

### Jacques Weiss und Michael Zutavern

PISA – inzwischen kann man nicht mehr sicher sein, ob bei diesem Wort die Vorstellung vom schiefen Turm ausgelöst wird oder mit ihm doch schon die Ergebnisse des ‚Program of International Student Assessment‘ der OECD assoziiert werden. Man bekommt den Eindruck, dass kein Vortrag ohne PISA-Zitat, keine Bildungsreform ohne Legitimation durch PISA mehr denkbar sind. Die 98-seitige Bibliografie, die das ‚Deutsche Institut für internationale pädagogische Forschung‘ (DIPF) gerade zusammengestellt hat ([www.bildungsserver.de/pdf/pisaliteratur.pdf](http://www.bildungsserver.de/pdf/pisaliteratur.pdf)), zeigt, dass auch die wissenschaftliche Diskussion durch die PISA-Studie überdurchschnittlich intensiv angeregt worden ist – und diese durchaus kontrovers verläuft (vgl. z.B. Baumert, 2001; Blum u.a., 2000; Helmke, 2000; Moser & Notter, 2000; Pekrun, 2002, Robin & Rochet, 2002).

PISA will Indikatoren für Basisfähigkeiten liefern und durch Informationen über den sozialen, institutionellen, demographischen und wirtschaftlichen Kontext, in dem Lernen geschieht, ergänzen. Durch regelmässige Datenerhebungen sollen Trendindikatoren zur Leistungsfähigkeit von Bildungssystemen geschaffen werden. Im Jahr 2000 wurden erstmals in 32 Ländern Jugendliche am Ende der obligatorischen Schulzeit auf ihre Grundbildung in den Bereichen Lesen, Mathematik und Naturwissenschaften getestet. Die Messungen wurden mit beträchtlichem Aufwand in weltweiter Kooperation von Bildungsfachleuten, insbesondere Fachdidaktikerinnen und Fachdidaktikern, vorbereitet. Sie orientierten sich zwar auch an schulischen Curricula, versuchten aber vor allem Kompetenzen zu definieren, die die Jugendlichen in ihrem Erwachsenenleben benötigen werden. Die Fähigkeit, auf unterschiedliche Situationsvorgaben adäquat reagieren zu können, Prozesse in den drei Bereichen zu beherrschen und Grundkonzepte verstanden zu haben, wurden geprüft. Diese Leistungsanforderungen wurden in Aufgaben mit unterschiedlichen Schwierigkeitsgraden operationalisiert, kontrolliert nach dem gleichen Verfahren in allen Ländern erhoben und mit aufwändigen statistischen Verfahren ausgewertet (Baumert u.a., 2001; OECD, 1999, 2001, 2002; Zahner u.a., 2002).

Neben den Leistungsdaten wurden Informationen über Lernbedingungen der Schülerinnen und Schüler in der Familie und in der Schule erfasst. So können z.B. Unterschiede der Leistungsergebnisse bei Schülerinnen und Schülern unterschiedlicher sozialer Herkunft aufgezeigt werden und Unterschiede in diesen Unterschieden zwischen den Ländern verglichen werden. Ausserdem wurden die Jugendlichen nach fachübergreifenden Fähigkeiten zum selbstregulierten Lernen und zu ihren Kenntnissen in der Informationstechnologie befragt. Auch das Leistungsselbstbild und Aussagen der Schülerinnen und Schüler über ihre Lernfähigkeiten konnten in Beziehung gesetzt werden zu ihren Leistungsergebnissen. Die Ergebnisse liegen in internationalen und nationalen Berichten vor. In der Schweiz wurde das Projekt von der EDK und vom Bundesamt für Statistik getragen. Ergänzend zum internationalen Vergleich wurden durch zusätzliche Stichproben auch innerschweizerische Vergleiche zwischen den Sprachregionen und zwischen einzelnen Kantonen möglich. Zur Zeit erscheinen vertiefende und ergänzende Studien, in denen z.B. Vergleichen mit den am besten abschneidenden Ländern nachgegangen wird oder Vergleiche zwischen PISA-Aufgaben und schulischen Curricula angestellt werden (vgl. [www.statistik.admin.ch/stat\\_ch/ber15/pisa/pisa\\_d.htm](http://www.statistik.admin.ch/stat_ch/ber15/pisa/pisa_d.htm)).

Obwohl bislang nur beschreibende Studien vorliegen, wird in der öffentlichen Diskussion bereits manches mit PISA erklärt und von ganz unterschiedlichen Seiten die heilsame Wirkung von Massnahmen versprochen, die angeblich aus PISA ableitbar seien: von der Einführung der Basisstufe, der stärkeren Kontrolle der Schule, der Förderung der Bibliotheken, der Abschaffung von Noten, der Einführung von standardisierten Tests, der Separation der fremdsprachigen Schüler, der Integration bis zur 9. Klasse – um nur einige der teilweise recht widersprüchlichen Beispiele zu nennen. Während die vertiefenden Studien noch in Arbeit sind und für die zweite Staffel von PISA – mit der mathematischen Grundbildung im Zentrum – gerade die Erhebungen laufen, besteht die Gefahr, dass der nüchterne Blick auf die Aussagekraft der bislang vorliegenden Ergebnisse verloren geht. PISA liefert eine Datenbasis, wie sie bisher nicht zur Verfügung stand – aber auch deren Aussagekraft unterliegt Beschränkungen. An einige soll erinnert werden:

- PISA testete 15-Jährige und Schülerinnen und Schüler der 9. Klasse – über die Leistungsfähigkeit jüngerer Schülerinnen und Schüler sagt sie nichts aus, auch nichts darüber, wo und wie eventuelle Leistungsmängel in den vorangegangenen Jahren entstanden sein könnten.
- Getestet wurden die erwähnten drei Leistungsbereiche, die sicher wichtige Basiskompetenzen umfassen, die für eine aktive Teilnahme der künftigen Erwachsenen in der Gesellschaft notwendig sind. Speziell die Lesefähigkeit hat dabei eine Schlüsselfunktion inne. Aber die Zielsetzungen von Schule sind selbstverständlich reichhaltiger. Viel Wissen und manche Fähigkeit, um deren Förderung sich Lehrerinnen und Lehrer bemühen, sind in PISA nicht geprüft worden. Das reicht vom Texteschreiben über historisches Verstehen bis

- zu musikalischen Fähigkeiten. Soziale Kompetenz wurde nicht erfasst und die Fähigkeit zum Lernen zwar erfragt, aber nicht getestet.
- Die jetzt vorliegenden Ergebnisse sind in Mittel- und Streuwerten, in Varianzen und Korrelationen ausgedrückt. Es werden also Unterschiede, Zusammenhänge und auffällige Unterschiede im Vergleich dieser Werte analysiert. Solche Analysen sollten bekanntermassen nicht mit Kausalitäten verwechselt werden. Allenfalls werden durch überzeugende theoretische Grundlagen kausale Zusammenhänge nahegelegt – ihre genauere Untersuchung steht jedoch noch aus.
  - Bei den Wiederholungen von PISA werden wieder die jetzt 15-Jährigen getestet. Es erfolgt also keine Messung der Leistungsentwicklung von Schülerinnen und Schülern, sondern allenfalls eine Art Längsschnitzaussage über Bildungssysteme.
  - Die PISA-Tests müssen in verschiedene Sprachen übersetzt und in unterschiedlichen Kulturen vergleichbar verständlich konzipiert werden. Dies geschieht in einem mehrphasigen und kontrollierten Prozess. Gleichwohl erscheint es nötig, in methodenkritischen Studien mögliche Einflüsse dieses Vorgehens auf die Ergebnisse zu analysieren.

Die Beiträge in diesem Themenheft tragen in diesem Sinne einige ergänzende Perspektiven zu PISA zusammen: Möglichkeiten und Grenzen der Studie werden kritisch analysiert und Daten aus zwei kantonalen Erweiterungsstudien zum Blick auf den besonders kritischen Bereich sozialer Ungleichheiten genutzt. Schliesslich wird an einem Beispiel gezeigt, dass es nicht nur auf die Ergebnisse von Bildungsevaluationen ankommt, sondern auch ihre Rezeption ein interessanter Gegenstand empirischer Forschung sein kann.

*Christian Nidegger*, cheville ouvrière de la recherche PISA en Suisse romande, (Nidegger 2002) a pu saisir, peut-être mieux que quiconque, la richesse d'une étude internationale, tant sur le plan scientifique que sur celui des coopérations régionales et nationales. Mais il est celui qui a pu également en mesurer la complexité, les limites et les dérives. L'auteur structure son article autour de trois entrées, l'organisation, l'impact politique, le défi scientifique. Au plan de l'organisation, PISA a offert l'opportunité aux institutions de recherche et aux chercheurs de Suisse romande de coopérer au sein d'une forme nouvelle d'organisation, créée pour l'occasion. Le consortium de recherche, ainsi nommé, a permis de fédérer des compétences et des ressources, et de développer une nouvelle manière de «chercher ensemble». Toutefois l'interpénétration des niveaux de gestion de la recherche, de l'international au cantonal, en passant par le national et le régional, est une source de tensions et de réglages permanents. Quant à l'impact politique des résultats, Christian Nidegger considère qu'il a été considérable en Suisse et en Suisse romande et que le débat, souvent très vif, a généré d'importantes et abusives interprétations. C'est au défi scientifique que l'auteur consacre l'essentiel de son propos en soulignant d'abord la complexité de la méthodologie de cette recherche. Il relève notamment que la recherche PISA n'a pas

produit des résultats surprenants pour les chercheurs, mais a plutôt confirmé des tendances observées dans des recherches antérieures dont l'interprétation, dans cette évaluation comme dans les précédentes études de ce type, ne permet pas de tirer de conclusions simples et définitives, mais des faisceaux de réponses possibles, que d'autres recherches, plus qualitatives permettront de compléter et d'éclairer.

Christian Nidegger a souligné la complexité scientifique et de gestion administrative de la recherche PISA, et a tenté de montrer les précautions qu'exigent l'interprétation des résultats et leur exploitation au plan de la pratique de l'enseignement. Pour sa part, *Daniel Bain* prend à son compte les interrogations de Christian Nidegger et analyse la validité didactique d'une telle enquête. Prenant pour objet d'étude la lecture, compétence-cible de l'évaluation PISA, il dresse un constat sévère et critique de cette enquête, et exprime quelques doutes quant à l'efficacité de tels programmes internationaux de recherche pour fournir une aide à la didactique. Son argumentation est étayée par une mise en question du cadre conceptuel de l'étude, et de son opérationnalisation dans l'évaluation, produit d'un compromis plutôt que d'un modèle consensuel de validation des items. Il conteste les mises en situation des évaluations qui n'en sont pas, et décrit les faiblesses des «niveaux de compétences en compréhension de lecture», construction hétérogène aux frontières arbitraires. Il regrette le choix de la méthode statistique qui donne priorité aux contraintes psychométriques plutôt qu'aux exigences critérielles correspondant à des tâches de lecture valides, socialement pertinentes. Sceptique quant à l'intérêt de telles enquêtes, Daniel Bain tient tout de même pour positif l'impulsion que pourraient donner ces enquêtes aux travaux de didactique relatifs à l'apprentissage de la lecture ou traitant de la littéracie.

*Christine Barré-De Miniac* engage un propos qui, sous l'impulsion du débat actuel sur les compétences en lecture, vient clarifier le notion de littéracie et montrer l'étendue des recherches à mener dans ce domaine. Les résultats de la recherche PISA en deviennent tout relatifs, et considérés à leur juste dimension. L'auteur expose avec rigueur et solides références, l'émergence de la notion, ses potentialités et sa souplesse pour la recherche en didactique. La notion est référée à diverses théories ou modèles, toujours en débat; et la littéracie générique, mise en question par des littéracies multiples. De surcroît nombre de questions sont encore sans réponses comme les liens existant entre syntaxe et capacités discursives ou le rôle du langage dans la littéracie en langue seconde. Des recherches nouvelles s'imposent donc, hors milieu scolaire, en milieu familial. Christine Barré-De Miniac souligne l'urgence de mener également des études descriptives des pratiques scolaires ou non-scolaires des apprenants mais aussi des enseignants, s'inscrivant soit dans le courant de l'étude des interactions verbales en classe, soit dans le courant ergonomique de l'analyse du travail de l'enseignant. La notion de littéracie est donc porteuse et prometteuse de recherches, et peut-être fédératrice des travaux en didactique, considérés par l'auteur comme encore

trop cloisonnés. Elle regretterait toutefois que la recherche «s'empare de cette notion dans le seul but de calculer des chiffres et des pourcentages, l'étalage de chiffres servant d'alibi à l'absence de véritables recherches».

*Christian Brühwiler und Erich Ramseier* werten die kantonalen Stichproben aus Bern und St. Gallen mit besonderem Blick auf Aspekte der sozialen Chancengerechtigkeit aus. In beiden Kantonen wurde zusätzlich ein Intelligenztest eingesetzt. In Pfadanalysen zeigt sich die erwartungsgemäss hohe Beziehung dieser kognitiven Grundfähigkeiten mit den Lese-, Mathematik- und naturwissenschaftlichen Leistungen in unterschiedlicher Weise, wobei die zentrale Rolle des Lesens bestätigt wird. Dabei kommt dem Schultyp, den die Schülerinnen und Schüler besuchen, eine besondere Rolle zu. Die Chance, das Gymnasium zu besuchen, hängt bei gleichen Intelligenzwerten stark von der Schichtherkunft ab. Darüber hinaus zeigen sich auch Unterschiede, die teilweise durch die unterschiedliche soziale Zusammensetzung von Schulen des gleichen Typs erklärt werden können. Interessanterweise scheint auch das Zugehörigkeitsgefühl zur Schule, das für individuelle Leistungsunterschiede keine Rolle spielt, den Schulleffekt zum Teil erklären zu können – was die Ergebnisse anderer Motivations- und Schulklimastudien bestätigen könnte (Rutter, 1980; Ryan & Deci, 2000).

Betrachtet man PISA als Diagnose von Bildungssystemen, wird schnell klar, dass diese nur etwas nützt, wenn ihre Ergebnisse auch diejenigen erreicht, die für eine Therapie zuständig sind. Und die Therapeuten müssen die Diagnosedaten dann auch noch verstehen, richtig interpretieren und ihre Massnahmen nach ihnen ausrichten wollen. All das scheint keinesfalls sichergestellt. Forschung sollte also auch für mehr Wissen über die Verarbeitung von Bildungsevaluationen bei den verschiedenen Verantwortlichen des Bildungswesens sorgen. *Friedrich-Wilhelm Schrader und Andreas Helmke* beschreiben eine solche Studie, in der sie eine landesweite Leistungsevaluation auf ihre Rezeption in Schulen untersuchten. Die Ergebnisse der Befragung von Schulleiterinnen und Schulleitern zeigt, dass zur aktiven Reaktion auf Evaluationsergebnisse weniger deren Resultat als die Einstellung zur Evaluation und die Motivation der Betroffenen zur Veränderung beitragen.

Die Fortsetzungen von PISA bieten reichliche Gelegenheiten, diese Einstellungen und den Veränderungswillen auf die Probe zu stellen. Zur Zeit werden die drei Grundfähigkeiten wieder erhoben – mit Mathematik im Schwerpunkt, für das 2000 gute schweizerische Vergleichswerte vorlagen. 2006 dagegen werden die naturwissenschaftlichen Fähigkeiten im Zentrum stehen, die im ersten Durchgang wie die Lesefähigkeiten durchschnittliche Werte zeigten und deren Umfang und Struktur in den schulischen Curricula nicht unumstritten sind. Die Redaktion dieser Zeitschrift wird auch weiterhin den kritischen Blick auf dieses bedeutende bildungswissenschaftliche Projekt richten.

- Baumert, J. (2001). Evaluationsmassnahmen im Bildungsbereich. Eine kritische Sicht auf mögliche Zugänge. *Newsletter der Österr. Gesellschaft für Forschung und Entwicklung im Bildungswesen*, 1, (2), 3-15.
- Baumert, J., Klieme, E., Neubrand, M. u.a. (2001). *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske + Budrich.
- Blum, W., Demmer, M., Groeben, A. & Tillmann, K.-J. (2000). Leistungsvergleiche – und kein Ende. *Schwerpunktthema Heft 12, Pädagogik*, 52, (12).
- Helmke, A. (2000). TIMSS und die Folgen: Der weite Weg von der externen Leistungsevaluation zur Verbesserung des Lehrens und Lernens. In U. P. Trier (Hrsg.), *Bildungswirksamkeit zwischen Forschung und Politik* (S. 135-164). Chur & Zürich: Rüegger.
- Moser, U. & Notter, P. (2000). *Schulische Leistungen im internationalen Vergleich*. Bern: Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK).
- Nidegger, Ch. (Ed.) (2002). *Compétences des jeunes romands, résultats de l'enquête PISA 2000 auprès des élèves de 9ème année*. Neuchâtel: Institut de recherche et de documentation pédagogique (IRDPI).
- OECD (1999). *Measuring student knowledge and skills: A new framework for assessment*. Paris: OECD.
- OECD (2001). *Lernen für das Leben. Erste Ergebnisse der Internationalen Schulleistungsstudie PISA 2000*. Paris: OECD.
- OECD (2002). *Reading for Change. Performance and engagement across countries. Results from PISA 2000*. Paris: OECD.
- Pekrun, R. (2002). Vergleichende Evaluationsstudien zu Schülerleistungen: Konsequenzen für zukünftige Bildungsforschung. *Zeitschrift für Pädagogik*, 48, (1), 111-128.
- Robin, I. & Rocher, T. (2002). La compétence en lecture des jeunes de 15 ans: une comparaison internationale. Données sociales 2002 à 2003 – Education, formation. Paris: INSEE, 93-101. [www.insee.fr/fr/ffc/docs\\_ffc/ds019fr.pdf](http://www.insee.fr/fr/ffc/docs_ffc/ds019fr.pdf)
- Rutter, M., Maughan, B., Mortimore, P. & Ouston, J. (1980). *Fünfehtausend Stunden: Schulen und ihre Wirkung auf Kinder*. Weinheim & Basel: Beltz.
- Ryan, R.M. & Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development and wellbeing. *American Psychologist*, 55, 68-78.
- Zahner, C., Meyer, A., Moser, U. u.a. (2002). *Für das Leben gerüstet? Die Grundkompetenzen der Jugendlichen – Nationaler Bericht der Erhebung PISA 2000*. Neuchâtel: Bundesamt für Statistik.
- Zahner, C., Meyer, A., Moser, U. et al (2002). *Préparé pour la vie ? Les compétences de base des jeunes – rapport national de l'enquêt PISA 2000*. Neuchâtel: Office fédéral de la statistique.