

Schulklima der Kooperation und Innovation, das Professionswissen der Schule insgesamt etc.) den Umgang mit Rückmeldungen bestimmt. Dies weist auf die immense Bedeutung von Unterstützung beim Aufbau einer vom Kollegium breit getragenen Evaluationskultur hin, die bei Fortbildungsmaßnahmen stärker in den Blick zu nehmen ist.

Reformen bedürfen der Stärkung der Kompetenz, den eigenen Wandel zu organisieren und dabei an die Professionalität (innovationsbereiter!) Lehrkräfte anzuknüpfen. In diesem Sinne hat Wales – angesichts der erheblichen Zweifel an der Wirksamkeit des am englischen Nachbarn orientierten Testsystems – die Abschaffung desselben zum Schuljahr 2008/2009 angekündigt. Die Überprüfung der Standards, die weiterhin Bestand haben, wird sich dann auf teilzentrale Abschlussprüfungen sowie insbesondere auf die Arbeit der Schulinspektion konzentrieren. Somit steht dem Ausbau outputorientierter Steuerung, z. B. in den deutschsprachigen Ländern, die teilweise Zurücknahme dieser Elemente bzw. die Konzentration auf prozessorientierte, unterstützende Steuerung in anderen, test- und standarderfahrenen Ländern gegenüber.

## Literatur

- Ackeren, Isabell van (2002): Von FIMS und FISS bis TIMSS und PISA. Schulleistungen in Deutschland im historischen und internationalen Vergleich. In: *Die Deutsche Schule*, Jg. 94, Heft 2, S. 157–175.
- Ackeren, Isabell van (2003): Evaluation, Rückmeldung und Schulentwicklung. Erfahrungen mit zentralen Tests, Prüfungen und Inspektionen in England, Frankreich und den Niederlanden. Münster: Waxmann.
- Ackeren, Isabell van (2006): Freie Wahl der Grundschule? Zur Aufhebung fester Schulbezirke und deren Folgen. In: *Die Deutsche Schule*, Jg. 98, Heft 3, S. 301–310.
- Akker, Jan van den/Letschert, Jos (2002): Schulentwicklung in den Niederlanden. In: *Pädagogik*, Jg. 54, Heft 12, S. 10–15.
- Arbeitsgruppe Internationale Vergleichsstudie (Hrsg.) (2003): Vertiefender Vergleich der Schulsysteme ausgewählter PISA-Teilnehmerstaaten. Berlin: Bundesministerium für Bildung und Forschung (BMBF). <http://www.bmbf.de/pub/pisa-vergleichsstudie.pdf> (Abruf: 4.5.2007).
- Ashby, John/Sainsbury, Marian (2001): How Do Schools Use National Curriculum Test Results? Slough: NFER. [http://www.nfer.ac.uk/publications/other-publications/downloadable-reports/pdf\\_docs/AICreport.PDF](http://www.nfer.ac.uk/publications/other-publications/downloadable-reports/pdf_docs/AICreport.PDF) (Abruf: 4.5.2007).
- CITO (Centraal Instituut voor Toetsontwikkeling) (2007): <http://www.cito.nl> (Abruf: 4.5.2007).
- Eurybase (2006): Het onderwijssysteem in Nederland. [http://www.eurydice.org/portal/page/portal/Eurydice/DB\\_Eurbase\\_Home](http://www.eurydice.org/portal/page/portal/Eurydice/DB_Eurbase_Home) (Abruf: 4.5.2007).
- Gleeson, Denis/Husbands, Chris (2001): The performing school. Managing teaching and learning in a performance culture. London und New York: Routledge.
- Huber, Christina/Späni, Martin/Schmellentin, Claudia/Criblez, Lucien (2006): Bildungsstandards in Deutschland, Österreich, England, Australien, Neuseeland und Südostasien. Literaturbericht zu Entwicklung, Implementation und Gebrauch von Standards in nationalen Schulsystemen. Aarau: Fachhochschule Nordwestschweiz. [http://www.edk.ch/PDF\\_Downloads/Harmos/Literaturanalyse\\_1.pdf](http://www.edk.ch/PDF_Downloads/Harmos/Literaturanalyse_1.pdf) (Abruf: 3.5.2007).
- Inspectie van het Onderwijs (2007): <http://www.kwaliteitskaart.nl> (Abruf: 4.5.2007).
- Jenkins, Edgar W. (2000): The impact of the national curriculum on secondary school science teaching in England and Wales. In: *International Journal of Science Education*, Jg. 22, Heft 3, S. 325–336.
- Kennisnet Ict op school (2007): <http://kerndo.clen.kennisnet.nl/> (Abruf: 4.5.2007).
- Kirkup, Catherine/Sizmur, Juliet/Sturman, Linda/Lewis, Kate (2005): Schools' use of data in teaching and learning. London: Department for Education and Skills (DfES).
- Klieme, Eckhard et al. (Hrsg.) (2003): Expertise zur Entwicklung nationaler Bildungsstandards. Berlin: BMBF.
- OFSTED (Office for Standards in Education) (2007): <http://www.ofsted.gov.uk> (Abruf: 4.5.2007).
- NC (National Curriculum for England) (2007a): <http://www.nc.uk.net> (Abruf: 4.5.2007).
- QCA (Qualifications and Curriculum Authority) (2007): <http://www.qca.org> (Abruf: 4.5.2007).
- QCA (2007b): Fallstudien In: <http://www.qca.org.uk/9117.html> (Abruf: 4.5.2007).
- Quinlan, Michael/Scharaschkin, Alex (1999): National curriculum testing. Problems and practicalities. Paper presented at the British Educational Research Association Annual Conference, Brighton: University of Sussex. <http://brs.leeds.ac.uk/~beiwww/beid.html> (Abruf: 04.03.2007).
- Roeleveld, Jaap/Otter, Martha E./Blok, Henk (2002): Leerlingvolgsystemen in de jaren 90: Secundaire analyses op gegevens uit PRIMA en IST. Amsterdam: SCO-Kohnstamm Instituut.
- Saunders, Lesley/Rudd, Peter (1999): Schools' use of «value added» data. Slough: NFER. <http://www.leeds.ac.uk/educol/documents/000001110.htm> (Abruf: 3.5.2007).
- Thomas, Sally/Peng, Wen-Jung (2004): The use of educational standards and benchmarks in indicator publications. In: *European Educational Research Journal*, Jg. 3, Heft 1, S. 177–212. [http://www.wwords.co.uk/pdf/viewpdf.asp?j=eerj&vol=3&issuc=1&year=2004&article=8\\_Benchmarking\\_Chapter\\_3\\_EERJ\\_3\\_1\\_web&cid=192.41.150.10](http://www.wwords.co.uk/pdf/viewpdf.asp?j=eerj&vol=3&issuc=1&year=2004&article=8_Benchmarking_Chapter_3_EERJ_3_1_web&cid=192.41.150.10) (Abruf: 3.5.2007).
- Tymms, Peter (2004): Are standards rising in English primary schools? In: *British Educational Research Journal*, Jg. 30, Heft 4, S. 477–494.

Visscher, Adrie J. (2001): Publieke schoolprestatie-indicatoren. De problemen op een rij. In: Anne Bert Dijkstra/Karsten, Sjoerd/Veenstra, René/Visscher, Adrie J. (Hrsg.): Het oog der natie. Scholen op rapport (S. 54–61). Assen: Van Gorcum.

Weeren, Jan van (2007): Wem nutzen Outputmessungen? Eine kritische Analyse ihrer Wirksamkeit und Nebeneffekte aus niederländischer Perspektive. In: Die Deutsche Schule, Jg. 99, Heft 2.