De l'usage et du mésusage du classement par rang en matière de médiatisation de l'évaluation internationale PISA

How the PISA rankings by country are used and... misused

Pierre Champollion et Angela Barthes

Édition électronique

URL : <http://journals.openedition.org/questionsvives/895>

DOI : 10.4000/questionsvives.895

ISBN : 978-2-8218-1392-2

ISSN : 1775-433X

Éditeur

Université Aix-Marseille (AMU)

Édition imprimée

Pagination : 15-27

ISSN : 1635-4079

Référence électronique

Pierre Champollion et Angela Barthes, « De l'usage et du mésusage du classement par rang en matière de médiatisation de l'évaluation internationale PISA », *Questions Vives* [En ligne], Vol.6 n° 16 | 2012, mis en ligne le 15 décembre 2011, consulté le 19 avril 2019. URL : <http://journals.openedition.org/questionsvives/895> ; DOI : 10.4000/questionsvives.895

Ce document a été généré automatiquement le 19 avril 2019.

Questions Vives est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

De l'usage et du mésusage du classement par rang en matière de médiatisation de l'évaluation internationale PISA

How the PISA rankings by country are used and... misused

Pierre Champollion et Angela Barthes

Introduction

- ¹ La presse spécialisée, voire les organisations représentatives enseignantes, sans compter les décideurs politiques, n'ont en France bien souvent voulu voir, dans les évaluations internationales successives PISA 2000, 2003, 2006 et 2009, que les éventuels biais et limites d'une enquête initialement fondée sur des pratiques enseignantes anglo-saxonnes¹. C'est bien ce qui ressort de la médiatisation initiale des résultats de PISA. Mais, au-delà de ce premier comportement « de défense », les principaux médias, tant généralistes que professionnels, n'ont surtout retenu de ces évaluations successives triennales qu'un classement par rang entre les différents pays enquêtés, dont le nombre – qui a plus que doublé entre 2000 et 2009 – s'est accru au fil des études, s'étendant aujourd'hui bien plus loin que le cercle initial restreint des pays « riches » membres de l'OCDE. L'essentiel de l'intérêt de PISA – indéniable par-delà les différents biais répertoriés – ne s'inscrit cependant pas dans ces premières réactions fréquemment médiatisées. D'autres angles d'analyse, insuffisamment évoqués malgré leur « évidence », semblent davantage susceptibles de mettre en lumière des résultats des enquêtes.
- ² L'hypothèse initiale que nous posons dans cet article est qu'un « hit-parade » des performances des différents systèmes scolaires évalués au travers d'un classement par rang des pays enquêtés ne constitue pas le meilleur moyen de rendre compte, y compris en terme de médiatisation, à des fins de vulgarisation comme de diffusion des résultats scientifiques, d'une enquête complexe comme PISA. En effet, les scores bruts –

compétence par compétence – des différents pays enquêtés, loin de séparer les performances des différents pays, ont plutôt tendance à les rassembler dans une même grande fourchette, au-delà de quelques cas s'en écartant significativement (Baudelot & Estabiet, 2009). À quoi s'ajoute que les différents classements par rang, qui n'ont pas grande signification statistique, sont de plus en plus mis à contribution par les responsables de l'éducation, en France notamment, pour tenter de légitimer, en s'appuyant sur les supposés mauvais résultats de PISA très médiatisés, des réformes éducatives locales.

- 3 Cette première approche effectuée, nous esquissons ici l'amorce d'une analyse des résultats des enquêtes visant à la recherche de facteurs explicatifs et comparatifs entre les pays. Les performances scolaires des élèves des différents pays enquêtés ont été rapportées à quelques caractéristiques des systèmes éducatifs évalués en croisant les résultats de PISA avec des indicateurs scolaires ou institutionnels comme la pratique généralisée du redoublement. Plus largement, les résultats PISA des pays évalués ont été situés d'abord par rapport à leurs niveaux de développement économique et social, puis par rapport à leurs types d'organisation scolaire respectifs, dont dépendent potentiellement une partie des résultats, dans le but de mieux apprécier la plus-value éventuelle des différents systèmes éducatifs nationaux. C'est ainsi, rappelons-le pour mémoire, que le ministère de l'Éducation nationale (MEN) lui-même procède, en France, pour – un peu – « sortir » les établissements scolaires (lycées) de leur classement journalistique médiatisé par la seule réussite brute au baccalauréat ou bien lorsqu'il publie des « indicateurs de pilotage des établissements secondaires » (IPES) qui rapportent le niveau de performance scolaire d'un établissement secondaire à ses caractéristiques contextuelles propres, de composition sociale notamment.

Fragilité de la médiatisation des classements par rang et approche critique de son utilisation politique

Fragilité de la vulgarisation médiatique des classements par rang

- 4 Dans l'éducation, on le sait, le classement par rang, à partir de notes chiffrées, relève quasiment du traditionnel ou de l'institutionnel historiques. C'est lui qui depuis toujours institue l'élève dans un statut flatteur ou stigmatisant en France. C'est lui qui, aujourd'hui, fonde par exemple la distinction des lycées sur la réussite au baccalauréat, distinction qui s'appuyait antérieurement souvent sur la seule rumeur. Il n'est donc pas surprenant de voir les évaluations successives PISA faire l'objet, dans la presse, y compris la presse éducative spécialisée, de classements permettant de facto d'encenser les premiers et de stigmatiser les derniers, même s'il ne s'agit pas de l'aspect le plus intéressant de PISA (Sauvageot & Dalsheimer, 2008). La communication officielle de l'État, du MEN en l'occurrence, n'est pas absente – voir là une litote ! – de ce type de présentation, comme on le verra plus précisément par la suite. Il n'est pas jusqu'à des ouvrages plus « sérieux » pour sous-titrer un chapitre « la France mal classée », juste avant d'affirmer que l'aspect le plus médiatisé de PISA, son côté hit-parade, est celui qui est le moins porteur de sens (Baudelot & Estabiet, 2009, pp. 15-16) !
- 5 Effectivement, déjà, les marges d'erreur de l'enquête susceptibles d'affecter les scores de PISA invalident, de l'aveu même des producteurs des évaluations effectuées dans ce cadre (Baudelot & Estabiet, 2009), les différenciations fondées sur des écarts inférieurs à cinq

points de part et d'autre d'un même score qui peut donc regrouper plus d'une dizaine de pays sans différence statistiquement significative entre eux. À titre d'exemples, pour les distributions 2009, cette marge de non-différenciation statistique à .05 autour des scores de la France correspond à 10 scores ou pays en culture mathématique et en culture scientifique et à 15 scores ou pays en compréhension de la langue écrite. Au-delà, si l'on encadre le score moyen des distributions des différentes enquêtes PISA effectuées depuis 2000² par un intervalle de confiance testant la significativité des scores moyens à .05³, on s'aperçoit facilement que cet encadrement rassemble tendanciellement, à chacune des quatre évaluations, une vingtaine de pays (tableau n° 1). Dans la mesure où les seules différences statistiquement significatives sont celles qui sortent de la fourchette, il est trivialement évident qu'un classement individualisé par rang n'a pas grande valeur scientifique. Et il en va de même lorsqu'on encadre les performances françaises par un intervalle du même type. La fourchette ne regroupe dans ce dernier cas pas moins de 38 pays, dont l'Allemagne, le Japon, et même la – si performante – Finlande, etc. Sa borne supérieure – 538 – va presque jusqu'à talonner la Corée du Sud dont le score s'élève à 539 !

Tableau n° 1 : encadrement à .05 des performances moyenne et française à la compréhension de la langue maternelle écrite des enquêtes PISA successives 2000, 2003, 2006 et 2009 ($m \pm z \cdot \sigma / \sqrt{N}$). Distributions respectives de 31 pays (2000), 43 pays (2003), 60 pays (2006) et 65 pays (2009)⁴.

	Enquête 2000	Enquête 2003	Enquête 2006	Enquête 2009
Intervalle de confiance encadrant la moyenne de la distribution à .05	434 < 471 < 508	423 < 459 < 495	372 < 420 < 468	393 < 435 < 477
Nombre de scores (pays) inclus dans cette fourchette	20 scores	18 scores	16 scores	24 scores
Intervalle de confiance encadrant la performance de la France à .05	468 < 505 < 542	460 < 496 < 532	439 < 487 < 525	454 < 496 < 538
Nombre de scores (pays) inclus dans cette fourchette	24 scores	31 scores	35 scores	38 scores

Source : OCDE. N.B. Moyennes des distributions et performances françaises mises en gras.

- 6 De la même façon, si au lieu de regarder les performances à la compréhension de l'écrit sur les quatre évaluations successives, on s'attardait sur la seule évaluation de 2009, en prenant en compte les trois types d'évaluation, on verrait facilement (tableau n° 2) que les intervalles de confiance calculés englobent à chaque fois au moins une vingtaine de pays, rendant de ce fait même tout classement par rang inopérant.

Tableau n° 2 : encadrement à .05 des performances moyenne et française en culture scientifique, culture mathématique et compréhension de la langue maternelle écrite de PISA 2009 ($m \pm z \cdot \sigma / \sqrt{N}$). Distribution comprenant 65 pays.

	Sciences	Mathématiques	Compréhension de l'écrit
Intervalle de confiance encadrant la moyenne de la distribution à .05	410 < 452 < 494	419 < 465 < 511	393 < 435 < 477
Nombre de scores (pays) inclus dans cette fourchette	22 scores	33 scores	24 scores
Intervalle de confiance encadrant la performance de la France à .05	456 < 498 < 540	451 < 497 < 543	454 < 496 < 538
Nombre de scores (pays) inclus dans cette fourchette	36 scores	36 scores	41 scores

Source : OCDE. N.B. Moyennes des distributions et performances françaises mises en gras.

- 7 En fait, la vulgarisation – assise principalement sur des classements par rang fragiles scientifiquement – à laquelle nous avons assisté depuis quelques années pourrait bien avoir d'autres objectifs que la seule médiatisation journalistique de résultats scientifiques réputés trop difficiles à présenter tels quels. Une certaine légitimation des mesures et des réformes entreprises et/ou projetées, via des résultats récurrents médiocres, s'aggravant même au fil des évaluations successives, pourrait bien être visée à travers de telles pratiques...

Médiatisation et utilisation politique des résultats de PISA

- 8 En effet, après avoir, au tout début de l'opération, largement ignoré la première évaluation PISA (2000), puis avoir mis l'accent sur ses imperfections et biais, voire dénigré à l'occasion de la sortie des deuxièmes résultats de l'évaluation (2003), la presse – aussi bien spécialisée que généraliste – a médiatisé de façon souvent outrancière des différents classements par rang des systèmes éducatifs évalués. En atteste, aujourd'hui encore, la récente parution, le 15 décembre 2010, dans *Le Monde Éducation*, d'un dossier complet « classant les pays de l'OCDE en fonction de leurs résultats au test PISA 2009 » intitulé « Les meilleures écoles du monde » ! Après quoi, devant la récurrence des résultats peu flatteurs pour la France, PISA a été de plus en plus mis à contribution par les responsables de l'éducation en France pour légitimer les mesures préconisées et les réformes entreprises.
- 9 Ce changement d'attitude, qui se produit majoritairement à partir de l'année 2005-2006, se poursuit les années suivantes. C'est à cette époque que la Direction de l'évaluation et

de la prospective (DEP), fondée en 1986 puis réorganisée depuis en DEPP (le dernier « P » ajouté signifie « performance »), pour mieux prendre en compte les recommandations de la « Loi organique relative aux lois de finances » (LOLF) qui, à partir du 1^{er} janvier 2006, a imposé de piloter le système éducatif français par des objectifs quantifiés (Pons, 2010). En témoigne ainsi, par exemple, la référence à PISA inscrite, pour la première fois, dans le décret du 11 juillet 2006 relatif au « socle commun de connaissances et de compétences » (SCCC) pris en application de la loi d'orientation et de programme du 23 avril 2005, cadre législatif actuel du fonctionnement et de l'organisation du système éducatif :

La définition du socle commun de connaissances et de compétences se réfère enfin aux évaluations internationales, notamment au Programme international pour le suivi des acquis des élèves (PISA), qui propose une mesure comparée des connaissances et des compétences nécessaires tout au long de la vie (p. 3).

- 10 Les toutes dernières références institutionnelles aux évaluations PISA des nombreuses communications officielles développées depuis le décret de juillet 2006 sur le SCCC en attestent également. Nous en citons sommairement ici deux, correspondant respectivement à un « avis » d'un organisme officiel de la République et au « rapport » livré à la suite d'une commande institutionnelle, à titre d'illustration de notre propos :

- D'une part, l'avis du Conseil économique, social et environnemental (CESE) sur « les inégalités à l'école » publié par les éditions du Journal Officiel le 20 septembre 2011 : « Au regard des indicateurs d'équité de l'enquête PISA 2009, le système scolaire français est plutôt mal positionné » (p. 11).

Au-delà de la référence générale précédente, bien d'autres renvois tant implicites qu'explicites à PISA, dans le but de fonder techniquement les « préconisations » émises, émaillent ce rapport du CESE sur la scolarité française. L'exposé des motifs initial, notamment, en est littéralement truffé.

- D'autre part, le rapport sur la « maîtrise de la formation initiale des enseignants, enjeux et bilan », établi à la demande du ministère de l'enseignement supérieur et de la recherche (MESR) par le président du comité de suivi « Master » et remis à son commanditaire le 10 octobre 2011 : « L'expérience internationale [PISA] nous montre clairement que toutes les nations sont elles aussi impliquées dans des processus récurrents de réforme de cette formation [...]. Réformer n'est donc en ce sens pas critiquable » (p. 5).

- 11 Ces deux derniers exemples précédents vont même plus loin que la référence initiale à PISA du SCCC : ils mettent clairement en évidence la fonction « légitimatrice » de l'appel aux évaluations internationales, dont PISA est aujourd'hui le vecteur majeur. Pourtant, rappelons-le, PISA n'a bien sûr jamais eu – y compris en creux, *id est* négativement ! – pour vocation essentielle de légitimer les réformes menées par les échelons politiques nationaux en charge de l'éducation, les mesures administratives prises par les échelons déconcentrés de l'État et les avis et propositions des institutions officielles.

- 12 Comme on le voit, on est donc passé en France, vis-à-vis des évaluations PISA, de l'ignorance pure et simple du début au dénigrement quasi-systématique par l'ensemble des acteurs de l'éducation, puis à la reconnaissance progressive de validité, pour aboutir enfin aujourd'hui à une sorte d'instrumentation politique de ses résultats en vue de légitimer les réformes entreprises et les mesures prises. Mais la France ne détient pas le monopole de ce type de réaction. Bien d'autres pays se sont appuyés sur PISA pour tenter de légitimer leurs réformes. Ce fut, par exemple, le cas du Luxembourg qui, effrayé par ses résultats 2003 dans le domaine de la compréhension de la langue écrite, décida

d'apprendre les langues française et allemande, non plus comme des langues maternelles, mais comme des langues étrangères⁵, ainsi que de l'Allemagne qui, après 2006, entreprit avec succès de rénover son système éducatif et vit, dès 2009, ses résultats améliorés⁶.

Analyse des performances des pays rapportés à quelques paramètres contextuels et organisationnels

- 13 Comme nous l'avons explicité dans la première partie, les différents classements par rang auxquels ont donné lieu les évaluations successives PISA ne constituent pas les approches les plus pertinentes pour rendre compte de l'intérêt scientifique de la démarche d'évaluation mise en œuvre. Il nous paraît donc nécessaire dans cette perspective de déplacer quelque peu la posture d'analyse en rapportant par exemple, les performances des différents pays à des éléments contextuels ou bien encore en croisant les scores bruts des pays dans les trois domaines d'évaluation avec des paramètres organisationnels. Au-delà du profil de connaissances et de compétences à la fin de la scolarité obligatoire⁷, nous tenterons, comme avant nous Olsen et Svein (2006), d'établir des liens entre les caractéristiques des écoles, notamment, et les résultats scolaires, ainsi que plus largement de rapporter les performances des pays à des indicateurs contextuels. Les éléments de contexte, en effet, sont susceptibles de rendre compte de l'enquête et d'amorcer des tentatives d'explications et de comparaison des facteurs de performance. Nous choisirons pour ce faire plusieurs paramètres contextuels dont nous faisons a priori l'hypothèse qu'ils sont corrélés avec les performances obtenues par les pays : en premier lieu, le niveau plus ou moins élevé du PIB par habitant, comme par exemple la pratique systématique ou non du redoublement et l'organisation plus ou moins centralisée des systèmes éducatifs. Nous utiliserons dans ce cadre comparatif les toutes dernières évaluations PISA, datant de 2009, qui regroupent le plus grand nombre de pays (65).
- 14 D'un point de vue méthodologique, notre démarche consistera donc à sélectionner des groupes de pays comparables en fonction des facteurs organisationnels susceptibles d'expliquer les performances constatées. Les pays séparés en sous-groupes feront après quoi l'objet de tris croisés classiques avec les performances ou scores PISA. Les résultats comparés pourront alors donner lieu à des tentatives d'explication. La méthode suivie s'articule ainsi en deux temps : la première phase consiste à dégager les facteurs contextuels majeurs influençant les résultats scolaires – comme par exemple les niveaux économiques des pays – pour être ensuite à même d'interroger les facteurs organisationnels de l'éducation. À titre d'exemple, deux facteurs organisationnels ont été ainsi extraits, mais de nombreux autres peuvent être bien sûr convoqués.

Le PIB par habitant, un facteur déterminant des résultats PISA ?

- 15 Dans la mesure où l'analyse a pour objectif d'amorcer des tentatives d'explications entre les facteurs de performance et les facteurs organisationnels de l'éducation, il nous a paru dans un premiers temps utile de « neutraliser » les aspects contextuels trop fortement déterminants comme les niveaux de développement économiques des pays. Nous avons donc effectué dans ce but des corrélations entre les scores bruts des pays (moyenne des trois domaines d'évaluation) et les indicateurs retenus, puis déterminé les groupes de pays pour lesquels les facteurs contextuels majeurs ne constituent pas à qualité un élément déterminant et où, donc, d'autres facteurs organisationnels peuvent être

avancés. Ainsi, les résultats PISA, dans les trois champs d'investigation, ont été appréciés par rapport à un indicateur synthétique du niveau de développement économique, social et culturel des pays concernés. L'objectif a été ici avant tout de déterminer des classes de pays comparables. Nous avons choisi de nous appuyer pour ce faire sur le PIB annuel par habitant⁸. La figure n° 1 présentée immédiatement ci-après montre que, globalement, la corrélation entre le niveau de score PISA et le niveau du PIB annuel par habitant est faible (0,39). Mais il n'en va pas du tout de même lorsqu'on corrèle le même niveau de PIB annuel avec les scores PISA des pays ayant un PIB par habitant au-dessus de la moyenne (figure n° 2) et au-dessous de la moyenne (figure n° 3). Dans le premier cas, la corrélation devient quasi-nulle (0,08), ce qui signifie qu'au-delà d'un niveau minimal, ici la moyenne (score de 450 environ), le niveau de PIB n'intervient plus guère dans les résultats PISA. Tandis que dans le second cas (figure n° 3), la corrélation s'élève à 0,7, ce qui montre au contraire une influence tout à fait significative.

Figure n° 1 : croisement des scores PISA 2009 (moyenne des math, culture scientifique et langue maternelle) et du niveau moyen de PIB par habitant annuel en US dollars des 65 pays enquêtés

Sources : OCDE pour les scores PISA ; FMI 2010 pour le PIB

Figure n° 2 : croisement des scores PISA 2009 (moyenne des culture mathématique, culture scientifique et langue maternelle) et du niveau moyen de PIB par habitant annuel en US dollars des 65 pays enquêtés situés au-dessus de la moyenne de la distribution du PIB par habitant annuel

Sources : OCDE pour les scores PISA ; FMI 2010 pour le PIB

Figure n° 3 : croisement des scores PISA 2009 (moyenne des math, culture scientifique et langue maternelle) et du niveau moyen de PIB par habitant annuel en US dollars des pays enquêtés situés en-dessous de la moyenne de la distribution du PIB par habitant annuel

Sources : OCDE pour les scores PISA ; FMI 2010 pour le PIB

- 16 Compte-tenu des résultats des corrélations, il paraît évident que les facteurs organisationnels de l'éducation peuvent être discriminés seulement sur les pays ayant un PIB supérieurs à la moyenne des PIB concernés, le facteur économique étant trop discriminant pour les autres pays. De la même manière, des corrélations ont été

effectuées sur les IDH (indices de développement humain). Il s'avère que cet indice permet de discrétiser plus finement les pays à faible développement et, donc, d'aborder de la même façon l'analyse des facteurs organisationnels de l'éducation sur d'autres groupes de pays.

L'influence potentielle des facteurs organisationnels : deux exemples

- 17 Ce travail préalable réalisé, il devient alors possible de croiser les résultats de PISA avec un certain nombre d'indicateurs scolaires ou institutionnels pouvant déboucher sur des hypothèses intéressantes en termes d'efficacité organisationnelle. En croisant ainsi la pratique systématique du redoublement en Europe et les résultats PISA 2009 (tableau n° 3), nous pouvons voir que seul le champ de la culture mathématique donne un résultat significatif, aussi bien à .05 qu'à .01, alors qu'il n'en va pas de même des autres champs d'investigation, dont les croisements respectifs avec la pratique du redoublement n'apparaissent pas significatifs. Sans doute peut-on penser que les mathématiques, qui servent aujourd'hui – après le latin avant-hier et l'allemand hier – de support souvent exclusif aux passages dans les classes supérieures, voient leurs évaluations peser plus fortement dans les pays qui pratiquent systématiquement le redoublement et, par conséquent, les efforts des élèves et des enseignants se densifier en faveur de cette discipline. C'est, en tous cas, vraisemblablement ce qui se passe en France.

Tableau n° 3 : tri croisé « pratique systématique du redoublement en Europe / résultats PISA 2009 » en Europe

	Sciences	Mathématiques	Compréhension de l'écrit
Pays avec pratique systématique du redoublement	513	521	521
Pays sans pratique systématique du redoublement	492	425	504

- 18 Lorsqu'on croise « centralisation de l'organisation scolaire » (par exemple degré d'autonomie des établissements scolaires) et « résultats PISA 2009 » en Europe, l'on n'observe pas, en revanche, de différence significative dans les résultats PISA en fonction de l'aspect plus ou moins centralisé du système éducatif (tableau n° 4).

Tableau n° 4 : tri croisé « organisation scolaire centralisée en Europe / résultats PISA 2009 » en Europe

	Sciences	Mathématiques	Compréhension de l'écrit

Organisation centralisée	scolaire	plutôt	509	483	491
Organisation décentralisée	scolaire	plutôt	485	497	495

- 19 Faudrait-il formuler l'hypothèse que le degré de centralisation ou de décentralisation n'aurait pas, en tant que tel, d'impact sur les performances scolaires ? Il serait sûrement utile dans cette perspective de conduire des études complémentaires à l'instar par exemple de ce qui a déjà été réalisé (Prenzel & Zimmer, 2006) en Allemagne où la constitution décentralisée rend nécessaire la différenciation des résultats PISA *Land* par *Land*⁹. Les différents éléments de contexte locaux restant bien sûr à corrélérer ultérieurement avec la distinction développée.

Conclusion

- 20 Ressortent potentiellement, on l'a vu, des analyses rapides que nous avons présentées dans cet article, d'autres perspectives qui vont bien au-delà de la seule photographie figée qui a bien souvent fait la une des médias et corrélativement suscité l'ire des professionnels de l'éducation. L'utilisation d'indicateurs contextuels économiques, sociaux, territoriaux et culturels pour apprécier les performances d'un pays paraît indispensable pour sortir des seuls classements par rang aux résultats scientifiques fragiles qui résonnent encore fortement dans la sphère médiatique éducative. De même, l'affinement des résultats de PISA par des analyses comparatives reposant sur des facteurs organisationnels liés à l'éducatif va évidemment dans le même sens.
- 21 Mais, au-delà de la vulgarisation médiatique encore dominante, plusieurs éléments entravent encore aujourd'hui les recherches de ce type en sciences de l'éducation. L'impact des différents facteurs contextuels (sociaux, institutionnels, territoriaux, notamment) reste globalement, d'abord, encore peu développé (Rhein, 2003 ; Arrighi, 2004 ; Ben Ayed, 2007 ; Lahire, 2012). Un certain nombre de chercheurs en sciences de l'éducation, ensuite, répugnent à utiliser des corpus qu'ils n'ont pas constitués eux-mêmes (Olsen & Svein, 2006). La pression médiatique environnante, enfin, réduit considérablement les marges de manœuvre socialement et financièrement disponibles pour aller vers de telles analyses.
- 22 Nous voudrions, enfin, ne pas passer sous silence tout l'intérêt des approches comparatives au sein des évaluations internationales, comme celles qui ont été récemment conduites (Rémond, 2006) entre PISA et PIRLS (Programme international de recherche en lecture scolaire), qui ne sont pas encore très répandues...

BIBLIOGRAPHIE

- Arrighi, J.-J. (2004). Les Jeunes dans l'espace rural : une entrée précoce sur le marché du travail ou une migration probable. *Formation-Emploi*, n° 87.
- Ben Ayed, C. (2007). Entre universalisme et développement local : l'émergence d'une nouvelle école ? *Revue internationale d'éducation de Sèvres*, n° 46, décembre 2007.
- Baudelot, C., & Establet, R. (2009). *L'élitisme républicain. L'école française à l'épreuve des comparaisons internationales*. Collection La République des idées. Paris : Seuil.
- Lahire, B. (2012). *Monde pluriel. Penser l'unité des sciences sociales*. Paris : Seuil.
- MEN-DEP (2001). Les élèves de 15 ans – Premiers résultats d'une évaluation internationale des acquis des élèves (PISA). *Note d'information* 01.52, MEN-DEP, décembre 2001.
- MEN-DEP (2004). Les élèves de 15 ans – Premiers résultats de l'évaluation internationale PISA 2003. *Note d'information* 04.12, MEN-DEP, décembre 2004.
- MEN-DEP (2008). L'évolution des acquis des élèves de 15 ans en culture mathématique et en compréhension de l'écrit – Premiers résultats de l'évaluation internationale PISA 2006. *Note d'information* 08.08, MEN-DEPP, janvier 2008.
- MEN-DEP (2008). Comparaisons internationales. *Éducation & Formations*, n° 78, MEN-DEPP, novembre 2008.
- MEN-DEP (2010). L'évolution des acquis des élèves de 15 ans en culture mathématique et en culture scientifique – Premiers résultats de l'évaluation internationale PISA 2009. *Note d'information* 10.23, MEN-DEPP, décembre 2010.
- MEN-DEP (2010) L'évolution des acquis des élèves de 15 ans en compréhension de l'écrit – Premiers résultats de l'évaluation internationale PISA 2009. *Note d'information* 10.24, MEN-DEPP, décembre 2010.
- Olsen, R.-V., & Svein, L. (2006). Les évaluations internationales et la recherche en éducation : principaux objectifs et perspectives. *Revue française de pédagogie*, 157, octobre-novembre-décembre 2006.
- Pons, X. (2010). L'enquête PISA sert plutôt à légitimer des réformes ou des positions préétablies. *La Lettre de l'éducation*, 687, 13 décembre 2010.
- Prenzel, M., & Zimmer, K. (2006). Études complémentaires de PISA 2003 en Allemagne : principaux résultats et enseignements. *Revue française de pédagogie*, 157, octobre-novembre-décembre 2006.
- Rémond, M. (2006). Éclairages des évaluations internationales PIRLS et PISA sur les élèves français. *Revue française de pédagogie*, 157, octobre-novembre-décembre 2006.
- Rhein, C. (2003). L'espace, les sociologues et les géographes. Déconstruire et reconstruire les disciplines. Les jeux de l'interdisciplinarité, *Sociétés contemporaines*, n° 49-50.
- Sauvageot, C., & Dalsheimer, N. (2008). La situation éducative de la France comparée à celle d'autres pays de l'UE ou de l'OCDE. *Éducation & Formations*, n° 78, MEN-DEPP, novembre 2008.

NOTES

1. Évaluations fondées sur des items correspondant plus à des compétences qu'à des connaissances par exemple.
 2. Sur les trois types d'évaluations – culture mathématique, culture scientifique et compréhension de la langue maternelle écrite – répétées tous les trois ans depuis 2000, soit au total aujourd'hui douze distributions, nous présentons ici les seules évaluations portant sur la compréhension de la langue écrite.
 3. En « normalisant » les distributions qui s'avèrent encore plus resserrées autour de la moyenne qu'une distribution gaussienne standard.
 4. Les intervalles de confiance à .05 pour les distributions 2000 et 2003, s'ils ne diffèrent pas vraiment en tendance des intervalles correspondant aux données 2006 et 2009, ne sont cependant donnés ici que pour illustration. Leurs effectifs sont trop faibles pour les rendre véritablement significatifs.
 5. Il faut dire que le Luxembourg scolarise dès trois ans à l'école maternelle – obligatoire – ses ressortissants en luxembourgeois (langue nationale issue du vieil haut allemand au XI^e siècle comme le francique, l'alsacien et le *Schweizer Deutsch*), puis le fait à partir de six ans à l'école élémentaire en allemand et enfin après 11 ans au collège en français !
 6. Remontant – dans le classement par rang ! – jusqu'à la treizième place en culture scientifique en 2009.
 7. En fin de secondaire 1 (collège) dans tous les pays européens.
 8. Malgré les critiques que l'on peut adresser à cet indicateur central qui – n'indiquant pas la dispersion des résultats – ne prend en compte que les éléments comptables de la situation d'un pays.
 9. Les compétences éducatives sont, pour l'essentiel, exercées par les différents *Länder* allemands.
-

RÉSUMÉS

L'article montre que, si les classements des pays par rang issus des évaluations PISA n'ont qu'une faible signification statistique, leur médiatisation a souvent été mise à contribution, notamment en France, pour essayer de légitimer des mesures et des réformes éducatives. Mais par-delà ce positionnement très médiatique – et en dépit des biais bien connus de l'enquête internationale – les évaluations PISA réalisées permettent d'avancer un certain nombre de facteurs explicatifs comparatifs relatifs aux aspects organisationnels de l'éducation qui s'avèrent à la fois plus intéressants, plus précis et plus robustes que les classements par rang à partir d'une méthodologie simple à visée exploratoire.

The article shows that the classification of countries by rank has only a weak statistical meaning but its mediatization in particular in France has been used to legitimize educational reforms. However, the results of the PISA evaluation allow us to search for explanatory and comparative organizational factors in education altogether more interesting, sounder and more precise than classification by rank, since a clear methodology is formulated.

INDEX

Mots-clés : classement par rang, instrumentation politique, légitimation des réformes, performance scolaire, politique éducative

Keywords : classification by rank, political instrumentation, reform legitimization, school performances, educational policy

AUTEURS

PIERRE CHAMPOLLION

Inspecteur d'académie, Habilité à diriger des recherches, sciences de l'éducation, EAM ADEF

ANGELA BARTHES

Maître de conférences, Habilitée à diriger des recherches, sciences de l'éducation, EAM ADEF