

Was bringen Bildungsstandards den Lernenden?¹

Urs Moser

Kompetenzzentrum für Bildungsevaluation und Leistungsmessung
an der Universität Zürich

Sehr geehrte Damen und Herren

Als mir die Organisatoren der Tagung diese Frage als Titel für ein Referat vorschlugen, war ich mit der Thematik sofort einverstanden. Als ich mir dann aber bei der Vorbereitung erste Gedanken zur Beantwortung der Frage machte, wurde ich skeptisch. Ist das nicht eine rhetorische Frage, quasi eine Testfrage für eine Person, die regelmässig Leistungstests entwickelt und sich dabei nicht überlegt, welches Unheil sie damit anrichten könnte? Beispielsweise Stress und Prüfungsangst bei den Schülerinnen und Schülern, oder schlimmer noch: Lehrpersonen, die aus Angst vor schlechten Testergebnissen mit ihren Klassen nur noch auf Testfragen büffeln, statt den Schülerinnen und Schülern die Möglichkeit zu bieten, sich konstruktiv mit allgemein bildenden Lerngegenständen auseinander zu setzen. Meinungen gegenüber Leistungstests, denen ich bei meiner Arbeit als empirischer Bildungsforscher schon lange vor PISA begegnet bin.

Skeptisch wurde ich auch, weil die Frage nach dem Nutzen für die Lernenden hoffentlich bereits vor längerer Zeit positiv beantwortet worden ist, schliesslich ist die Einführung von Bildungsstandards zumindest für die obligatorische Schulzeit beschlossene Sache. Und bei genauer Betrachtung scheint eine outputorientierte oder *Result-Based-Reform* (RBR), wie das Eva Baker vom National Center for Research on Evaluation, Standards and Student Testing (CRESST) von der University of California nennt, eine einfache Angelegenheit zu sein (Baker, 2004).

- Figure out what should be taught
- Be prepared to teach it
- Help students to learn it
- Measure their learning
- Continue the cycle until desired improvement is met

So what could be bad? (Baker, 2004, S. 3)

Einige Kantone der Schweiz haben sich diese Frage auch gestellt: Was ist schlecht daran, die Schülerinnen und Schüler zu testen und zu erfahren, welche Leistungen sie erbringen? Der Kanton Aargau beispielsweise mochte nach dem PISA-Schock nicht auf national ver-

¹ Referat an der Kadertagung des Schweizerischen Instituts für Berufspädagogik (SIBP) und der Schweizerischen Zentralstelle für die Weiterbildung der Mittelschullehrpersonen (WBZ) zum Thema «Bildungsstandards – wo steht die Sekundarstufe II?» vom 26./27. Januar in Murten.

ordnete Bildungsstandards und Leistungstests warten. Er hat für die fünften Primarschulklassen bereits Leistungstests angeordnet, zurzeit noch auf freiwilliger Basis, und deshalb wohl auch mit bemerkenswertem Erfolg. Im Herbst 2004 beteiligten sich rund zwei Drittel der angefragten Lehrpersonen an «Check 5», so die anglophile Bezeichnung für Tests in Mathematik, Deutsch, kooperativem Problemlösen und selbstreguliertem Lernen. Im Jahr 2005 waren es bereits 80 Prozent Freiwillige, und die Reaktionen der Lehrpersonen auf dieses Checken waren durchwegs positiv, wie folgendes Zitat belegt:

«Die persönlichen Stärken und Schwächen der Schülerinnen und Schüler von einer externen Seite beurteilt zu haben interessiert mich, um eine umfassende individuelle Förderung anstreben zu können.»

(Tresch & Moser, 2005, S. 46).

Da haben Sie die Antwort auf die Frage nach dem Nutzen aus erster Hand. PISA hat in der Schweiz allerdings mehr ausgelöst als kantonale Checks. Die mittelmässigen Ergebnisse im internationalen Schulleistungstest waren ohne Zweifel der eigentliche Grund, weshalb die ergebnisorientierte Reform zur nationalen Angelegenheit wurde. Die Methodik des internationalen Schulleistungsvergleichs bildet zudem eine wichtige Grundlage für das Vorgehen in einem föderalistischen, mehrsprachigen Land, sowohl bei der Entwicklung als auch bei der Überprüfung von Standards. Ich werde deshalb in einem ersten Teil die Gemeinsamkeiten und die Unterschiede bei der Entwicklung von internationalen Schulleistungstests und nationalen Bildungsstandards im Hinblick auf den Nutzen für die Lernenden besprechen. Im zweiten Teil gehe ich auf die wesentlichen Besonderheiten der Reform in der Schweiz ein, um abschliessend drei Punkte zu erwähnen, die beachtet werden müssen, sofern die Reform den Lernenden etwas bringen soll.

1. Gemeinsamkeiten und Unterschiede bei der Entwicklung von internationalen Schulleistungstests und nationalen Bildungsstandards

Drei Gemeinsamkeiten

1. Bildungsstandards konkretisieren die Ziele des Lehrplans in Form von Kompetenzanforderungen und legen fest, über welche Kompetenzen die Schülerinnen und Schüler verfügen müssen. Diese Anforderungen sollen in der Schweiz auf der Grundlage von empirisch validierten Kompetenzmodellen bestimmt werden. Ein Kompetenzmodell ist eine Beschreibung des kumulativen Aufbaus der Anforderungen, dient also der systematischen Ordnung der Kompetenz in hierarchischer Folge. Mit dem Begriff der Kompetenz wird ausgedrückt, dass Bildungsstandards – anders als Lehrpläne – nicht auf Listen von Inhalten zurückgreifen, sondern auf Fähigkeiten und Fertigkeiten zur Lösung verschiedenster Aufgabenstellungen und Probleme zielen, die sich in der Ausbildung und im Alltag stellen (Klieme, 2003). Die hierarchische Beschreibung der Kompetenzen anhand von abgestuften Niveaus erleichtert es zum einen, Forschungsergebnisse sinnvoll

zu interpretieren. Zum andern wird dadurch für die Lernenden Klarheit über die Anforderungen geschaffen.

2. Sowohl für den internationalen Vergleich als auch für die nationale Reform muss ein allgemeines Verständnis der Kompetenzen gefunden werden. Begriffe müssen definiert und übersetzt werden, so dass alle Beteiligten unter einer Kompetenz das Gleiche verstehen und die Kompetenzen auch in allen Sprachen entsprechend gemessen und ausgewiesen werden. Die Kantone, aber auch die Sprachregionen, ja sogar die Landesgrenzen sind irrelevante Kategorien, wenn es um die Beurteilung von Kompetenzen sowie um die Zertifizierung und die Schaffung von Transparenz geht. Dies hilft, die Ergebnisse des Bildungssystems besser zu verstehen und die Fähigkeiten der Schülerinnen und Schüler nachvollziehbar zu kommunizieren.
3. Standards können sowohl als Input – zur Orientierung der Vermittlung von Kompetenzen im Unterricht – wie auch als Output – zur Überprüfung der Ergebnisse des Unterrichts – genutzt werden (Herrmann, 2003). Die Überprüfung, inwieweit die Standards erreicht werden, ist dann sinnvoll, wenn dies – wie bei PISA – auf der Basis einer Testtheorie geschieht (Rost, 2004). Kompetenzen werden anhand von geeigneten Testaufgaben gemessen und somit konkretisiert und präzisiert. Die Bewertung der Qualität des Bildungssystems, aber auch die Beurteilung der Kompetenzen der Schülerinnen und Schüler wird objektiviert. Ausschliesslich auf Grund von Tests zu beurteilen kann nicht das Ziel sein, beurteilen ohne zu messen muss aber in Zukunft verhindert werden – sowohl im Interesse des Bildungssystems als auch im Interesse der Lernenden.

Drei Unterschiede

1. PISA geht zwar von einer kompetenzorientierten Grundbildung aus, wie sie in der Rahmenkonzeption für die Erfassung von Wissen und Fähigkeiten der OECD nachzulesen ist (OECD, 2000), nicht aber von einem vorgängig hierarchisch organisierten Kompetenzmodell. Die Hierarchie des Modells wurde im Nachhinein festgelegt und in Form von Stufen qualitativ umschrieben. Das Verfahren war nur beschränkt deduktiv. Beim Projekt HarmoS ist das geplante Vorgehen hoffentlich auch nur beschränkt deduktiv, und die Fachdidaktiker sind hoffentlich emotional in der Lage, an ihren postulierten Modellen einige Abstriche zu machen oder sie auf Grund der empirischen Validierung entsprechend anzupassen. Das Problem liegt allerdings weniger bei der erstmaligen Validierung von Kompetenzmodellen mit Hilfe eines Leistungstests als vielmehr im enormen Aufwand, das Kompetenzmodell mit einem zweiten Test zu replizieren. Bei PISA werden die unterschiedlichen Tests immer durch bewährte Testaufgaben verbunden («Link-Items»). Ein System, das methodisch anspruchsvoll ist und zudem verlangt, dass die Testaufgaben absolut vertraulich behandelt werden. Kommt kein solches Verfahren zur Anwendung, dann werden die Kompetenzen der Lernenden zwar gemessen, aber ohne Bezug zu Standards und Kompetenzmodellen. Denn es ist kaum möglich, dass *ein* Kompetenzmodell mit *zwei* Tests, die komplett verschieden sind, also nicht eine be-

stimmte Anzahl identischer Aufgaben enthalten, empirisch validiert werden kann. Diese methodische Herausforderung könnte dazu führen, dass zwar mit unterschiedlichen Testverfahren gemessen wird, andererseits aber auf den empirischen Nachweis des Zusammenhangs von Tests, Bildungsstandards und Kompetenzmodellen verzichtet wird. Konkret würde dies bedeuten: Fortsetzung der Beurteilung von Lernenden anhand von wenig transparenten und willkürlichen Massstäben.

2. PISA geht vom Konzept «Literacy» aus, und das ist die grosse Stärke des Programms, der Schlüssel zur Akzeptanz eines internationalen Schulleistungsvergleichs. Die Standards oder Kompetenzen werden rein funktional definiert, mehr oder weniger losgelöst von Lehrplänen und Lehrmitteln; sie orientieren sich nicht explizit an Bildungszielen, sondern am Nutzen fürs künftige Leben. Auch wenn das Konzept «Literacy» durchaus Teile einer humanistischen Bildung umfasst, gilt es zu beachten, dass es sich nur um eine *Auswahl* von Kompetenzen handeln kann – Kompetenzen, die mit einem Leistungstest erfasst werden können. Sofern Bildungsstandards im Sinne eines Outputs auch überprüft werden müssen, ist vermutlich aus ökonomischen Gründen eine Einschränkung auf das Grundlegende, auf absolute Basiskompetenzen, nicht zu umgehen. Vor allem bei den Sprachen wird es zudem kaum möglich sein, Bildungsstandards so zu formulieren, dass sie unabhängig von explizit erwähnten Inhalten, Texten oder ähnlichen sprachlichen Produkten verständlich kommuniziert und objektiv überprüft werden können. All dies ist kein Problem für die Lernenden, wenn sich die Bildungsstandards einerseits auf ein definiertes Minimum beschränken und im Unterricht eine Nebenrolle spielen, und wenn andererseits die Testergebnisse im schulischen Umfeld jenen Stellenwert erhalten, der ihnen gebührt. Genau dies wage ich allerdings zu bezweifeln. PISA hat aufgedeckt, dass sich weder die Politik noch die Wissenschaft an die Regeln der Statistik halten und Daten äusserst selektiv interpretieren. Zudem halten Lehrpersonen ihren eigenen Notenmassstab für den einzig richtigen, was es schwierig machen wird, sie davon zu überzeugen, dass Testergebnisse im Gegensatz zu ihren Noten mit einem ausgewiesenen Messfehler behaftet sind und entsprechend vorsichtige Interpretationen verlangen. Keine schönen Aussichten für die Lernenden.
3. Die Definition der Kompetenz nach Weinert (2001) schliesst sehr viel mehr ein, als beispielsweise mit einem PISA-Test gemessen wurde oder mit einem Leistungstest gemessen werden kann. Kompetenzen werden definiert als die bei Individuen verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten, bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösung in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können (Weinert, 2001). Eine solche Definition ist empirisch kaum umsetzbar, weil sich Motivationen in Tests integriert nicht messen lassen. Leistungstests definieren eine spezifische motivationale Situation, die sich nicht auf die Motivation einer Person verallgemeinern lässt. Es ist sehr zu hoffen, dass den Schülerinnen und Schülern nicht auch noch durch Standards vorgeschrieben wird, wie motiviert sie zu sein haben und was sie wollen müssen.

Weniger ist mehr, dies gilt nicht nur für die Ausrichtung der aktuellen ergebnisorientierten Reform, sondern auch für die damit verbundenen Erwartungen von Politik und Öffentlichkeit, womit ich beim zweiten Teil meiner Ausführungen, den Besonderheiten der Reform in der Schweiz angelangt bin.

2. Besonderheiten der Reform in der Schweiz

Die Schweiz plant die Einführung von Mindeststandards, um den Anteil an Risikoschülerinnen und -schülern, deren Grundbildung für einen reibungslosen Übergang in den Arbeitsmarkt nicht ausreicht, zu reduzieren. Angenommen, das Projekt HarmoS verläuft erfolgreich – Bildungsstandards werden lehrbuchmässig entwickelt und überprüft; die Erkenntnisse aus dem Bildungsmonitoring führen zu angemessenen Massnahmen; die Schülerinnen und Schüler mit ungenügender Grundbildung erhalten die Möglichkeit, ihre Versäumnisse nachzuholen – dann profitieren zum einen unsere Schule und die Gesellschaft, zum anderen aber auch die – eher schlechten – Schülerinnen und Schüler. Die Ausrichtung der Reform an Mindestkompetenzen ist richtig. Denn, wie Heinz-Elmar Tenorth festhält, unser Bildungssystem beziehungsweise unsere Gesellschaft kann es sich gar nicht leisten, Jugendliche aus der Schule zu entlassen, die die Mindeststandards nicht erreicht haben (Tenorth, 2004, S. 109f.).

Die Ausrichtung der Reform an Mindestkompetenzen ist aber auch gefährlich, vor allem weil die Bildungsstandards für die Schnittstellen formuliert werden. Die Probleme des schweizerischen Bildungssystems liegen tatsächlich bei den Schnittstellen. Die erste Selektion findet statt, bevor die Schule begonnen hat. Im Kanton Zürich schaffen rund 10 Prozent der Kinder den Eintritt in die reguläre erste Klasse nicht, rund 15 Prozent haben hingegen beim Schuleintritt bereits die Ziele der ersten Klasse im Lesen und in der Mathematik erreicht (Moser, Stamm & Hollenweger, 2005). Nach der zweiten Schnittstelle am Ende der Primarschule wird das Öffnen der Schere zwischen guten und schlechten Schülerinnen und Schülern forciert, die Weichen für die Schul- und Berufslaufbahn werden gestellt. Mit der Konzentration der Bildungsstandards auf die Schnittstellen laufen wir Gefahr, dass Tests nicht primär zur Förderung der Lernenden, sondern einzig zur Selektion eingesetzt werden. Selbstverständlich spricht nichts dagegen, die Anforderungen an den Schnittstellen, die heute nach Kanton, Gemeinde oder Schule sehr unterschiedlich sein können, konkret zu umschreiben und das Erreichen der Anforderungen zu überprüfen. Die Definition von Bildungsstandards für die Schnittstellen könnte aber auch zur Folge haben, dass unser nicht nur glücklich selektionierendes und von Stütz- und Förderangeboten dominiertes Bildungssystem in seiner Ausrichtung noch gestärkt wird. Im Kanton Zürich haben bereits am Ende der dritten Klasse über die Hälfte der Schülerinnen und Schüler eine Sonderklasse besucht oder ein Förderangebot in Anspruch genommen (Moser, Keller & Tresch, 2002). Bildungsmonitoring auf der Grundlage von Bildungsstandards ist ein wichtiges Instrument für die Bildungspolitik, weil es aufzeigt, welcher Anteil der Schülerinnen und Schüler die minimalen Ziele erreicht hat und ob allenfalls Anpassungen des Bildungssystems notwen-

dig sind. Bildungsstandards sind aber vor allem dann im Interesse der Lernenden, wenn sie nicht nur eine Erwartung an den Adressaten formulieren oder aufzeigen, wie die Schullaufbahn zu verlaufen hat, sondern wenn sie die Schulen gleichzeitig mit einer Bringschuld verpflichten (Tenorth, 2004, S. 106). Dieses Verantwortungsbewusstsein der Schulen ist sozusagen eine Voraussetzung dafür, dass die Reform den Lernenden einen Mehrwert bringt.

Ohne Zweifel helfen Bildungsstandards und Tests auch an den Schnittstellen, indem die Lernenden und die Lehrenden objektivierte Informationen über den Lernstand erhalten. Ihre Vorteile, nämlich Transparenz und die objektivierende Ergänzung der von subjektiven Fehlern behafteten Beurteilung durch die Lehrpersonen, sind unbestritten und schon aus Gründen der Fairness in Zukunft nicht mehr wegzudenken. Doch was nützt es, wenn ein Schüler erst dann erfährt, ob er den Standard erreicht hat, wenn er in seiner Schullaufbahn bei der Schnittstelle angekommen ist? Geradezu irrelevant für die Lernenden ist das Überprüfen der Bildungsstandards am Ende der Volksschule, weil sich die Jugendlichen bereits in der 8. Klasse für die zukünftige Schul- und Ausbildungslaufbahn bewerben müssen.

Während sich ein gesamtschweizerisches Bildungsmonitoring auf die Leistungsmessung an den Schnittstellen konzentrieren soll, müssen die Lernenden mit objektivierten Informationen *zwischen* den Schnittstellen unterstützt werden. Was das Schweizer Bildungssystem benötigt, sind nicht zwingend Instrumente, um die Schülerinnen und Schüler einzuteilen, sondern um sie zu fördern, ihnen aufzuzeigen, was sie können und können müssen, und sie dadurch zu motivieren.

Nun ist mir klar, dass die ergebnisorientierte Reform nicht nur Mindeststandards zur Folge haben wird und mehr als ein Instrument für die Verhinderung von zu schlecht ausgebildeten Schülerinnen und Schülern sein soll. So werden die Kantone die Möglichkeit haben, ihre eigenen Standards zu definieren und die Standards sollen gesamtschweizerisch einen grossen Harmonisierungseffekt auf Lehrpläne, Lehrmittel und Tests ausüben (Maradan & Mangold, 2005). Es ist zu hoffen, dass die Begeisterung für Kompetenzmodelle in der Praxis gross und die Ressourcen für die Validierung solcher Modelle vorhanden sein werden. Ist dies nicht der Fall, so wird es in der Schweiz kaum zu einer kantonsunabhängigen Transparenz und zur Standardisierung der Beurteilung der Lernenden kommen.

3. Ergebnisorientierte Reform zu Gunsten der Lernenden

«What could be bad?» war die rhetorische Frage von Eva Baker (2004) auf die fünf Schritte der ergebnisorientierten Reform. Was kann daran schlecht sein, Ziele zu operationalisieren, sie im Unterricht anzustreben und mit Tests zu überprüfen, Anpassungen im Unterricht vorzunehmen und diesen Zyklus so lange fortzusetzen, bis die Ziele erreicht sind? Wenn die ergebnisorientierte Reform den Lernenden tatsächlich etwas bringen soll, dann müssten nach Eva Baker (2004) folgende Punkte geklärt sein:

- benefits or sanctions
- goals or standards
- intended beneficiaries
- desired level of operational attainments equivalent to targets
- the rate of progress
- the inputs
- the resources
- the operations
- the measures
- the reported results (Baker, 2004, S. 4)

Der Nutzen und die mit der Reform verbundenen Sanktionen, Ziele und Standards, Nutzniesserinnen und Nutzniesser (beispielsweise Schülerinnen und Schüler, Eltern, Lehrpersonen, Behörden, Bildungspolitik), gewünschter Zielzustand und gewünschter Fortschritt (anhand des Kompetenzmodells), Input im Sinne von «opportunity-to-learn-standards» (beispielsweise Lehrmittel), Ressourcen (beispielsweise Weiterbildung), operative Durchführung (beispielsweise Umsetzung der Standards im Unterricht oder die Durchführung von Leistungstests), Testinstrumente sowie die Kommunikation der Ergebnisse. Ich gehe abschliessend auf drei Punkte von Eva Baker ein, die ich als wesentliche Voraussetzung dafür erachte, dass die Lernenden von Bildungsstandards und Tests profitieren.

1. *Zu den Sanktionen* – Die Einführung von Bildungsstandards und Leistungstests wird nicht ohne Wirkung bleiben. Damit Bildungsstandards zu besseren Kompetenzen bei den Schülerinnen und Schülern führen, sollten sie nicht mit Sanktionen verbunden werden. Amrein und Berliner (2002) konnten für die USA eindrücklich nachweisen, dass durch die Verbindung von Testergebnissen mit der Schullaufbahnplanung, mit der Mittelvergabe an Schulen oder gar mit der Beförderung der Lehrpersonen das Qualitätssystem untergraben wird. Je wichtiger ein quantitativer Indikator für Entscheidungen im sozialen Bereich wird, umso grösser ist die Wahrscheinlichkeit, dass der soziale Prozess verzerrt und korrumpiert wird. In den USA entwickeln sich die Schulleistungen in Staaten mit «high stake testing» (das heisst, die Testergebnisse sind mit Konsequenzen verbunden) statistisch signifikant schlechter als in Staaten mit Tests zur Qualitätssteigerung (Amrein & Berliner, 2002).
2. *Zum Unterricht* – Mit der Einführung von Bildungsstandards werden Auswirkungen auf den Unterricht erwartet, wobei diese Auswirkungen fälschlicherweise immer despektierlich mit «teaching to the test» gleichgesetzt werden. Am besten erforscht sind die Auswirkungen des Testens auf den Unterricht im Bereich des Sprachenlernens unter dem Begriff «Washback» (Rückfluss). Es gibt Studien, die einen negativen Effekt des Testens auf den Unterricht nachweisen, aber auch solche, die einen positiven Effekt nachweisen (Cheng, Watanabe & Curtis, 2004). Es liegt in der Natur der Sache, dass sich eine Person auf eine Prüfung oder einen Test vorbereitet. Es ist auch nicht falsch, wenn die Schülerinnen und Schüler von den Lehrpersonen auf die Prüfungen vorberei-

tet werden. Dies wurde immer gemacht und wird auch in Zukunft gemacht werden. Das «teaching to the test» wirkt sich erst dann für die Lernenden negativ aus, wenn Prüfungen oder Tests einseitig sind. Diese Gefahr ist systemimmanent und hat nur bedingt etwas mit der Reform, aber viel mit dem Handeln der Lehrpersonen zu tun. Im Interesse der Lernenden wird die Gefahr nun erstmals intensiv diskutiert, was hoffentlich zu einer Verbesserung der Prüfungskultur und des Unterrichts führt.

3. *Zur Kommunikation der Ergebnisse* – Wenn Ergebnisse von Tests erfolgreich kommuniziert werden sollen, dann muss dies in der Sprache der Praxis geschehen. Eine ergebnisorientierte Reform wird vor allem dann Erfolg haben, wenn (1) die Bildungsstandards und die Anforderungen gemäss Kompetenzmodell von den betroffenen Lehrpersonen, aber auch von den Eltern und den Schülerinnen und Schülern in Worte gefasst werden können, wenn (2) die Ergebnismeldungen den Lehrpersonen sowie den Lernenden konkrete Anhaltspunkte über die individuelle Förderung im Unterricht geben und wenn (3) die Ergebnisse in den Tests dazu führen, dass die Schülerinnen und Schüler die Möglichkeit erhalten und dazu motiviert werden, das Bildungssystem auf Grund ihrer ausgewiesenen Kompetenzen schneller oder langsamer zu durchlaufen.

Die Einführung von Bildungsstandards und die damit verbundenen Anstrengungen zur Qualitätssicherung und Evaluation gehören zweifelsohne zu den einschneidenden Reformen des Schweizer Bildungswesens, und sie bergen die Gefahr des Scheiterns. Diese Gefahr ist dann gering, wenn der Nutzen für die Lernenden im Zentrum steht.

Literatur

- Amrein, L.A. & Berliner, D.C. (2002). High-Stakes Testing, Uncertainty, and Student Learning. *Education Policy Analysis*, 10(18), 1. Verfügbar unter: <http://epaa.asu.edu/epaa/v10n18/> (14. 3. 2005).
- Baker, E.L. (2004). *Aligning Curriculum, Standards, and Assessments: Fulfilling the Promise of School Reform. CSE Report 645*. Center for the Study of Evaluation. National Center for Research on Evaluation, Standards, and Student Testing. University of California. Verfügbar unter: <http://cresst96.cse.ucla.edu/index4.htm> (14. 3. 2005).
- Cheng, L., Watanabe, Y. & Curtis, A. (2004). *Washback in Language Testing. Research Contexts and Methods*. Mahaw, NJ: Lawrence Erlbaum Associates.
- Cizek, G.J. (2001). *Setting Performance Standards. Concepts, Methods, and Perspectives*. Mahaw, NJ: Lawrence Erlbaum Associates.
- EDK (Schweizerische Konferenz der kantonalen Erziehungsdirektoren) (2004). *HARMOS: Zielsetzungen und Konzeption*. Juni 2004. Bern: EDK.
- Herrmann, U. (2003). Bildungsstandards – Erwartungen und Bedingungen, Grenzen und Chancen. *Zeitschrift für Pädagogik*, (4) 49, 625–639.
- Klieme, E. et al. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Bundesministerium für Bildung und Forum. Verfügbar unter: <http://www.bmbf.de> (14. 3. 2005).
- Maradan, O. & Mangold, M. (2005). *Bildungsstandards in der Schweiz. Das Projekt HarmoS. ph Akzente*, (2), 3. Verfügbar unter: http://www.phzh.ch/content,208,-r1557,_Dz.html (25. 1. 2006).
- Moser, U., Keller, F. & Tresch, S. (2003). *Schullaufbahn und Leistung. Bildungsverlauf und Lernerfolg von Zürcher Schülerinnen und Schülern am Ende der 3. Volksschulklasse*. Bern: h.e.p. Verlag.
- Moser, U., Stamm, M. & Hollenweger, J. (2005). *Für die Schule bereit? Lesen, Wortschatz, Mathematik und soziale Kompetenz beim Schuleintritt*. Aarau: Sauerländer.
- Organisation für wirtschaftliche Zusammenarbeit und Entwicklung [OECD] (2000). *Schülerleistungen im internationalen Vergleich. Ein neues Rahmenkonzept für die Erfassung von Wissen und Fähigkeiten*. Paris: OECD.
- Organisation für wirtschaftliche Zusammenarbeit und Entwicklung [OECD] (2001). *Lernen für das Leben. Erste Ergebnisse von PISA 2000*. Paris: OECD.
- Organisation für wirtschaftliche Zusammenarbeit und Entwicklung [OECD] (2004). *Lernen für die Welt von morgen. Erste Ergebnisse von PISA 2003*. Paris: OECD.
- Rost, J. (2004). *Lehrbuch Testtheorie / Testkonstruktion*. Bern, Göttingen, Toronto, Seattle: Huber.

- Tenorth, H.-E. (2004). Bildungsziele, Bildungsstandards und Kompetenzmodelle. Kritik und Begründungsversuch. In D. Diskowski & E. Hammes-Di Bernardo (Hrsg.), *Lernkulturen und Bildungsstandards. Kindergarten und Schule zwischen Vielfalt und Verbindlichkeit*. Baltmannsweiler: Schneider Verlag Hohengehren.
- Tresch, S. & Moser, U. (2005). *Check 5. Schlussbericht zuhanden des Departements Bildung, Kultur und Sport des Kantons Aargau*. Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich. Verfügbar unter: <http://www.kbl.unizh.ch> (20.01.2006).
- Weinert, F.E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F.E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S.17–31). Weinheim und Basel: Beltz.