Introduction

Les évaluations internationales et l'éducation comparée

Marc DEMEUSE
Institut d'Administration Scolaire
Université de Mons,
Belgique

Les grandes évaluations internationales fournissent des données très riches aux chercheurs (Olsen et Lie, 2006 ; Loye, 2011 ; Demeuse, 2012), mais elles nourrissent aussi les prises de positions des décideurs politiques et des institutions nationales et internationales, à propos de comment fonctionne l'école et, surtout, comment celle-ci devrait fonctionner ou ce qu'il conviendrait de faire pour qu'elle fonctionne mieux. Comme le précise Ozga (2012, p. 168) dans l'introduction d'un récent numéro de l'*European Educational Research Journal* consacré à PISA, « ces politiques (à de rares exceptions) ne se considèrent pas comme redevables, face à leur électorat, des faibles performances établies par PISA. Ils utilisent plutôt PISA pour justifier les orientations politiques qu'ils se voient déjà adopter dans leurs propres contextes ». En quelque sorte, passant très vite à côté d'un dangereux accroissement de la reddition de comptes par rapport au passé (*accountability*) (Dupriez et Mons, 2011), certains décideurs orienteraient insidieusement l'usage de ces évaluations internationales vers une forme de justification de ce qui devrait advenir et de la nécessaire mobilisation de bonnes pratiques. Mais les choses ne sont sans doute pas si simples…

Si les chercheurs ont longtemps gardé le monopole des travaux comparatifs dans le domaine de l'évaluation, jusqu'à l'avènement du *Programme international pour le Suivi des Acquis des élèves* (PISA) de l'*Organisation de Coopération et de Développement Economiques* (OCDE) (Felouzis et Charmillot, 2012), ils doivent à présent partager aussi bien

les choix stratégiques que les décisions méthodologiques en la matière.

En soi, cela n'a rien d'anormal et si ces études ont gagné en notoriété auprès d'un large public, y compris à travers les classements internationaux (Champollion et Barthes, 2012), les chercheurs ont, à présent, à la fois à s'intéresser à la manière dont les données sont récoltées et analysées, et aux éventuels effets sur les systèmes éducatifs des résultats que certains d'entre eux contribuent à produire. L'un des objectifs essentiels est donc aujourd'hui pour les chercheurs en éducation d'interroger aussi bien les résultats (le premier objectif proposé par Olsen et Lie (2006), au-delà des classements ou des certitudes affichées par certains à propos de recettes qui auraient fait leurs preuves ailleurs et qu'il suffirait d'adapter, que les changements que ces résultats peuvent induire. Pour Olsen et Lie (2006), cela n'interdit donc pas de s'intéresser à l'amélioration des systèmes éducatifs et à l'accroissement de leur efficacité, ce qui constitue, selon ces deux auteurs, le second objectifs de ces études.

Certains chercheurs se sentent pourtant dépouillés ou exclus, même si, face aux grandes évaluations internationales du type PISA, « il est sans doute trop facile d'accuser le développement des 'statistiques d'Etat' de 'déposséder' le milieu des sociologues de l'usage des statistiques (comme le suggère Poupeau, 2003). Car rien n'empêche les chercheurs d'exploiter ces données, à condition d'en être capable techniquement »… puisque les données PISA sont disponibles, gratuitement, en ligne, comme le rappelle Duru-Bellat (2007, p. 12). Et de poursuivre, cette fois en pointant l'absence relative de débat quant aux données récoltées pour privilégier au mieux une critique de principe, au pire une sorte de rejet sans autre forme de procès : « refuser l'objectivation chiffrée (ce qui n'interdit pas la nuance), publique et donc critiquable, renverrait à un obscurantisme préférant le huis clos entre enseignants et élèves où l'ineffable qui y est produit n'est pas nécessairement positif et juste » (Duru-Bellat, 2007, p. 12).

Les chercheurs ne doivent donc pas se priver de développer des modèles alternatifs par rapport à ceux qui sont proposés par les instances nationales (par exemple, en France, dans le cadre de la Loi organique aux lois de finances – LOLF) ou internationales comme les Regards sur l'Education de l'OCDE et les documents de la Commission européenne, notamment dans le cadre de la Stratégie de

Lisbonne (European Commission, 2006). Les chercheurs doivent aussi réinterroger les rapports qu'entretiennent l'Ecole et la Société dans laquelle elle est implantée, en s'appuyant autant que faire ce peut sur des données robustes, notamment celles produites par les enquêtes internationales. Ils doivent ainsi examiner, de manière critique, certaines prises de positions officielles, en interrogeant les données empiriques mobilisées pour justifier des prises de positions politiques réputées élaborées sur la base de preuves (*evidence-based policy*).

Ces constats soulignent la grande richesse et la complexité de l'éducation comparée ainsi que l'intérêt de confronter les visées prescriptives/normatives de certaines prises de position officielles aux résultats des recherches qu'elles mobilisent ou invoquent. Ces approches impliquent également de recontextualiser les informations collectées dans une perspective internationale pour leur rendre ensuite une dimension nationale, appréhendable par les acteurs à ce niveau (Cytermann & Demeuse, 2005), ou pour en considérer une dimension particulière, bénéficiant non seulement de données comparatives, mais aussi d'autres sources, comme dans l'analyse, par Merle (2012), des mécanismes de ségrégation scolaire en France. Cela oblige naturellement à mettre en rapport de nombreuses informations, portant sur des niveaux différents et impliquant des méthodes différentes pour éclairer un même objet (Lafontaine et Blondin, 2004).

Pour y parvenir, les chercheurs en éducations doivent disposer de qualifications très variées, ce qui les place devant une certitude : il n'est plus possible, dans ce domaine comme dans d'autres, de mener seul une telle entreprise. Il faut en effet, à la fois, avoir une bonne maîtrise des domaines évalués, du fonctionnement des systèmes éducatifs, des techniques de mesure et d'échantillonnage, des méthodes d'analyse des données… mais aussi développer des compétences permettant de travailler en équipe, de maîtriser des outils informatiques, de lire dans d'autres langues et dans d'autres domaines que sa propre spécialité.

Pour Duru-Bellat (2007, p. 12), « une approche quantifiée peut également préserver de prises de position plus idéologiques que scientifiques, qui sont encore légion, dès qu'on traite d'éducation ». Pour cela, une bonne maîtrise des différents niveaux de production des statistiques et des indicateurs est indispensable et l'on peut sans doute, comme elle, souscrire à l'idée que « ces difficultés ne devraient

pas dissuader de recourir aux statistiques, au contraire », même si « l'essentiel reste, dans le champ de l'éducation comme dans d'autres, la qualité des interrogations » et des interprétations qui résultent de ces analyses. Mais, il ne suffit pas d'acquérir de meilleures données, ni même une meilleure compréhension des mécanismes à l'œuvre pour parvenir à améliorer la situation d'un système éducatif. Il serait naïf de croire qu'on peut « gouverner l'éducation par les nombres », pour reprendre le titre de l'ouvrage coordonné par Felouzis et Hanhart (2011). Cependant, comme le précise ces deux auteurs, il serait tout aussi contre-productif d'adopter « une attitude de rejet de toute évaluation sous prétexte que les résultats de l'éducation échappent en grande partie au mesurable et que seuls les professionnels sont légitimes pour juger d'autres professionnels 'en interne' » (p. 23).

Face aux problèmes qui ne manquent pas de se présenter dans de telles entreprises, s'ils ont été détectés, plusieurs attitudes sont possibles. La première, celle du « puriste », consisterait à n'entreprendre aucune analyse ou comparaison. Ce serait sans doute la seule position aisée à défendre dans le milieu académique. Malheureusement, cette attitude conduirait à peu de chose près, soit à renoncer à toute analyse quantifiée, soit à recourir exclusivement à une approche « littéraire ». Dans ce dilemme, comme Duru-Bellat le précise, « renoncer à des analyses solides sur le plan statistique pour se tourner vers des approches qualitatives pas moins construites et pas moins entachées d'artéfacts amènerait à délaisser certaines interrogations fondatrices de la sociologie de l'éducation, en se résignant à des monographies distrayantes ou à des expertises de moindre portée » (Duru-Bellat, 2007, p. 13).

Il est donc nécessaire d'examiner des solutions alternatives qui permettent, à défaut de certitudes, d'étudier des tendances et des hypothèses fortes permettant d'instruire des questions souvent posées de manière très générale, et parfois même sous la forme d'affirmations péremptoires. L'approche comparative, bien menée, permet également, pour une part, de dépasser la controverse entre les tenants et les adversaires d'une politique donnée en les amenant sur le terrain des faits, et non plus seulement sur celui des idées.

Mons (2007, p. 423), reprenant à son compte la conclusion d'Epstein (1988) qui réfute le lien de nécessité entre les positions épistémologiques (particularisme vs positivisme) et les méthodes

utilisées (quantitatif vs qualitatif), réaffirme la complémentarité des méthodes aux différents stades de chaque recherche. Elle souligne également les apports importants des comparaisons internationales, malgré leurs conclusions encore trop limitées, et précise utilement les améliorations nécessaires dans ce type de démarches comparatives, au-delà de leur composante quantitative et de la maîtrise des outils et des concepts statistiques.

Pour elle, le niveau d'analyse ne doit pas être exclusivement le niveau de l'Etat-nation, celui-ci ne correspondant par exemple pas au niveau de décision en matière d'éducation dans la plupart des pays connaissant une structure fédérale (Allemagne, Belgique, Suisse, Etats-Unis d'Amérique…). Par ailleurs, la complexité des politiques éducatives doit davantage être capturée par des indicateurs *ad hoc* de nature typologique que par une analyse unidimensionnelle, en veillant à la création de concepts communs au-delà des catégorisations nationales (ce qui permet d'éviter sans doute une part de « naturalisation » qui s'imposerait par l'évidence de pratiques trop exclusivement contextualisées). Pour Mons (2007), la nécessité de développer une approche néo-institutionnelle ne doit pas se limiter à une analyse formelle et exclusivement descriptive. Les performances des systèmes scolaires doivent être appréciées sous différents angles et pas seulement sous le seul angle de l'efficacité. L'équité, mais aussi la diversité des performances en regard du territoire ou l'importance numérique de certains groupes (par exemple, les plus forts ou, au contraire, les plus faibles), doivent pouvoir être pris en compte, aussi bien dans le domaine cognitif que non cognitif. Les résultats doivent être interprétés dans une perspective relativiste, privilégiant une forme de « généralisation contextualisée » à la recherche de règles immuables et indépendantes des conditions nationales et locales, sans pour autant verser dans l'incomparabilité et le particularisme absolu qui dénieraient à l'approche comparative tout intérêt en dehors de l'établissement de « monographies exotiques ». Le recours à des données factuelles est donc essentiel. On ne peut pas se limiter à des « avis d'experts indigènes » qui seraient juxtaposés.

C'est ce que vont tenter de mettre évidence les articles de ce dossier, en insistant chacun sur des aspects différents, mais complémentaires, même s'ils portent sur des sujets en apparence parfois éloignés. Ce dossier d'Education comparée compte cinq articles. Dans le premier article, Gérard Lassibille (CNRS et Université de Bourgogne), met en

évidence les problèmes posés par la mise en œuvre d'évaluations internationales dans les pays en développement. S'il existe également des évaluations internationales dans ces pays, notamment à l'initiative de l'Organisation des Nations Unies pour l'Education, la Science et la Culture (UNESCO) ou de la Conférence des Ministres de l'Education des pays ayant le français en partage (CONFEMEN), celles-ci font bien moins que le PISA de l'OCDE ou les travaux de l'Association internationale pour l'évaluation du rendement scolaire (IEA), l'objet de publications. Cet article, à travers un cas particulier, celui du Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC), dans le contexte malgache, pose la question de la fiabilité de ce type d'évaluations externes dans les pays en développement.

Le second article, proposé par Marie Duru-Bellat (SciencePo Paris), interroge ce qui doit être comparé lors d'entreprises de comparaisons internationales. Au-delà des inégalités internes au système éducatif (inégalités scolaires), l'auteur aborde le problème de la reproduction sociale que peuvent renforcer ou limiter certains systèmes éducatifs. Mettant en évidence à la fois des systèmes éducatifs peu inégalitaire dans des sociétés favorisant la mobilité sociale et d'autres systèmes éducatifs beaucoup moins égalitaires, implantés dans des sociétés relativement inégalitaires, ce qui semble assez logique, Duru-Bellat montre aussi des situations plus inattendues, l'Ecole étant tantôt beaucoup inégalitaire que la Société dans laquelle elle est implantée ou, au contraire, plutôt beaucoup plus équitable que l'analyse du fonctionnement social l'aurait laissé prévoir.

Au-delà du choc Pisa (Pons, 2012), le troisième article, rédigé par Dominique Lafontaine et Ariane Baye (Université de Liège), pose la question des effets des évaluations internationales sur les systèmes éducatifs. Pour ces entreprises à large échelle, s'agit-il principalement de mesurer les changements, d'en être les témoins, ou au contraire, de les provoquer, d'en être donc les instruments ? Les relations entre la production de résultats et les changements qui pourraient être observés dans le système éducatif ne sont pas simples. Pour commencer parce que « la réception d'une enquête internationale dans un pays ne se réduit pas à une simple transmission verticale de messages du niveau international vers le niveau national » (Pons, 2012, p. 35). Mais aussi parce que, pour une part, comme le précise Pons (p. 38), « si Pisa met la pression sur les systèmes éducatifs, c'est bien parce que les acteurs qui les composent, par leurs interactions et

interdépendances réciproques, concourent collectivement à lui conférer cette pression ».

Anne West (London School of Economics and Political Science) et Annamari Ylönen (University of Exeter), pour leur part, choisissent de comparer deux systèmes éducatifs assez contrastés, le système anglais et le système finlandais, observé sous toutes les coutures et maintes fois pris comme référence depuis les premiers résultats de PISA en 2000, y compris par des enseignants curieux qui finiront pas lui consacrer un ouvrage (Robert, 2008). Cette comparaison ne repose pas, comme d'autres, sur les seuls résultats des enquêtes quantitatives, mais elle vient utilement éclairer, en les mettant en regard, deux systèmes éducatifs qu'a priori tout oppose… en interrogeant les véritables différences qui peuvent effectivement être observées, ou pas.

Marc Demeuse, Natacha Duroisin et Sabine Soetewey (Université de Mons), présentent un texte d'une nature un peu différente par rapport aux précédents. Dans cet article, les auteurs ne s'intéressent en effet pas directement aux évaluations internationales, mais à l'analyse des programmes de l'enseignement des sciences en Belgique francophone. Cependant, en réinterrogeant les référentiels communs qui servent de base à l'élaboration de ces programmes dans un système où chaque réseau d'enseignement élabore ses propres programmes, ils mettent en évidence une très grande similitude entre les méthodes qui sont mobilisées dans le cadre des évaluation internationales et celles qui devraient être mise à l'œuvre pour procéder à des évaluations à l'intérieur même d'un système très peu régulé en dehors des « règles » d'un quasi-marché, conduisant même à rendre impossibles des évaluations « nationales » en science au grade 10 (la quatrième année de l'enseignement secondaire en Belgique francophone).

Références bibliographiques

CHAMPOLLION, P., & BARTHES, A. (2012). De l'usage et du mésusage du classement par rang en matière de médiatisation de l'évaluation internationale PISA. *Questions Vives, vol. 6* n°16, 2-11.

CYTERMANN, J.R., & DEMEUSE, M. (2005). *La lecture des indicateurs internationaux en France.* Paris: Haut Conseil de l'Evaluation de l'Ecole (n° 18, mai 2005).

DEMEUSE, M. (2012). *Comparer les systèmes éducatifs européens. Une entreprise complexe mais indispensable*. Diversité (Ville Ecole Intégration), 169 (juillet 2012), 39-44.

DEMEUSE, M., BAYE, A. (2007). La Commission européenne face à l'efficacité et l'équité des systèmes éducatifs européens. *Éducation et sociétés, vol. 2*, n° 20, 105–119.

DUPRIEZ, V., MONS, N. (2011). Les politiques d'accountability. Du changement institutionnel aux transformations locales. *Education comparée, 5*, 7-16.

DURU-BELLAT, M. (2007). Progrès statistiques et glissements conceptuels dans l'analyse des inégalités sociales à l'école. *Revue française de Pédagogie, 161,* 5-14.

EPSTEIN, E. (1988). The problématic meaning of 'comparison' in comparative education'. In J. Schriewer & B. Holmes (eds), *Theories and Methods in Comparative Education.* Frankfurt : Verlag Peter Lang.

EUROPEAN COMMISSION (2006). *Progress Towards the Lisbon Objectives in Education and Training. Commission staff working paper. 2006 Report*. Brussels : Commission of the European Communities, Education and Culture.

FELOUZIS, G., & CHARMILLOT, S. (2012). *Les enquêtes PISA.* Paris : Presses universitaires de France, coll. « Que sais-je ? ».

FELOUZIS, G., & HANHART, S. (éds.) (2011). *Gouverner l'éducation par les nombres? Usages, débats et controverses*. Bruxelles: De Boeck, coll. "Raisons éducatives".

LAFONTAINE, D., & BLONDIN, C. (2004). *Regards sur les acquis des élèves en Communauté française Apports des enquêtes de l'I.E.A, de PISA et des évaluations externes.* Bruxelles : De Boeck.

LOYE, N. (2011). Panorama des programmes actuels d'enquêtes à large échelle. *Mesure et évaluation en éducation, 34*(2), 3-24.

MERLE, P. (2012). *La ségrégation scolaire.* Paris : Editions La Découverte, coll. « Repères ».

MONS, N. (2007). L'évaluation des politiques éducatives. Apports, limites et nécessaire renouvellement des enquêtes internationales sur les acquis des élèves. *Revue Internationale de Politique comparée, 14(*3), 409-423.

OLSEN, R.V., & LIE, S. (2006). Les évaluations internationales et la recherche en éducation : principaux objectifs et perspectives. *Revue française de pédagogie, 157*, 11-26.

OZGA, J. (2012). Assessing PISA. *European Educational Research Journal, Volume 11* Number 2, 166-171.

POUPEAU F. (2003). *Une sociologie d'Etat. L'école et ses experts en France.* Paris : Edition Raisons d'agir.

PONS, X. (2012). Qu'est-ce qu'un effet Pisa ? Réflexions sur la politique française d'évaluation en éducation. *Diversité (Ville intégration éducation), 169,* 34-38.

ROBERT, P. (2008). La Finlande : un modèle éducatif pour la France ? Les secrets de la réussite. Paris : ESF éditeur, coll. « Pédagogies ».