

Lerngelegenheiten und Grundbildung: Sekundär-analyse der Schweizer PISA-Daten und Implikationen für die Überprüfung von Bildungsstandards

Urs Moser

Der internationale Schulleistungsvergleich PISA wurde anhand der alltagsbezogenen Grundbildung in Lesen, Mathematik und Naturwissenschaften durchgeführt, die auf den Nutzen im täglichen Leben ausgerichtet ist. Inwieweit diese alltagsbezogene Grundbildung zu den Lerngelegenheiten im Unterricht gehört, wurde nicht berücksichtigt. In einer Sekundäranalyse wurde deshalb überprüft, welche der PISA-Testaufgaben von Expertinnen und Experten sowie von Lehrpersonen als curricular valide bezeichnet wurden und wie sich das Ergebnis der Schweiz verändert, wenn die Parameter ausschliesslich anhand curricular valider Testaufgaben geschätzt werden. Die Ergebnisse sind für das geplante Bildungsmonitoring insofern von Bedeutung, als sie zeigen, dass die Einführung von Bildungsstandards – durch die Bildungsziele in Form von Kompetenzanforderungen konkretisiert werden und dabei festgelegt wird, welche Kompetenzen Kinder oder Jugendliche bis zu einer bestimmten Jahrgangsstufe mindestens erworben haben sollen (EDK, 2004, S.10; Klieme et al., 2003, S. 9) – kaum ohne Anpassung von Lehrplänen und Lehrmitteln erfolgreich sein wird.

Einleitung

Im Vergleich zu früheren internationalen Studien wurde die Frage nach der curricularen Validität der im Rahmen von PISA eingesetzten Tests auf internationaler Ebene kaum thematisiert (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung [OECD], 2001). Während bei der Third International Mathematics and Science Study (TIMSS) auch ein internationaler Vergleich anhand jener Aufgaben durchgeführt wurde, die von den beteiligten Ländern als curricular valide bezeichnet worden waren und die am kleinsten gemeinsamen Nenner der Lehrpläne der beteiligten Länder ausgerichtet waren (Beaton & al., 1996a; Beaton & al., 1996b), verzichtete PISA in Übereinstimmung mit der theoretischen Grundlage des eingesetzten Tests auf eine zusätzliche Auswertung anhand curricular valider Testaufgaben und führte den Ländervergleich einzig anhand der alltagsbezogenen Grundbildung durch (OECD, 1999).

Die alltagsbezogene Grundbildung wird mit Kompetenzen umschrieben, die für die Bewältigung authentischer Anwendungssituationen von besonderer Bedeutung sind (Baumert, Stanat & Demmrich, 2001, S. 19). Das heisst, dass PISA über die an Lehrplänen orientierte Leistungsmessung hinausgeht und aufzeigt, inwieweit die Jugendlichen in Lesen, Mathematik und Naturwissenschaften jene allgemeinen Kenntnisse und Fähigkeiten erworben haben, die sie später als Erwachsene benötigen werden (OECD, 2001, S. 11). Für den internationalen Vergleich der Schulleistungen setzt sich die OECD mit der für PISA definierten Grundbildung somit über die nationalen Curricula hinweg.

Weil mit PISA also nie beabsichtigt wurde, die Curricula der ausgewählten Unterrichtsfächer der teilnehmenden Länder abzubilden, muss davon ausgegangen werden, dass einige der im Test verwendeten Aufgaben hinsichtlich der gestellten Anforderungen beziehungsweise der Art der Aufgabenstellung von dem abweichen, was in den nationalen Lehrplänen explizit thematisiert wird (Artelt, Brunner, Schneider, Prenzel & Neubrand, 2003, S. 78). Dass die nationalen Curricula für die Ergebnisse der Länder in internationalen Schulleistungsvergleichen von Bedeutung sind, wurde mit verschiedenen Sekundäranalysen der TIMSS-Daten nachgewiesen (Moser, Ramseier, Keller & Huber, 1997; Schmidt, McKnight, Cogan, Jakwerth & Houang, 1999). Trotzdem wurde die Frage nach der curricularen Validität der in PISA eingesetzten Tests auch nach der Veröffentlichung der Ergebnisse zumindest in der Schweiz kaum gestellt. Die im Rahmen von PISA erfasste Grundbildung erfuhr – sozusagen als internationaler Standard – eine hohe Akzeptanz, die sich vielleicht mit der ökonomischen Bedeutung der Grundbildung erklären lässt. Lesekompetenzen beispielsweise hängen positiv mit dem Einkommen und negativ mit der Arbeitslosigkeit zusammen: Je besser eine Person bei gleicher Ausbildung lesen kann, desto mehr verdient sie, und je schlechter eine Person bei gleicher Ausbildung lesen kann, desto grösser ist ihre Chance, arbeitslos zu werden (OECD, 2001, S. 22). Ausserdem besteht in der Schweiz auch ein positiver Zusammenhang zwischen Lesekompetenz und Erwerbseinkommen (Notter, Bonerad & Stoll, 1999). Solche Erkenntnisse zeigen, dass die Grundbildung für die soziale Wohlfahrt einer Gesellschaft von hoher Relevanz ist.

Sowohl aus wissenschaftlichem Interesse als auch aus bildungspolitischer Verantwortung heraus muss allerdings bei einer Diskussion der Ergebnisse von PISA berücksichtigt werden, dass beim normativen Vorgehen der OECD unbekannt bleibt, wie gut die PISA-Tests den Curricula der teilnehmenden Länder entsprechen. Für eine zuverlässige Interpretation der PISA-Ergebnisse sind Angaben über die curriculare Validität der Tests deshalb eine notwendige Voraussetzung. Wenn ein Land als Folge mangelnder curriculärer Validität der Tests bei einem internationalen Vergleich schlecht abschneidet, dann lassen sich die Ursachen für die Ergebnisse weder durch Unterrichtsmerkmale noch durch Merkmale des Bildungssystems erklären. Vielmehr muss der Sachverhalt vorerst auf Differenzen zwischen Curriculum und international definierter Testleistung

zurückgeführt werden. Falls beispielsweise der PISA-Lesetest in keiner Weise den Intentionen und Inhalten der Lehrpläne entspricht, führen Zusammenhänge zwischen Merkmalen des Bildungssystems oder der Unterrichtsqualität und den Ergebnissen im internationalen Vergleich kaum zu den erhofften zuverlässigen Erklärungen. Rückschlüsse auf die Qualität von Bildungssystem und Unterricht sind vor allem dann sinnvoll, wenn die Qualität am intendierten Output gemessen wird. Ansonsten können die nachgewiesenen Zusammenhänge in der Tat auch durch nationale curriculare Besonderheiten erklärt werden. Eine Verbesserung der Ergebnisse im internationalen Vergleich würde dann auch durch Anpassungen der Lehrpläne und Lehrmittel erreicht.

Erkenntnisse über den Zusammenhang zwischen Lerngelegenheiten und Leistungen, wie sie mit Tests erfasst werden, sind für die Evaluation des Bildungssystems und für ein gesamtschweizerisches Bildungsmonitoring auf der Grundlage von nationalen Bildungsstandards insofern von Bedeutung, als im Rahmen der Harmonisierung der obligatorischen Schule (HarmoS) die Lehrpläne nicht durch Bildungsstandards ersetzt werden sollen (Schweizerische Konferenz der kantonalen Erziehungsdirektoren [EDK], 2004). Bildungsziele der Lehrpläne dienen als Orientierung beziehungsweise zur Legitimation von Bildungsstandards: «Ohne Bezug auf allgemeine Bildungsziele wären Kompetenzanforderungen reine Willkür oder blosser Expertenmeinung. Erst die Orientierung an diesen Zielen legitimiert die Bestimmung von erwünschten Niveaustufen und die daraus resultierenden Testverfahren. Insbesondere sollten die Standards von einem Verständnis des Bildungsauftrags der jeweiligen Fächer ausgehen, das expliziert werden muss» (Klieme & al., 2003, S. 23). Während die Lehrpläne eine Grundlage für die Bildungsstandards sind, werden Unterricht und Evaluation im Nachhinein an den Bildungsstandards ausgerichtet. Bildungsstandards sind für den Unterricht grundlegend, weil sie Kompetenzanforderungen beziehungsweise Ergebnisse von Lernprozessen konkretisieren, die von den Schülerinnen und Schülern erreicht werden müssen. Damit empirisch überprüft werden kann, wie gut die Standards im Unterricht erreicht wurden, müssen auch die eingesetzten Aufgabenstellungen und Testverfahren an den Bildungsstandards ausgerichtet sein. Vorausgesetzt, es wird nach der expliziten Beschreibung des Verhältnisses von Bildungsstandards und Lehrplan gehandelt, kann davon ausgegangen werden, dass die Lehrpläne ihre Bedeutung als strukturierendes Element von Unterricht nicht verlieren werden (Klieme & al., 2003, S. 93) und auf eine ausschliesslich output-orientierte Steuerung des Bildungssystems verzichtet wird. Dass sich die postulierte Beziehung zwischen Bildungsstandards und Lehrplänen auf Grund der kantonal unterschiedlichen Lehrpläne in der Praxis nicht so einfach umsetzen lässt, zeigen die Ergebnisse einer Sekundäranalyse der Daten der Erhebung PISA 2000, mit der die Beziehungen zwischen den kantonalen Curricula und dem PISA-Test sowie zwischen Curricula und Leistungen analysiert wurden (Moser & Berweger, 2003).

Methode

Erfassung der curricularen Validität der PISA-Testaufgaben

Um abschätzen zu können, wie gut der PISA-Test mit den Inhalten und Anforderungen in Schweizer Schulen der Sekundarstufe I übereinstimmt, wurde eine Befragung von Lehrplanexpertinnen und -experten sowie von Lehrpersonen zur Erfassung der curricularen Validität der PISA-Testaufgaben durchgeführt. Mit dem Begriff «curricular» wird darauf hingewiesen, dass die Validität der PISA-Tests nicht einfach an den fachlichen Zielen und Inhalten überprüft werden kann, wie sie in Dokumenten des Bildungssystems festgehalten sind. Während Lehrpläne vorwiegend Lerninhalte zur Erreichung bestimmter Lernziele enthalten, wird mit dem Begriff Curriculum ein umfassendes Konzept zur effektiven Planung und Durchführung des Unterrichts bezeichnet, das Aussagen über Lernziele, Lerninhalte, Lernmethoden, Medien und Verfahren zur Lernzielkontrolle enthält (Schröder, 1992, S. 50f.). Weil Schweizer Lehrpläne gemäss dieser Definition kaum als Curricula bezeichnet werden können, lässt sich die curriculare Validität der PISA-Tests ohne Berücksichtigung der Umsetzung von Zielen und Inhalten im Unterricht nur beschränkt bestimmen. Schweizer Lehrpläne enthalten zwar verbindliche Vorgaben, die allerdings meist sehr offen gestaltet sind. Der Lehrplan des Kantons Zürich enthält beispielsweise zum Lesen auf der Sekundarstufe I verschiedene Angaben über den Umgang mit Texten wie «längere Texte selbständig still lesen» oder «schriftliche Anleitungen, Anweisungen und Aufträge verstehen und ausführen» (Erziehungsdirektion des Kantons Zürich, 1993, S. 148ff.). Welche Anforderungen an die Schülerinnen und Schüler zu stellen sind, ist jedoch nur sehr vage formuliert. Weil Angaben darüber, ob oder wann die Ziele erreicht sind, gänzlich fehlen, liegen diese Entscheidungen in der Hand der beurteilenden Lehrpersonen. Die Erfassung der Leistungserwartungen von Lehrpersonen liefert deshalb eine Information, die die Unterrichtspraxis wesentlich zuverlässiger repräsentiert als Umschreibungen in Lehrplänen. Zum Curriculum wurden für die vorliegende Untersuchung deshalb auch die im Unterricht behandelten Inhalte und die gestellten Anforderungen gezählt.

Unter Einbezug der Leistungen der Schülerinnen und Schüler konnten somit folgende drei Aspekte des Curriculums berücksichtigt werden:

- (1) die in Lehrplänen und Lehrmitteln schriftlich festgehaltenen Inhalte und Anforderungen (angestrebtes Curriculum)
- (2) die von den Lehrpersonen im Unterricht behandelten Inhalte und die gestellten Anforderungen (umgesetztes Curriculum)
- (3) die von den Schülerinnen und Schülern beherrschten Inhalte und Anforderungen (erreichtes Curriculum)

Zur Beurteilung der curricularen Validität wurden den Lehrplanexpertinnen und -experten sowie den Lehrpersonen sämtliche Testeinheiten mit den zugehö-

rigen Aufgaben vorgelegt. In Anlehnung an eine ähnliche Untersuchung in Deutschland (Artelt & al., 2003) mussten die Testaufgaben nach verschiedenen Kriterien auf ihre curriculare Validität hin beurteilt werden. Gegenstand der Befragung waren die Aufgaben der drei Tests zu den Lesekompetenzen, zur mathematischen und zur naturwissenschaftlichen Grundbildung. Pro Testaufgabe wurde erfasst, (1) in welcher Klassenstufe die Art des Textes beziehungsweise der Inhalt der Aufgabe in der Regel behandelt wird, (2) in welcher Klassenstufe Texte beziehungsweise Inhalte dieses Schwierigkeitsgrades behandelt werden, (3) ab welcher Klassenstufe erwartet wird, dass die Anforderung der Testaufgabe bewältigt werden und (4) wie vertraut die Schülerinnen und Schülern mit der Art und Weise der Aufgabenstellung sind. Die Expertinnen und Experten beantworteten die Fragen auf Grund der in Lehrplänen und Lehrmitteln festgehaltenen Inhalte und Anforderungen. Die Lehrpersonen beantworteten die Fragen auf Grund der im Unterricht behandelten Inhalte und gestellten Anforderungen. Weil die Urteile zum Teil uneinheitlich ausfielen, wurde eine Aufgabe dann als curricular valide gewertet, wenn 80 Prozent der Befragten zur gleichen Einschätzung kamen.

Die Befragung war national ausgerichtet und bezog sich auf die Schülerinnen und Schüler der 9. Klassen aus Schulen mit Grundansprüchen (Realschulen) und aus Schulen mit erweiterten Ansprüchen (Sekundarschulen). An der Befragung beteiligten sich 56 Lehrplanexpertinnen und Lehrplanexperten sowie 109 Lehrpersonen der Kantone Bern, St. Gallen, Tessin und Zürich sowie aus den Kantonen der Zentralschweiz und der französischen Schweiz. Weil die Ergebnisse eindeutig darauf hinweisen, dass zwischen den Kantonen innerhalb einer Sprachregion relativ grosse Übereinstimmung in der Einschätzung der curricularen Validität der PISA-Testaufgaben vorliegt, die Unterschiede zwischen den Sprachregionen jedoch sehr gross sind, wurden die Ergebnisse nach Sprachregionen getrennt analysiert. Zusätzlich zur Differenzierung nach Sprachregion wurden die Ergebnisse auch nach Schultypen beziehungsweise nach Schulen mit erweiterten Ansprüchen und Schulen mit Grundansprüchen getrennt dargestellt.

Skalierung der PISA-Daten

Zur Beantwortung der Frage, wie das Ergebnis der Schweiz ausgefallen wäre, wenn in PISA nur Testaufgaben eingesetzt worden wären, die den Schülerinnen und Schülern auf Grund ihrer Textform und ihrer Anforderungen vertraut sind, wurden die Ergebnisse in PISA anhand einer curricular validen Testversion berechnet. Die Rohdaten der Schülerinnen und Schüler wurden entsprechend dem internationalen Verfahren (Adams, 2002) nach dem Rasch-Modell, einem Skalierungsverfahren auf der Grundlage der probabilistischen Testtheorie, skaliert (Rost, 2004). Die Leistungen der Schülerinnen und Schüler der 9. Klasse wurden in einem zweidimensionalen Modell geschätzt. Die PISA-Testaufgaben wurden aufgeteilt in solche, die auf Grund der Curricula für die Schülerinnen und Schüler der 9. Klasse entweder lösbar sind (erste Dimension des Tests), und sol-

che, die nicht lösbar sind (zweite Dimension des Tests). Eine der grundlegenden Annahmen des Rasch-Modells wird mit dem Begriff der spezifischen Objektivität umschrieben (Rost, 2004, S. 38). Die spezifische Objektivität bezeichnet die Unabhängigkeit der Testergebnisse von der Auswahl der Testaufgaben. Dahinter steht die Überlegung, dass jeder Test nur eine begrenzte Anzahl von Aufgaben umfassen kann, das Testergebnis aber nicht nur etwas über die Fähigkeit zur Beantwortung der begrenzten Auswahl von Aufgaben aussagen soll, sondern etwas über die Fähigkeit zur Beantwortung von Aufgaben des gleichen Typs insgesamt (ebd. S. 38). Damit die Ergebnisse anhand der curricular validen Testversion mit den Ergebnissen im originalen PISA-Test verglichen werden konnten, wurden die internationalen Schwierigkeitsparameter als Anker benutzt. Durch diese Vorgehensweise können die Ergebnisse auf der internationalen Skala, die auf einen Mittelwert von 500 Punkten mit einer Standardabweichung von 100 Punkten normiert ist, abgebildet werden (Adams, 2002).

Tabelle 1: Stichprobe zur Überprüfung der Bedeutung der curricularen Validität des PISA-Tests

Region	Schülerinnen und Schüler der 9. Klasse					
	erweiterte Ansprüche			Grundansprüche		
	L	M	N	L	M	N
Deutschschweiz Bern, St. Gallen, Zürich	1509	834	836	1125	626	633
Französische Schweiz Freiburg, Genf, Neuenburg	901	494	498	519	297	294

Anmerkung: L = Lesen, M = Mathematik, N = Naturwissenschaften

Zur neuen Berechnung der Ergebnisse der Schweiz auf Grund curricular valider PISA-Tests wurden die Daten von Schülerinnen und Schülern der 9. Klassen einbezogen, sofern sie einem der Schultypen mit Grundansprüchen (beispielsweise Realschule) oder erweiterten Ansprüchen (beispielsweise Sekundarschule) zugeordnet werden konnten und sofern die Kantone sich mit einer kantonalen Stichprobe an PISA 2000 beteiligt hatten. Schülerinnen und Schüler des Langzeitgymnasiums wurden ausgeschlossen. Dies führte zu einer Teilstichprobe der Schülerinnen und Schüler der Deutschschweiz (vgl. Tabelle 1). Für die Überprüfung der Bedeutung der curricularen Validität des PISA-Tests wurden Testversionen des originalen PISA-Tests nach Region, Schultyp und Fachbereich zusammengestellt.

Ergebnisse

Curriculare Validität der PISA-Testaufgaben

Die in PISA eingesetzten Tests stimmen mit den Zielvorgaben der Schweizer Lehrpläne und deren Umsetzung im Unterricht je nach Region und Schultyp unterschiedlich gut überein. Die curriculare Validität der PISA-Tests ist in der Deutschschweiz am höchsten, in der französischen Schweiz am geringsten. Für die Schülerinnen und Schüler aus Schulen mit Grundansprüchen bezeichnen die Lehrplanexpertinnen und Lehrplanexperten und die Lehrpersonen weit weniger Aufgaben als lösbar als für die Schülerinnen und Schüler aus Schulen mit erweiterten Ansprüchen.

Im Lesen waren die Schülerinnen und Schüler vorwiegend bei Aufgaben, die sich auf nicht kontinuierliche Texte bezogen, ungenügend vorbereitet. Als unlösbar wurden vor allem Aufgaben zu Tabellen, Karten, Graphen oder Formularen eingeschätzt. Hauptsächlich in der französischen Schweiz werden nicht kontinuierliche Texte im Sprachunterricht kaum eingesetzt. Leseaufgaben zu kontinuierlichen Texten, beispielsweise zu Erzählungen oder Darlegungen, wurden für die Schülerinnen und Schüler dann als nicht lösbar eingeschätzt, wenn sie sich auf Texte von mehreren Seiten bezogen. Vor allem die Schülerinnen und Schüler aus Schulen mit Grundansprüchen waren auf solche Aufgaben schlecht vorbereitet. Aufgaben, bei denen es komplexe und lange Texte kritisch zu bewerten gilt und die verlangen, dass Hypothesen über den Text aufgestellt werden, sind für diese Schülerinnen und Schüler auf Grund des Curriculums nicht lösbar. Zudem wird von ihnen im Unterricht kaum verlangt, dass sie über Texte reflektieren müssen.

In der Mathematik wurden nahezu sämtliche Aufgaben zu den Fachbereichen Geometrie und Messen als lösbar eingeschätzt. Sowohl mit dem Inhalt als auch mit der Art und Weise der Aufgabenstellung sind die Schülerinnen und Schüler vertraut. Aufgaben, die als nicht lösbar eingeschätzt wurden, gehören zu den Fachbereichen Algebra, Statistik und Funktionen. Die meisten Aufgaben verlangen, dass die Schülerinnen und Schüler mathematische Zusammenhänge herstellen oder in alltäglichen Situationen enthaltene Mathematik erkennen und umsetzen. Keine Probleme bieten Wissens- und Routineaufgaben.

Als am geringsten wurde die curriculare Validität des naturwissenschaftlichen Tests beurteilt. Vor allem in der französischen Schweiz wurden auf Grund der Curricula weniger als die Hälfte der Aufgaben für die Schülerinnen und Schüler als lösbar eingeschätzt. Dabei liessen sich bei jenen Aufgaben, die auf Grund von Lehrplänen und deren Umsetzung im Unterricht als nicht lösbar eingestuft wurden, keine speziellen Muster erkennen.

Ergebnisse auf der Basis von curricular validen PISA-Tests

Wie wäre nun das Ergebnis der Schweiz ausgefallen, wenn in PISA 2000 nur Leseaufgaben eingesetzt worden wären, die auf Grund ihres Inhaltes und ihrer An-

forderungen den Schülerinnen und Schülern vertraut sind? Hatte die – auf Grund von Lehrplänen und Lehrmittel – zum Teil schlechte Vorbereitung der Schweizer Schülerinnen und Schüler auf die PISA-Tests für die Ergebnisse im internationalen Vergleich eine Bedeutung? Der Vergleich der Ergebnisse in PISA 2000 mit den Ergebnissen anhand curricular valider PISA-Tests zeigt, dass die Schweiz mit auf die Curricula ausgerichteten Tests zum Teil wesentlich besser abschneidet. Tabelle 2 enthält die Ergebnisse des Vergleichs der Leistungen im Lesen, in der Mathematik und in den Naturwissenschaften: einmal berechnet auf Grund des PISA-Test, einmal berechnet auf Grund der curricular validen Versionen des PISA-Tests. Die Prüfung auf signifikante Unterschiede erfolgte über die Vertrauensintervalle. Dafür wurden die Grenzen ermittelt, innerhalb welcher sich ein wahrer Mittelwert mit hoher Wahrscheinlichkeit befindet (Konfidenzintervall). Die Konfidenzintervalle um den Mittelwert wurden bei einem Signifikanzniveau von $\alpha = 0.05$ durch Addition und Subtraktion von 1.96-mal den Stichprobenfehler des Mittelwerts berechnet. Die Ergebnisse in den beiden Testversionen unterscheiden sich dann signifikant, wenn sich die Konfidenzintervalle überschneiden.

Tabelle 2. Vergleich der Ergebnisse zwischen dem PISA-Test und curricular validen Versionen des PISA-Tests

Lesen	PISA-Test 2000			curricular valider PISA-Test		
	M	SE	SD	M	SD	SE
<i>Erweiterte Ansprüche</i>						
Deutschs Schweiz	519.4	3.2	67.6	525.8	3.1	78.0
Französische Schweiz	502.6	3.5	57.4	507.1	4.1	71.4
<i>Grundansprüche</i>						
Deutschs Schweiz	416.6	3.3	76.8	423.3	3.6	83.6
Französische Schweiz	420.3	6.6	66.7	426.0	7.0	77.9
<i>Mathematik</i>						
	M	SE	SD	M	SE	SD
<i>Erweiterte Ansprüche</i>						
Deutschs Schweiz	552.9	3.2	69.3	563.0	3.0	73.4
Französische Schweiz	536.7	2.7	59.6	547.1	3.3	67.9
<i>Grundansprüche</i>						
Deutschs Schweiz	457.5	3.4	75.9	487.7	3.1	72.1
Französische Schweiz	472.4	5.8	65.1	499.4	4.9	69.4
<i>Naturwissenschaften</i>						
	M	SE	SD	M	SE	SD
<i>Erweiterte Ansprüche</i>						
Deutschs Schweiz	509.8	3.5	71.5	511.1	3.8	84.9
Französische Schweiz	499.1	3.7	69.9	499.3	4.1	82.7
<i>Grundansprüche</i>						
Deutschs Schweiz	421.8	3.3	71.8	440.2	3.4	78.5
Französische Schweiz	416.4	7.3	72.3	440.5	6.4	76.5

Anmerkung: M = Mittelwert, SE = Stichprobenfehler, SD = Standardabweichung

Im Lesen weichen die Ergebnisse im curricular validen PISA-Test weder statistisch signifikant, noch besonders stark vom ursprünglichen Ergebnis im PISA-Test ab. Die Differenzen betragen zwischen rund 5 und 7 Punkten. Die grösste Bedeutung hat die Ausrichtung der PISA-Tests an einer alltagsbezogenen Grundbildung für die Ergebnisse in der Mathematik. Die Schülerinnen und Schüler in Schulformen mit Grundansprüchen beider Regionen erreichen in der curricular validen Version des PISA-Tests statistisch signifikant bessere Ergebnisse. Die Differenzen betragen in der französischen Schweiz rund 30 Punkte. Der gleiche Befund zeigt sich bei den Naturwissenschaften. Während es für die Ergebnisse der Jugendlichen in Schulformen mit erweiterten Ansprüchen kaum eine Rolle spielt, welche Testversion eingesetzt wird, erreichen die Jugendlichen im curricular validen PISA-Test statistisch signifikant bessere Ergebnisse.

Diskussion

Die vorliegenden Ergebnisse sind für das Ziel der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren, die Lerninhalte zu harmonisieren und die Qualität des Bildungssystems durch Standards zu sichern, in verschiedener Hinsicht relevant.

Wenn die Grundbildung der Schülerinnen und Schüler im Sinne der Definition von PISA ein wünschenswertes Ziel darstellt, dann sind sowohl im Sprachunterricht als auch in der Mathematik und in den Naturwissenschaften curriculare Anpassungen notwendig. Diese Anpassungen sind insbesondere für den Unterricht in den Schulen mit Grundansprüchen notwendig. Ein grosser Teil der Aufgaben war für diese Schülerinnen und Schüler deshalb nicht lösbar, weil entweder der Inhalt der Aufgaben im Unterricht nicht durchgenommen worden war oder weil die gestellten Anforderungen zu hoch waren. Der nachgewiesene Zusammenhang zwischen Lehrplan und Lehrmitteln einerseits und schulischen Leistungen andererseits ist ein Beleg dafür, dass die Anpassung der Curricula beziehungsweise die Einführung von Bildungsstandards Wirkung zeigen kann. Transparenz über die in den Schulen vermittelten und angestrebten Kompetenzen beziehungsweise Anforderungen, wie sie durch die geplanten Bildungsstandards geschaffen wird, kann nach den vorliegenden Ergebnissen zu einer Verbesserung der Bildungsqualität führen. Es bestehen kaum Zweifel, dass die heutigen Lehrpläne der Schweiz, die inhaltlich innerhalb der Sprachregionen bereist in hohem Masse harmonisiert, aber meist auch wenig konkret sind, eine Ergänzung in Form von Bildungsstandards benötigen, die Klarheit über die Anforderungen im Unterricht und in Prüfungen schaffen.

Die Ergänzung der Lehrpläne durch Standards ist vor allem dann wirksam, wenn sich die Standards und die damit verbundenen Kompetenzbeschreibungen in Leistungstests, aber auch im Unterricht wieder erkennen lassen (Klieme & al., 2003, S. 23, S. 50). Dazu fehlen den Lehrpersonen zurzeit noch die nötigen Ins-

trumente, die für diagnostische Tätigkeiten oder die Leistungsbeurteilung im Unterricht eingesetzt werden können. Die Entwicklung von Bildungsstandards sollte in Übereinstimmung mit der Unterrichtsentwicklung und der Prüfungskultur im Unterricht erfolgen. Dies ist deshalb wichtig, weil es den Lehrpersonen nicht leicht fällt, anhand von konkreten Aufgabenbeispielen die Leistungsfähigkeit ihrer Schülerinnen und Schüler zuverlässig einzuschätzen (Klieme & al., 2003, S. 30; Moser & Berweger, 2003). Erwartungen der Lehrpersonen, was Schülerinnen und Schüler zu leisten imstande sind oder wie viele Lernangebote ihnen in einer bestimmten Zeitspanne zuzumuten sind, tragen dazu bei, wie Standards und Ziele im Unterricht verfolgt werden (Schrader & Helmke, 2001, S. 47). Sind die Erwartungen zu hoch oder zu niedrig, wirkt sich das auf die Leistungsfähigkeit ungünstig aus. Individualisieren und Differenzieren im Unterricht verlangt, dass Erkenntnisse über den Leistungsstand der Schülerinnen und Schüler vorhanden sind. Wie genau und realistisch Erwartungen sind, hängt von den Diagnosekompetenzen der Lehrpersonen ab. Zur diagnostischen Kompetenz gehört, dass die Fähigkeiten der Schülerinnen und Schüler sowie die Schwierigkeit von Aufgaben zutreffend eingeschätzt werden können (Schrader & Helmke, 2001, S. 48f.).

Die Absicht, dass die Entwicklung von Lehrplänen und Lehrmitteln Aufgabe der Kantone bleiben soll (EDK, 2004, S. 4) ist auf Grund der vorliegenden Ergebnisse nur beschränkt zu begrüssen. Obwohl PISA einen internationalen Bildungsstandard auf der Grundlage eines Kompetenzmodells setzt und deshalb in der Diskussion zu den Bildungsstandards auch als exemplarisches Beispiel zitiert wird (Klieme & al., 2003, S. 76), haben Lehrpläne und Lehrmittel eine Bedeutung für das, was die Schülerinnen und Schüler in der Schule lernen und leisten. Um das Ziel der Steuerung des Bildungssystems über den Output erreichen zu können, sollten die Leitfunktion nationaler Bildungsstandards und die Orientierungsfunktion von Lehrplänen systematisch gekoppelt werden. Weil in der Schweiz bereits eine grosse Anzahl kantonaler Lehrpläne vorliegt und Bildungsstandards die Bildungsziele in Form von Kompetenzmodellen konkretisieren, wird sich der Prozess der Harmonisierung auf die Lehrpläne ausweiten müssen. Gleich wie die Trennung zwischen der Entwicklung von Standards und der Entwicklung von Leistungstests die Gefahr birgt, dass die im Unterricht verfolgten Standards nicht mit den durch Leistungstests überprüften Standards übereinstimmen, birgt eine mangelnde Koordination zwischen Lehrplänen und Standards zur Gefahr, dass sich Lehrpersonen nur an Standards oder den damit verbundenen Testaufgaben oder nur an Lehrplänen orientieren. Um dies zu verhindern, sollte eine möglichst grosse Kongruenz zwischen Bildungszielen, Bildungsstandards, Unterrichtsentwicklung und Leistungstests herrschen.

Ein grosser Teil der Kompetenzen kann vermutlich durchaus mit unterschiedlichen Lehrplaninhalten vermittelt und erreicht werden. Vor allem Lesekompetenzen können durch unterschiedlichste Inhalte und Texte gefördert werden. Bei der Mathematik und den Naturwissenschaften ist die Situation aller-

dings anders zu beurteilen. Die Analysen haben gezeigt, dass Lücken in Lehrplänen und Lehrmitteln insbesondere in diesen Fächern eine Folge haben. Wenn im Mathematikunterricht die Themen «Statistik» oder «Funktionen» nicht behandelt werden, dann werden die Schülerinnen und Schüler Testaufgaben zu diesen Themen auch kaum lösen können. Dies erklärt zumindest zu einem Teil die PISA-Ergebnisse der Schweizer Schülerinnen und Schüler in Schulen mit Grundansprüchen (BFS & EDK, 2002). Neben der Beschreibung von Kompetenzanforderung ist eine Harmonisierung der Lehrpläne zumindest prüfenswert. Ein auf Bildungsstandards ausgerichtetes Qualitätsmanagement benötigt nicht nur die Kongruenz von Standards und Evaluationsinstrumenten, sondern auch die Übereinstimmung von Lehrplänen und Bildungsstandards. Die Harmonisierung durch nationale Bildungsstandards kann deshalb kaum ohne Anpassung beziehungsweise Harmonisierung der Lehrpläne erfolgreich sein, was in Anbetracht der Beschränkung der Standards auf Mindestkompetenzen einfach zu verwirklichen sein sollte. Die Kongruenz wird zudem am einfachsten dann erreicht, wenn die Entwicklung von Standards mit der Harmonisierung der Lehrpläne und der Entwicklung von Testaufgaben *gleichzeitig* erfolgt. Wie PISA nämlich gezeigt hat, führt die empirische Überprüfung von Kompetenzmodellen anhand von Leistungstests immer auch zu Anpassungen der theoretischen Modelle (Adams, 2002; OECD, 1999, 2001). Die Verbindung der Entwicklung von Kompetenzmodellen mit der Entwicklung von Leistungstests – was in der Schweiz vorgesehen ist (EDK, 2004) – ist eine Voraussetzung dafür, dass die an Bildungsstandards gestellten Kriterien der Fachlichkeit, Fokussierung, Kumulativität, Verbindlichkeit, Differenzierung Verständlichkeit und der Realisierbarkeit (Klieme & al., 2003, S. 25) erfüllt werden. Erst die empirische Überprüfung von Kompetenzmodellen zeigt, was konkret unter den Anforderungen verstanden wird, ob die Kompetenzen aufeinander aufbauen und ob das Ziel der Kongruenz von Bildungszielen, Bildungsstandards und Leistungstests erreicht wurde.

Literatur

- Adams, R. (2002). Scaling PISA Cognitive Data. In R. Adams & M. Wu (Eds.), *PISA 2000 Technical Report* (pp. 99-108). Paris: Organisation for Economic Co-operation and Development.
- Artelt, C., Brunner, M., Schneider, W., Prenzel, M. & Neubrand, M. (2003). Literacy oder Lehrplanvalidität? – Ländervergleiche auf der Basis lehrplanoptimierter PISA-Tests. In J. Baumert, C. Artelt, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, K.-J. Tillmann & M. Weiss (Hrsg.), *PISA 2000 – Ein differenzierter Blick auf die Länder der Bundesrepublik Deutschland* (S. 77-108). Opladen: Leske und Budrich.
- Baumert, J., Stanat, P. & Demmrich, A. (2001). PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In Deutsches PISA-Konsortium (Hrsg.), *PISA 2000 – Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 15-68). Opladen: Leske und Budrich.

- Beaton, A.E., Martin, M.O., Mullis, I.V.S., Gonzales, E.J., Kelly, D.L. & Smith, T.A. (1996a). *Science Achievement in the Middle School Years: IEA's third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College.
- Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzales, E.J., Kelly, D.L. & Smith, T.A. (1996b). *Mathematics Achievement in the Middle School Years: IEA's third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Boston College.
- BFS & EDK (2002). *Für das Leben gerüstet? Die Grundkompetenzen der Jugendlichen – Nationaler Bericht der Erhebung PISA 2000*. Neuchâtel: Bundesamt für Statistik.
- EDK (2004). *HARMOS: Zielsetzungen und Konzeption* [Juni 2004]. Bern: EDK.
- Erziehungsdirektion des Kantons Zürich (1993). *Lehrplan für die Volksschule des Kantons Zürich*. Zürich: Lehrmittelverlag.
- Klieme, E. et al. (2003). *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.
- Moser, U. & Berweger, S. (2003). *Lehrplan und Leistung. Thematischer Bericht der Erhebung PISA 2000*. Neuchâtel: Bundesamt für Statistik.
- Moser, U., Ramseier, E., Keller, C. & Huber, M. (1997). *Schule auf dem Prüfstand. Eine Evaluation der Sekundarstufe I auf der Grundlage der «Third International Mathematics and Science Study»*. Chur/Zürich: Rüegger.
- Notter, Ph., Bonerad, E.M. & Stoll, F. (1999). *Lesen – eine Selbstverständlichkeit? Schweizer Bericht zum «International Adult Literacy Survey»*. Chur/Zürich: Rüegger.
- OECD (Organisation for Economic Co-Operation and Development) (1999). *Measuring Student Knowledge and Skills: A New Framework for Assessment*. Paris: Organisation for Economic Co-operation and Development.
- OECD (Organisation for Economic Co-Operation and Development) (2001). *Lernen für das Leben. Erste Ergebnisse von PISA 2000*. Paris: Organisation for Economic Co-operation and Development.
- Rost, J. (2004). *Lehrbuch Testtheorie Testkonstruktion*. Bern: Huber.
- Schmidt, W.H., McKnight, C.C., Cogan, L.S., Jakwerth, P.M. & Houang, R.T. (1999). *Facing the Consequences. Using TIMSS for a Closer Look at U.S. Mathematics and Science Education*. Dordrecht: Kluwer Academic Publishers.
- Schrader, F.-W. & Helmke, A. (2001). Alltägliche Leistungsbeurteilung durch Lehrer. In F.E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 45-58). Weinheim und Basel: Beltz Verlag.
- Schröder, H. (1992). *Grundwortschatz Erziehungswissenschaft*. Ein Wörterbuch der Fachbegriffe von Abbilddidaktik bis Zielorientierung. München: Ehrenwirth.

Schlagnote: Bildungsstandards, Bildungsmonitoring, Bildungsevaluation, Leistungsmessung, PISA, Lehrplan, Curriculare Validität

Opportunités d'apprendre et formation de base: Analyse secondaire des données PISA pour la Suisse et implications pour la révision de standards éducatifs

Résumé

L'enquête PISA a évalué la lecture, les mathématiques et les sciences, dans la perspective d'une formation de base, nécessaire à la vie quotidienne et adaptée aux besoins de cette vie quotidienne. La question de savoir si cette éducation relève

de l'enseignement scolaire n'a pas été posée. Une analyse secondaire a donc demandé à des experts et à des enseignants quelles épreuves PISA pouvaient être considérées comme relevant des curricula scolaires. On a examiné comment variaient les résultats, lorsque les paramètres étaient exclusivement évalués en fonction des épreuves pertinentes aux curricula. Les résultats obtenus jusqu'à présent sont significatifs en termes de monitoring de la formation. Ils démontrent que l'introduction de standards de formation (par lesquels on établit concrètement des objectifs éducatifs, sous la forme de compétences attendues des élèves d'un degré d'enseignement donné) ne saurait réussir sans adaptation des plans d'étude et des moyens pédagogiques.

Mots clés: Standards de formation, Monitoring de la formation, Evaluation de l'enseignement, Mesure des performances, PISA, Plans d'étude, Validité des curricula

Opportunità di apprendimento e formazione di base: analisi secondaria dei dati svizzeri di PISA e implicazioni per la verifica degli standard formativi

Riassunto

Lo studio comparativo internazionale PISA ha preso in esame la formazione di base nei campi della lettura, della matematica e delle scienze naturali con particolare riferimento alla sua funzionalità nelle situazioni di vita quotidiana. Non si è però considerato in che misura questi riferimenti alla vita quotidiana coincidano con le opportunità di apprendimento effettive. Ecco perché in un'analisi secondaria si è voluto verificare quali tra i compiti compresi nel test PISA vengano considerati validi sul piano curricolare da esperti ed insegnanti e come si modifichino i risultati se si prendono in considerazione solo i compiti ritenuti esplicitamente validi. Questi risultati sono importanti ai fini di un monitoring della formazione in quanto mostrano come l'introduzione di standard formativi – intesi come delle concretizzazioni delle mete formative sotto forma di competenze minime richieste ai giovani a differenti livelli d'età (CDPE, 2004, p. 10; Klieme 2003, p. 9) – non potrà avere successo senza un adattamento dei programmi e dei sussidi didattici.

Parole chiave: standard formativi, monitoring della formazione, valutazione della formazione, valutazione, PISA, piano di formazione, validità curricolare

Learning situations and general education: Secondary analyses of the Swiss PISA data. Implications for the evaluation of educational standards

Summary

The international comparisons of students' school achievement within the PISA study have been carried out on subjects which are useful in everyday life – reading, math and science. Whether or not this life-related general education is part of learning situations within tuition was not part of the analyses. Therefore a secondary analysis was carried out in which teachers and other experts made an appraisal on the curricular validity of PISA test items, and these same experts made an estimation on how different the Swiss results would be if the tests would be of higher curricular validity. The results are meaningful with regards to the planned education monitoring insofar that they show that the implementation of educational standards are unlikely to be successful, unless curricula and study materials will be adjusted.

Key words: educational standards, monitoring, evaluation, testing, curriculum, curricular validity, PISA