

REVUE FRANÇAISE
DE
PÉDAGOGIE

Revue française de pédagogie

Recherches en éducation

157 | octobre-décembre 2006

PISA : analyses secondaires, questions et débats
théoriques et méthodologiques

Introduction

Jean-Yves Rochex



Édition électronique

URL : <http://journals.openedition.org/rfp/391>

DOI : 10.4000/rfp.391

ISSN : 2105-2913

Éditeur

ENS Éditions

Édition imprimée

Date de publication : 1 décembre 2006

Pagination : 5-10

ISBN : 978-2-7342-1075-7

ISSN : 0556-7807

Référence électronique

Jean-Yves Rochex, « Introduction », *Revue française de pédagogie* [En ligne], 157 | octobre-décembre 2006, mis en ligne le 01 décembre 2010, consulté le 03 mai 2019. URL : <http://journals.openedition.org/rfp/391> ; DOI : 10.4000/rfp.391

« performances » de leurs systèmes éducatifs, ce qui n'est pas sans poser problème quand on sait que ces données ont été produites à l'aide d'épreuves et selon des méthodologies d'enquête sur lesquelles les pays ne disposant pas ou guère d'expérience, ni de compétences en la matière, n'ont pu avoir que très peu d'influence.

Ces enquêtes et les données qu'elles permettent de produire ne sont pas considérées, par les organismes intergouvernementaux et gouvernementaux, par les décideurs et les experts qui les promeuvent et les réalisent, comme de seuls outils de connaissance, mais aussi, si ce n'est d'abord, comme des moyens et des occasions d'évaluer et d'infléchir les politiques éducatives et le fonctionnement des systèmes éducatifs des pays concernés. Le présupposé est qu'évaluer et comparer les « performances » des différents systèmes éducatifs (et, au sein de ceux-ci, de leurs différentes composantes et unités), au regard d'épreuves et d'indicateurs communs, serait le moyen le plus pertinent d'améliorer leur efficacité et, par là, d'améliorer la compétitivité économique des pays concernés. Les technologies de l'évaluation et de l'*accountability* sont dès lors considérées et mises en œuvre comme instruments au service de la gouvernance et de l'amélioration des systèmes éducatifs, à l'échelle nationale et internationale, de même qu'au sein des différents pays, les résultats produits par les travaux s'inscrivant dans le courant de la *School Effectiveness Research* et visant à évaluer l'efficacité des établissements sont pensés et utilisés comme devant permettre d'améliorer cette efficacité. Ces travaux, leurs fondements théoriques ou idéologiques, leurs méthodologies, leurs résultats, leurs effets positifs ou pervers sur les systèmes éducatifs ou les établissements, ont été et demeurent au centre de nombreuses publications et de nombreux débats, dont le dossier que nous publions, consacré à PISA ne prétend bien sûr pas rendre compte de manière exhaustive, et sur lesquels la *Revue française de pédagogie* se propose de revenir (2).

On pourra néanmoins lire dans ce dossier six articles abordant et rendant compte sous différents angles : de questions et débats théoriques et méthodologiques portant sur l'enquête PISA elle-même ; des effets que la publication et les commentaires des résultats de PISA ont pu avoir sur les politiques scolaires de différents pays ; ou encore des possibilités d'analyses secondaires auxquelles ces données peuvent donner ou ont déjà donné lieu, que ces analyses concernent la confrontation des résultats de PISA à ceux produits par d'autres enquêtes internationales, par des enquê-

tes et des tests complémentaires menées à l'échelle nationale, ou par d'autres modes d'investigation visant à saisir, au-delà des scores obtenus, les modes de faire des élèves avec les épreuves qui leur étaient soumises (3). Trois de ces articles concernent plus spécifiquement la littéracie et son évaluation, auxquelles la première édition de PISA, en 2000, accordait un poids prépondérant, qui sont de longue date l'objet de nombreuses autres enquêtes nationales ou internationales, et qui dès lors ont pu être au cœur de différentes analyses secondaires et d'importants débats théoriques et méthodologiques.

Ainsi Rolf V. Olsen et Svein Lie s'efforcent-ils de présenter un panorama structuré des perspectives de recherches, primaires et secondaires, ouvertes par les enquêtes internationales telles que PISA ou TIMSS. Plaidant pour que les données produites par ce type d'enquêtes soient mieux prises en considération par les chercheurs, et pour un meilleur dialogue entre chercheurs et décideurs, ainsi qu'entre objectif de recherche et objectif d'amélioration de l'efficacité des systèmes éducatifs, ils dessinent un cadre général permettant de classer les perspectives et méthodologies d'analyses secondaires en six types d'objectifs et d'approches de recherche, en donnant de nombreux exemples concernant l'enseignement des sciences, lequel est leur principal domaine de recherche, tout en souhaitant que ces différents types d'analyses secondaires soient mieux soutenus et financés par les gouvernements nationaux.

Pour autant, comme le reconnaissent ces auteurs, les données des enquêtes internationales sont non seulement d'une grande richesse mais d'une extrême complexité, tant elles sont le produit de l'interaction entre la manière dont sont conceptualisées et définies les compétences que l'on vise à mesurer, et les méthodologies fort sophistiquées que l'on met en œuvre pour ce faire. L'article de Pierre Vrignaud s'efforce de faire le point sur ces interactions concernant la mesure de la littéracie dans PISA. Présentant les théories psychométriques et le modèle de mesure utilisés dans PISA, il insiste non seulement sur l'aspect très technique de ces choix et procédures – lesquels rendent difficile la discussion des modalités et des résultats de l'enquête par les non-spécialistes –, mais aussi sur les conséquences de cette approche sur l'analyse des résultats. Il interroge pour finir le postulat d'unidimensionnalité de la compétence évaluée et la conception même de la littéracie qui sous-tend l'enquête PISA, ses épreuves et consignes de correction et de traitement statistique, faisant ainsi écho à certaines critiques d'ordre méthodologique (Bonnet,

2002 ; Goldstein, 1995 & 2004 ; Rocher, 2003) et à certaines analyses qui se demandent si le souci et l'exigence psychométriques n'ont pas pris le dessus sur le souci et l'exigence d'élaboration théorique de ce qu'est la littéracie, voire s'ils ne masquent pas la faiblesse ou « l'absence d'une théorie élaborée organisant les différentes dimensions de la compréhension de l'écrit en tant qu'activité langagière et proposant un point de vue déterminé sur les opérations de lecture » (Bain, 2003 ; Goody, 2001 ; cf. également sur ce point Rochex, 2006).

Les deux articles suivants rendent compte des effets ou des retombées de l'enquête PISA, et des prolongements auxquels elle a donné lieu en Suisse et en Allemagne. La publication des résultats de PISA a en effet suscité dans chacun de ces deux pays de très vives réactions et de très importants débats politiques, médiatiques, pédagogiques et institutionnels (concernant l'Allemagne, cf. Gruber, 2006), liés sans doute au fait que ces résultats tendaient à montrer que les systèmes éducatifs de ces deux pays seraient à la fois moins efficaces et moins équitables que la moyenne des pays de l'OCDE. Ces constats et les débats auxquels ils ont donné lieu ont entraîné des investigations complémentaires et contribué à des modifications structurelles ou curriculaires dont rendent en partie compte ces deux articles. Le premier, que nous devons à Jean Moreau, Christian Nidegger et Anne Soussi, s'interroge sur l'adéquation du modèle et des variables statistiques utilisés par PISA pour rendre compte de spécificités locales, à partir du constat selon lequel certains items considérés comme étant de difficulté identique obtenaient néanmoins des pourcentages de réussite très différents dans des populations scolaires pourtant comparables, ou encore du constat selon lequel le lien statistique entre niveau de performance des élèves (ici en mathématiques) et indice socio-économique s'avère être très variable d'un canton à l'autre. Les auteurs présentent ensuite les grandes lignes d'un projet d'harmonisation structurelle et curriculaire à l'échelle fédérale, qui s'inspire largement de démarches mises en œuvre dans PISA, malgré des objectifs différents.

Manfred Prenzel et Karin Zimmer, quant à eux, présentent les investigations complémentaires auxquelles a donné lieu l'enquête PISA 2003 en Allemagne, concernant particulièrement l'enseignement des sciences et la résolution de problèmes : élargissement de l'échantillon permettant de situer les seize *Länder* et leurs spécificités, au regard des données internationales et nationales, mais aussi permettant de disposer de données sur des catégories de

population originaires de pays différents ou arrivées en Allemagne à des époques différentes ; tests et questionnaires complémentaires permettant d'élargir l'éventail et la nature des données recueillies ; répétition de certains de ces tests permettant de recueillir des données sur l'évolution des compétences des individus ou des classes au terme d'une année scolaire. L'ensemble de ces données ont rendu possible l'évaluation de la conformité des tests nationaux et internationaux aux programmes des différents filières de scolarisation dans les différents *Länder*, ou l'étude des disparités entre *Länder* concernant leur niveau moyen dans les différents domaines évalués, les rapports entre les différents types de tests ou encore le lien, variable d'un *Land* à l'autre, entre origine sociale et performances des élèves.

Quant à Martine Rémond, elle met en rapport les résultats des élèves français à deux enquêtes internationales portant sur la *reading literacy*, l'enquête PISA 2000 portant sur des élèves de quinze ans, et l'enquête PIRLS, réalisée par l'IEA, et portant sur des élèves d'une dizaine d'années, résultats qui sont également rapprochés de ceux des dernières évaluations-bilans réalisées à l'échelle nationale par la Direction de l'évaluation et de la prospective du ministère de l'Éducation nationale. Ces mises en rapport et analyses croisées permettent de mettre au jour des profils de résultats et des caractéristiques propres à la population des élèves français, des écueils ou des comportements obérant leur réussite aux épreuves d'évaluation qui leur ont été soumises. La confrontation de ces données avec les réponses fournies par les maîtres ayant en charge les élèves évalués par PIRLS à un questionnaire les interrogeant sur leurs pratiques en matière d'exposition de leurs élèves à l'écrit, permet ensuite à l'auteur de s'interroger sur l'efficacité de telles pratiques déclarées, et de formuler quelques recommandations sur les contenus et les stratégies d'enseignement susceptibles de contribuer à l'amélioration des compétences des élèves en compréhension et en production d'écrit.

Enfin, Élisabeth Bautier, Jacques Crinon, Patrick Rayou et Jean-Yves Rochex rendent compte d'un travail d'analyse secondaire et complémentaire des épreuves de PISA 2000 portant sur les compétences en littéracie. Cette recherche a conjugué un travail d'analyse *a priori* de ces épreuves et de leurs modalités de correction et de cotation, un re-traitement statistique des réponses de plus de 800 élèves, mené à partir d'hypothèses théoriques différentes de celles qui ont présidé à la réalisation de l'enquête PISA, et un recueil et une analyse d'entretiens visant à mieux

comprendre les modes de faire de différentes catégories d'élèves face aux épreuves considérées, à mieux identifier les univers auxquels ces élèves se réfèrent pour comprendre les textes et répondre aux questions qui leur étaient proposés. L'ensemble des analyses ainsi produites ont conduit leurs auteurs à mettre au jour le caractère problématique et incertain des rapports entre résultats et modes de faire des élèves ; à mettre en évidence que le postulat de stabilité et d'homogénéité des compétences que les épreuves de PISA permettaient d'évaluer n'était vérifié que pour les élèves les plus performants ou les plus faibles, mais qu'il ne pouvait être soutenu concernant plus de 40 % des élèves ; et par conséquent, à interroger de manière critique tant les présupposés méthodologiques que les conceptions théoriques concernant la littéracie qui ont présidé à l'élaboration et à la mise en œuvre de l'enquête PISA.

Pour conclure, il convient de dire que les articles composant ce dossier ne représentent qu'un éventail partiel et fort incomplet des recherches et analyses auxquelles PISA et les enquêtes internationales ont donné et peuvent donner lieu, ou encore des travaux, débats et réflexions que suscitent ce type d'enquêtes ainsi que leurs (més-)usages sociaux, politiques, institutionnels ou médiatiques. Questions de recherche, débats scientifiques et enjeux sociaux et politiques qui appellent le développement de travaux, d'échanges et de controverses dont ce dossier n'est qu'un jalon, et auxquels la *Revue française de pédagogie* entend bien contribuer de nouveau dans l'avenir.

Jean-Yves Rochex
Université Paris 8
Équipe ESCOL

NOTES

- (1) Le titre de la version anglaise de ce même ouvrage est *Monitoring Educational Achievement*.
- (2) Parmi les publications s'inscrivant dans ce débat, citons, sans souci d'exhaustivité, Ball, 1990 ; Goldstein, 2001 & 2002 ; Goldstein & Myers, 1997 ; Goldstein & Woodhouse, 2000 ; Normand, 2003

- & 2004 ; Reynolds & Teddlie, 2000 ; Teddlie & Reynolds, 2000 ; Thrupp, 2001.
- (3) Pour un autre exemple d'analyse secondaire, cf. Duru-Bellat, Mons & Suchaut, 2004.

BIBLIOGRAPHIE

- BAIN D. (2003). « PISA et la lecture : un point de vue de didacticien ». *Revue suisse des sciences de l'éducation*, vol. 25, n° 1, p. 59-78.
- BALL S. J. (1990). « Management as moral technology : a Luddite analysis ». In S. J. Ball (éd.), *Foucault and Education : disciplines and knowledge*. London : Routledge, p. 153-166.
- BONNET G. (2002). « Reflexion in the critical eye : on the pitfalls of international assessment ». *Assessment in Education*, vol. 9, p. 387-399.
- BOTTANI N. & VRIGNAUD P. [éd.] (2005). *La France et les évaluations internationales*. Rapport établi à la demande du Haut conseil d'évaluation de l'école. Paris : Haut conseil de l'évaluation de l'école.
- DURU-BELLAT M. ; MONS N. & SUCHAUT B. (2004). « Caractéristiques des systèmes éducatifs et compétences des jeunes de 15 ans. L'éclairage des comparaisons entre pays ». *Les Cahiers de l'IREDU*, n° 66.
- GEWIRTZ S. (1998). « Can All Schools be Successful ? An exploration of the determinants of "school success" », *Oxford Review of Education*, vol. 24, n° 4, p. 439-457.
- GEWIRTZ S. ; BALL S. J. & BOWE R. (1995). *Markets, Choice and Equity in Education*. Buckingham : Open University Press.
- GOLDSTEIN H. (1995). « Interprétation des comparaisons internationales des résultats scolaires ». In UNESCO, *Études et documents d'éducation*. Paris : UNESCO, p. 11-32.
- GOLDSTEIN H. (2001). « Using pupil performance data for judging schools and teachers : scope and limitations ». *British Educational Research Journal*, vol. 27, n° 4, p. 433-442.
- GOLDSTEIN H. (2004). « International comparisons of students attainment : some issues arising from the PISA study ». *Assessment in Education*, vol. 11, n° 3, p. 319-330.
- GOLDSTEIN H. & MYERS K. (1997). « School effectiveness research : a bandwagon, a hijack or a journey towards enlightenment ? ». Paper presented at *British Educational Research Association* meeting, York, 11-14 september.
- GOLDSTEIN H. & WOODHOUSE G. (2000). « School Effectiveness Research and Educational Policy », *Oxford Review of Education*, vol. 26, n° 3-4, p. 353-363.
- GOODY J. (2001). « Competencies and Education : Contextual Diversity ». In D. S. Rychen & L. S. Salganik (éd.), *Defining and Selecting Key Competencies*. Göttingen : Hogrefe & Huber, p. 175-190.
- GRUBER K. H. (2006). « The German "PISA-shock" : some aspects of the extraordinary impact of the OECD's PISA study on the German education system ». In H. Ertl

- (ed.), *Cross-national Attraction in Education. Accounts from England and Germany*. Oxford: Symposium Books, p. 195-208.
- MASTERS G. N. (1995). « Élaboration d'échelles de mesure et agrégation dans les études de l'IEA : critique du rapport du Professeur Goldstein ». In UNESCO, *Études et documents d'éducation*. Paris : UNESCO, p. 33-39.
- NORMAND R. (2001). « Le mouvement de la *school effectiveness* et sa critique dans le monde anglo-saxon ». *Revue de l'institut de sociologie [de l'Université libre de Bruxelles]*, n° 1-4, p. 135-166.
- NORMAND R. (2004). « Les comparaisons internationales de résultats : problèmes épistémologiques et questions de justice ». *Éducation et sociétés : revue internationale de sociologie de l'éducation*, n° 12, p. 73-89.
- ORIVEL F. (2005). « Foreword ». In N. Postlethwaite, *Monitoring Educational Achievement*. Paris : UNESCO.
- POSTLETHWAITE N. (2005). *Monitoring Educational Achievement*. Paris : UNESCO.
- REYNOLDS D. & TEDDLIE C. (2000). « Reflections on the critics and beyond them ». Paper presented at *American Educational Research Association Annual Meeting*, New Orleans, 24-28 april.
- ROCHER T. (2003). « La méthodologie des évaluations internationales de compétences ». *Psychologie et psychométrie*, vol. 24, n° 2-3, p. 117-147.
- ROCHEX J.-Y. (2006). « Social, Methodological and Theoretical Issues Regarding Assessment Lessons From a Secondary Analysis of PISA 2000 Literacy Tests ». *Review of Research in Education*, vol. 30, p. 163-212.
- TEDDLIE C. & REYNOLDS D. (2000). « Responses to the criticisms of school effectiveness research contained in Slee, Weiner & Tomlinson (1998) and Thrupp (1999) ». Paper presented at *American Educational Research Association Annual Meeting*, New Orleans, 24-28 april (quoted in Thrupp, 2001).
- THRUPP M. (2001). « Recent School Effectiveness Counter-critiques : problems and possibilities », *British Educational Research Journal*, vol. 27, n° 4, p. 443-457.