

Editorial: Bildungsstandards, Systemevaluation und Qualitätsentwicklung

Michel Nicolet und Max Mangold

Die Einführung von Standards als Instrument zur besseren Steuerung von Bildungssystemen und zur Steigerung der Unterrichtsqualität im Rahmen nationaler Bildungsreformen hat inzwischen und wohl nicht zuletzt auf dem Hintergrund der grossen internationalen Leistungstests wie PISA an bildungspolitischer Relevanz gewonnen. Diese «empirische Wende der Bildungspolitik» (Lange, 1999) wurde zum Anlass genommen, in verschiedenen Ländern für das jeweilige Bildungssystem Standards zu entwickeln. Mittlerweile liegt ebenfalls eine ganze Anzahl von Forschungsergebnissen aus Kanada, Schweden, Finnland und weiteren Ländern vor, die bereits einschlägige Erfahrungen mit Bildungsstandards gemacht haben (Fitzner, 2004).

Die Mittel zur Steuerung und Qualitätssicherung zeigen im internationalen Vergleich eine eindeutige Richtung: Zielsteuerung, Entwicklung von Standards, interne und externe Evaluation sowie die «Umstellung der Ausbildungskultur am Ergebnis» (Oelkers, 2005). Galten die bisherigen schulischen Standards als eher unverbindlich und wenig präzise und standen unter dem Verdacht sich jeder wirklichen Kontrolle zu entziehen, verbindet sich mit dem «modernen» Begriff Standard die Vorstellung einer präzisen Erfassung schulischer Leistung über die Messung des Outcomes. Die bislang traditionelle Input-Orientierung soll zunehmend auf das Prinzip der Wirkungsorientierung verlagert werden. Die meisten Konzepte gehen diesbezüglich von Leistungs- bzw. Performancestandards aus, also von klar definierten Erwartungen im Hinblick auf schulische Kompetenzen (Ravitch, 1995). Der Begriff Standard steht also nicht für Leistungen, sondern für einen Massstab, an dem die Leistungen zu messen sind. Besteht in dieser Hinsicht eine mehr oder weniger grosse konzeptionelle Übereinstimmung, ist dies für die Art oder das Zielniveau der Leistungsstandards nicht mehr der Fall.

Während etwa in Deutschland Leistungsstandards als Regelstandards festgelegt werden, die das durchschnittliche Erwartungsniveau an alle Schülerinnen und Schüler vorgeben, sollen im schweizerischen Kontext sogenannte Mindeststandards entwickelt werden. Diese im Rahmen des EDK-Projekts HarmoS

(Harmonisierung der obligatorischen Schule – Schweizerische Konferenz der kantonalen Erziehungsdirektoren) zu leistende Arbeit sieht vor, für die Bereiche Erstsprache, Fremdsprachen, Mathematik und Naturwissenschaften der 2., 6. und 9. Jahrgangsstufe die für alle Schülerinnen und Schüler zu erreichenden Mindestziele auf der Basis von wissenschaftlich erarbeiteten Kompetenzmodellen zu erarbeiten (EDK, 2004). Die zugrunde liegenden Kompetenzmodelle sollen die Rückbindung der Zielvorgaben an spezifische Formulierungen von fachlichen und überfachlichen Kompetenzen erlauben, die progressiv aufbauend das Erreichen der Standards ermöglichen sollen. Bildungsstandards diesen Zuschnitt sollen sowohl der zunehmenden Mobilität der Bevölkerung als auch den gestiegenen schulischen Anforderungen Rechnung tragen. Zudem machen Forderungen der Eltern und Erwartungen aus Bildungskreisen und der Wirtschaft eine Präzisierung und Harmonisierung der Lernergebnisse notwendig.

Die Entwicklung von nationalen Bildungsstandards machen aber nur dann wirklich Sinn, wenn sie die Evaluation der Wirksamkeit des Schulsystems ermöglichen, indem die Ergebnisse einer national repräsentativen Stichprobe analysiert werden können. Im Projekt HarmoS geht es nicht um die Evaluation der einzelnen Schule oder ihrer Lehrpersonen, sondern darum, die Leistungen des Schulsystems zu überprüfen und allenfalls die Wirkung einer Schulreform präzise zu dokumentieren. Vorrangig soll das Erreichen der festgelegten Standards überprüft werden um allenfalls notwendige Unterstützungsmassnahmen zu bestimmen. Eine «empirisch konsequente Orientierung» muss sich auch der Wirkung der durchgeführten Massnahmen versichern (Helmke & Hosenfeld, 2004).

Diese Ergebnisse fliessen in ein gesamtschweizerisches Bildungsmonitoring ein. Dabei geht es um die Systematisierung und Ergänzung der für die verschiedenen Bildungstufen und Ausbildungsgänge gesammelten qualitativen und quantitativen Daten. Aufbereitungen und Analysen der Daten sollen als Entscheidungsgrundlagen für die bildungspolitischen Behörden bei Bund und Kantonen dienen.

Eine von der EDK durchgeführte Tagung vom März 2004 in Murten hatte zur Zielsetzung, neben der rein technischen Seite, wie etwa der Konstruktion eines Messinstrumentariums auch die internationalen Erfahrungen bei der Entwicklung von Bildungsstandards generell und Konzepte der Systemüberprüfung sowie auch Fragen der Implementation und gesellschaftlichen Akzeptanz zu erörtern. Die in diesem Themenband vorliegenden Beiträge sind überarbeitete Vorträge anlässlich dieser Tagung und sollen zu erweiterten theoretischen, sozialen und pädagogischen Überlegungen rund um die Evaluation von Bildungssystemen animieren, und sich in die aktuellen Überlegungen zur Erarbeitung von Bildungsstandards in der Schweiz einreihen.

Der thematische Teil dieser Ausgabe befasst sich in vier Beiträgen mit einer Anzahl von Fragen zur Gewichtung, Einführung und Verwendung von Standards aus der Perspektive der Steuerung des Schulsystems. Urs Moser diskutiert

in seinem Beitrag auf dem Hintergrund einer kritischen Analyse von PISA das Verhältnis zwischen tatsächlicher Kompetenzmessung und aktuellen Lehrplänen. Er zeigt auf, dass die gemessenen Leistungsdifferenzen von Schülergruppen der PISA-Studie mit der fehlenden Übereinstimmung zwischen den gemessenen Inhalten und den Lehrplänen zusammenhängen. Moser plädiert für die Zugrundelegung einer Kohärenz zwischen der Standarddefinition, den definierten Zielen in den Lehrplänen und der durchzuführenden Leistungsmessung.

Der Artikel von Günther Schneider befasst sich mit dem Europäischen Referenzrahmen für den Sprachenunterricht. Er zeigt auf, dass sich das etablierte Kompetenzmodell aus einer Kombination von qualitativen und quantitativen Bereichen zusammensetzt und ein generelles und nicht spezifisches Modell darstellt. Er verteidigt die Idee, dass Kompetenzniveaus auf empirischer und nicht nur theoretischer Grundlage entwickelt werden sollen und macht dabei auch auf die Problematik der Zusammenstellung einer Referenzstichprobe aufmerksam, die für das Gelingen der Arbeit bedeutungsvoll ist. Geht man von einer Schülerstichprobe im Sinne der «erstbesten Wahl» aus, riskiert man ein unterschiedliches Ergebnis zu derjenigen Stichprobe, welche sich aus Schülerinnen und Schülern zusammensetzt, die in den Genuss eines mehr immersiven Unterrichts kommen oder nach einer integralen sprachlich-didaktischen Sichtweise unterrichtet wurden. Auf jeden Fall scheint es wünschenswert, dass diesbezüglich die aktuelle Situation im Sinne einer mehr evolutiven und innovativen Weiterentwicklung überwunden wird. Schneider vertritt hierbei die These, dass sich die Zusammensetzung der Schülerstichprobe aus dem innovativen Erziehungskontext konstituieren sollte.

Dieses Problem unterstreicht zugleich die Schwierigkeit mit welcher man sich bei der Arbeit im Hinblick auf die Etablierung von Standards konfrontiert sieht. Es ist nicht immer leicht, konsistent die Phase der auf einer empirischen Grundlage entwickelten Kompetenzmodelle von derjenigen, welche sich auf die Art und Weise der Gewichtung einer Referenzpopulation bezieht zu trennen.

Der Beitrag von Michel Carbonneau beschäftigt sich mit der Frage von Standards und der damit verbundenen Problematik der Steuerung und Implementation einer Reform. Auf dem Hintergrund der in Quebec durchgeführten Curriculumsreform macht Carbonneau auf das Risiko aufmerksam, dass sich die Logik von Standards schlussendlich gegenteilig auf die angestrebten Kompetenzen auswirken können.

Auf dem Hintergrund einer vergleichenden Analyse zweier dezentralisierter Schulsysteme – England und der frankophone Teil Belgiens – stellt sich Vincent Dupriez die Frage nach der politischen Wirkung hinsichtlich der angestrebten Begründung der schulischen Systemsteuerung durch Resultate und die damit verbundenen Auswirkungen für die professionelle Praxis der Lehrkräfte. Unter Bezugnahme eines Beispiels aus England betont er, dass eine der Konsequenzen in der Infragestellung der professionellen Legitimität der Lehrkräfte zugunsten einer Legitimation des techno-demokratischen Typs stattgefunden hat. Dupriez

vertritt die Ansicht, dass die zentrale Herausforderung für ein System, welches durch Ergebnisse gesteuert wird, in der Abstimmung bzw. Übereinkunft mit der Rolle der am Prozess beteiligten Lehrkräften besteht.

Dupriez stellt ebenfalls die Frage nach der Verwendung der erhaltenen Leistungsergebnisse. Diesbezüglich macht er auf die Situation in England aufmerksam, wo die von den Schülerinnen und Schülern erbrachten Leistungen ans Bildungsministerium weitergegeben werden und im Falle von nicht befriedigenden Ergebnissen auf die Schulen Druck ausgeübt werden kann, auch seitens der Eltern.

Das Projekt HarmoS, in welchem Leistungsstandards im Hinblick auf Qualitätsentwicklung entwickelt werden, sagt bezüglich Verwendungszusammenhänge wenig aus. Es erscheint notwendig, über diesen Punkt eine wissenschaftliche Debatte zu führen. K. Bähr (2005) hat anlässlich der Plenarversammlung der CODICRE-CH (Schweizerische Konferenz der Leiter/innen von Arbeitsstellen für Schulentwicklung und Bildungsforschung) die Idee einer klaren Trennung hinsichtlich der Ebenen in die Bereiche Bildungspolitik und Verwaltung sowie Schulen und Unterricht stark gemacht. Gemäss seinen Vorstellungen wäre die Übersetzung der Ergebnisse für die Ebene der Unterrichtsentwicklung und für diejenige welche bildungspolitisch ausgerichtet ist Aufgabe und Sache einer dafür spezialisierten Instanz.

Zum Abschluss möchten wir dem Wunsch Ausdruck geben, dass dieser mit der aktuellen schweizerischen Bildungspolitik verbundene Themenband durch die Hervorhebung einiger Fragen dazu beiträgt, eine breite Debatte und weiterführende Überlegungen und Impulse für die wissenschaftliche Arbeit anzuregen.

Literatur

- Bähr, K. (2005). Erwartungen zentraler Steuerungsinstanzen an Evaluationsergebnisse. Exposé fait dans le cadre de l'assemblée plénière, Codicre-CH, Aarau, April.
- CDIP, (2004). *HarmoS – Finalités et conception du projet*. Berne: CDIP.
http://www.edk.ch/PDF_Downloads/Harmos/Harmos_Weissbuch_f.pdf
- Helmke, A.& Hosenfeld, I. (2004). Bildungsstandards und Unterrichtsqualität. *Pädagogische Führung*, 4, 173 – 176.
- Lange, H. (1999). Qualitätssicherung in Schulen. *Die Deutsche Schule*, 91, 144-159.
- Fitzner, T. (Hrsg.) (2004). *Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform*. Bad Boll: edition akademie.
- Oelkers, J. (2005). Von Zielen zu Standards. *Friedrich Jahresheft*, XXIII. 18 -19.
- Ravitch, D (1995). *National Standards in American Education*. Washington (D.C.): Brookings Institution Press.

Editorial: Standards de formation, évaluation des systèmes et développement de la qualité

Michel Nicolet et Max Mangold

La question de l'introduction de standards pour mieux piloter les systèmes de formation et accroître la qualité de l'enseignement dans le cadre de réformes nationales de l'éducation n'a cessé de prendre de l'importance dans le débat sur la politique de l'éducation. La raison en tient sans doute à l'existence de grands tests internationaux de performances tels que PISA. Plusieurs pays ont saisi l'occasion de ce « tournant empirique dans la politique éducationnelle » (Lange, 1999) pour développer des standards pour leurs propres systèmes éducatifs. On dispose depuis des résultats de nombreuses recherches entreprises au Canada, en Suède, en Finlande et dans bien d'autres pays encore qui ont permis de réunir une vaste expérience en la matière (Fitzner, 2004).

La comparaison au niveau international montre que les outils de pilotage et d'assurance de la qualité suivent une direction claire : pilotage des objectifs d'apprentissage, développement de standards, évaluation interne et externe, et « changement de cap de la culture éducative vers la prise en compte du résultat » (Oelkers, 2005). Si l'on peut reprocher aux standards utilisés habituellement dans le monde scolaire leur imprécision et leur caractère peu contraignant, et les soupçonner de vouloir se soustraire à un véritable contrôle, le concept « moderne » de standard vise quant à lui une saisie précise de la performance scolaire, mesurée sur la base des produits obtenus. Le principe de l'efficacité prend progressivement le pas sur l'orientation traditionnelle axée sur l'*input*. La majorité des concepts se basent désormais sur des standards de performance, donc sur des attentes clairement définies en matière de compétences scolaires (Ravitch, 1995). Le concept de standard ne se réfère donc pas à des performances particulières mais à une échelle servant à mesurer des performances. Si un consensus, au moins partiel, se fait jour actuellement autour de cette définition, il n'en est pas de même de la fixation du niveau des objectifs de ces standards de performance.

Alors que l'Allemagne, par exemple, définit les standards de performance comme des standards (*Regelstandards*) prescrivant le niveau moyen des performances attendues de tous les élèves, on se propose en Suisse de développer des standards dits minimaux. Les travaux entrepris dans le cadre du projet HarmoS

(Harmonisation de l'école obligatoire) de la Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP) prévoient l'élaboration d'objectifs minimaux à atteindre par tous les élèves à partir de modèles de compétences élaborés sur une base scientifique pour les domaines de la langue 1, des langues étrangères, des mathématiques et des sciences naturelles pour les degrés 2, 6 et 9 (CDIP, 2004). Ces modèles de compétences devront permettre de coupler les objectifs d'apprentissage à des formulations spécifiques de compétences disciplinaires et transversales qui aideront les élèves à atteindre progressivement le standard. Les standards de formation de ce type permettront de tenir compte aussi bien de la mobilité croissante de la population que du renforcement des exigences scolaires. De plus, les revendications des parents tout comme les attentes des milieux de la formation et de l'économie plaident également en faveur d'une plus grande précision des objectifs attendus et d'une harmonisation des résultats d'apprentissage.

La création de standards nationaux de formation n'est toutefois pertinente que dans la mesure où ces standards offrent la possibilité d'évaluer l'efficacité du système éducatif sur la base d'analyses de résultats pour un échantillonnage représentatif au niveau national. Dans le projet HarmoS il ne s'agit en aucun cas d'évaluer les établissements et leurs enseignants mais de vérifier les performances du système scolaire et éventuellement de documenter précisément l'efficacité d'une réforme scolaire. En premier lieu, il s'agit d'évaluer l'atteinte des standards donnés, de manière à prendre les mesures de soutien qui s'avèreraient nécessaires. Une «orientation empirique conséquente» doit s'assurer aussi de l'efficacité des mesures appliquées (Helmke & Hosenfeld, 2004).

Les diverses données recueillies vont s'inscrire au sein d'un monitoring national du système de formation. Il consistera en une amélioration et une systématisation des données qualitatives et quantitatives récoltées pour chaque degré d'enseignement et chaque filière d'étude, permettant une analyse approfondie et fournissant aux autorités politiques des bases de décision pertinentes.

Un congrès organisé par la CDIP en mars 2004 à Morat avait pour objectif à la fois d'étudier des aspects purement techniques, comme par exemple la construction d'un ensemble d'instruments de mesure, et de débattre des expériences internationales réalisées dans le développement de standards de formation, des concepts de vérification des systèmes, des questions d'implémentation et de l'acceptation par la société. Les textes des conférences ont été retravaillés et réunis dans ce numéro thématique. Destinées à susciter une réflexion dynamique sur les aspects théoriques, sociaux et pédagogiques de l'évaluation des systèmes éducatifs, ces contributions s'insèrent dans le débat actuel sur l'élaboration de standards de formation en Suisse.

Les quatre articles formant la partie thématique de ce numéro abordent un ensemble de questions en rapport avec la mesure ou l'utilisation des standards dans la perspective du pilotage du système éducatif. Dans sa contribution, U. Moser discute des relations entre les mesures de compétences effectuées, en se ré-

férant de manière critique aux travaux de PISA, et les plans d'études en vigueur en montrant qu'une bonne partie des différences mesurées dans l'étude PISA au niveau des performances de certains groupes d'élèves tiennent à l'absence de correspondance entre les contenus mesurés et les plans d'études. Il plaide pour l'instauration d'une cohérence entre la définition des standards, les objectifs définis dans les plans d'études et les mesures de performances réalisées.

G. Schneider dans son article basé sur le cadre européen de référence pour l'enseignement des langues montre que le modèle de compétences établi par une combinaison d'approches qualitative et quantitative aboutit de fait à un modèle général et non spécifique. Il défend l'idée que les échelles soient constituées empiriquement et non seulement sur une base théorique. Il soulève alors le problème de la constitution de l'échantillon de référence pour mener à bien ce travail. Le fait de prendre un échantillon d'élèves «tout venant» risque de conduire à un résultat différent de celui auquel on pourrait aboutir en choisissant des élèves ayant bénéficié d'un enseignement plus immersif ou s'inscrivant dans une perspective de didactique intégrée des langues. Si l'on souhaite ne pas se figer dans la situation actuelle mais se placer dans une visée évolutive et innovatrice, il défend l'idée de se doter d'échantillons d'élèves placés dans des contextes éducatifs innovants.

Ce problème souligne également la difficulté à laquelle est confronté le travail d'établissement de standards où il n'est pas toujours aisé de dissocier la phase consistant à établir des modèles de compétences sur une base empirique de celle visant à mesurer la manière dont une population de référence se situe par rapport à eux.

Avec l'article de M. Carbonneau, on s'écarte de la question des standards proprement dite pour aborder la problématique du pilotage et de l'implémentation d'une réforme. Se référant à l'importante réforme du curriculum menée au Québec, M. Carbonneau souligne le risque que la logique des standards puisse aller à fin contraire d'une approche par compétences.

S'appuyant sur l'analyse comparative de deux systèmes scolaires décentralisés, l'Angleterre et la Belgique francophone, V. Dupriez questionne l'impact des politiques visant à instaurer un pilotage du système éducatif par les résultats, sur l'implication des enseignants dans leur pratique professionnelle; se basant sur l'exemple anglais, il souligne que l'une des conséquences de ce mouvement est la mise en cause de la légitimité professionnelle des enseignants au profit d'une légitimité de type techno-démocratique. Il défend alors l'idée selon laquelle le défi majeur résiderait pour un système marqué par un pilotage par les résultats dans la place accordée aux enseignants dans ce processus.

Prenant appui sur la situation anglaise dans laquelle les mesures d'acquis scolaires des élèves sont transmises à la fois au Ministère, qui peut exercer une pression sur les établissements dont les résultats sont considérés comme insuffisants, et aux parents eux-mêmes, qui peuvent réagir en faisant pression sur l'école, V. Dupriez pose la question de l'usage qui est fait des mesures de performances. Le

projet Harnos est pour l'instant relativement discret sur l'utilisation qui sera faite des mesures effectuées dans la perspective du développement de la qualité. Il apparaît important qu'un débat scientifique puisse être ouvert à ce propos. K. Bähr (2005) dans une conférence présentée à l'assemblée plénière de la CODICRE-CH (Conférence suisse des directeurs de centres de développement scolaire et de recherche en éducation) a défendu l'idée d'une dissociation radicale des plans constitués par les domaines de la politique scolaire et de l'administration, des établissements et de l'enseignement. Il appartiendrait alors selon lui à une instance spécialisée de traduire ces mesures à la fois sur le plan du développement de l'enseignement et sur celui de l'orientation de la politique scolaire.

Pour terminer, nous exprimons le souhait que ce dossier thématique en lien avec l'actualité de la politique éducative suisse soit parvenu à souligner un ensemble de questions appelant un plus large débat et une réflexion prenant appui sur des travaux scientifiques.

Bibliographie

- Bähr, K. (2005). Erwartungen zentraler Steuerungsinstanzen an Evaluationsergebnisse. Exposé fait dans le cadre de l'assemblée plénière, Codicre-CH, Aarau, April.
- CDIP, (2004). *HarmoS – Finalités et conception du projet*. Berne: CDIP.
http://www.edk.ch/PDF_Downloads/Harnos/Harnos_Weissbuch_f.pdf
- Helmke, A. & Hosenfeld, I. (2004). Bildungsstandards und Unterrichtsqualität. *Pädagogische Führung*, 4, 173 – 176.
- Lange, H. (1999). Qualitätssicherung in Schulen. *Die Deutsche Schule*, 91, 144-159.
- Fitzner, T. (Hrsg.) (2004). *Bildungsstandards. Internationale Erfahrungen – Schulentwicklung – Bildungsreform*. Bad Boll: edition akademie.
- Oelkers, J. (2005). Von Zielen zu Standards. *Friedrich Jahresheft*, XXIII. 18 -19.
- Ravitch, D (1995). *National Standards in American Education*. Washington (D.C.): Brookings Institution Press.