

2 LA PROBLÉMATIQUE DE L'ÉVALUATION PAR LES STANDARDS INTERNATIONAUX



Jean-Marie De Ketele

Professeur émérite UCL

DHC Chaire UNESCO de l'UCAD

MOTS CLES: *Évaluation des acquis, évaluation des compétences, évaluations internationales*

est fondamental, car la norme ultime de tout système éducatif réside bien dans ce qu'il a permis d'apprendre aux élèves dont il a la charge. Mais la question mérite d'être élargie dans l'optique de ce que l'on appelle une « culture de l'évaluation » et qui prend en compte non seulement l'efficacité en termes de résultats, mais aussi d'autres dimensions du processus éducatif, dont deux nous paraissent importantes : la capacité du système éducatif à générer plus d'équité (ce qui relève des aspects éthiques) et sa capacité à s'améliorer (ce qui relève de son pouvoir d'action). Il s'agit donc bien ici de l'impact différentiel des enquêtes internationales selon les pays qui y participent.

L'évaluation des performances

L'accent de nombreuses publications a été porté sur le palmarès comparatif des performances, ainsi que sur l'évolution comparative des performances en termes de progrès ou de régression à travers le temps (tout particulièrement entre les résultats PISA des enquêtes menées en 2000 et 2003). Certains ont même déploré la dérive médiatique de ce qu'ils ont appelé « l'Eurovision des systèmes éducatifs ». Par contre, trop peu de publications se sont penchées sur le contenu des tests adressés aux élèves et sur la pertinence de ce qui est évalué.

Il serait faux de croire que les responsables des enquêtes internationales ont accordé peu d'importance à la question du choix de ce qui doit être évalué. Baye, Demonty, Fagnant, Lafontaine, Matoul et Monsieur (2004) mettent bien en évidence ce que les concepteurs de l'enquête PISA 2003 ont voulu évaluer. Ainsi,

I. Regard à partir des enquêtes internationales

Comme le souligne Lise Demailly dans ses travaux récents (2006a et 2006b), les offreurs d'évaluation se sont multipliés, tant au niveau national qu'international. Elle en conclut avec raison qu'il est urgent d'évaluer ces évaluations.

Mais qu'évaluent donc les enquêtes internationales ? Cette question revêt plusieurs sens complémentaires, dont deux sont l'objet de cette contribution. La première peut être formulée comme suit : quelles sont les performances des élèves réellement évaluées par les enquêtes internationales ? Et donc corollairement, quelles sont les « performances » (les guillemets soulignent le caractère restrictif du terme) des élèves qui ne sont pas prises en compte dans de telles enquêtes ? Il va de soi que ce type de questions

l'évaluation de « la culture mathématique chez les jeunes de 15 ans » est définie à travers la combinaison de trois dimensions: la complexité conceptuelle, le contenu à mobiliser et les réalités ou situations envisagées. Ainsi, la complexité conceptuelle est envisagée à travers trois groupes de compétences de plus en plus complexes: la reproduction de connaissances acquises ; les connexions à établir entre l'énoncé d'un problème et les connaissances mathématiques pour traduire le problème de départ en un problème mathématique; la « réflexion élaborée» qui implique des capacités de modélisation, d'argumentation, d'abstraction et de généralisation. Cette dimension « groupes de compétences » va être croisée avec quatre types potentiels de contenus : le domaine des « quantités », celui de « l'espace et des formes », celui de « l'incertitude» (les statistiques et les probabilités) et celui des « variations et relations». Ces deux dimensions seront aussi combinées avec des situations issues de quatre champs : les réalités personnelles de l'élève ; le champ éducatif ou professionnel ; les réalités dites publiques ; et le champ des réalités scientifiques. Cet aperçu montre bien que la conception des items du test PISA repose sur un cadre conceptuel élaboré. Il en va de même pour chacune des « disciplines» évaluées : la résolution de problèmes ; la culture scientifique ; et la compréhension de l'écrit. Ainsi, pour donner un second exemple, la compréhension de l'écrit repose également sur trois dimensions à combiner : le type de processus de lecture dont les modalités sont « retrouver des informations », « interpréter le texte », « réfléchir sur le contenu du texte » ; le type de textes, à savoir « les textes continus » et « les textes non continus » ; l'usage pour lequel le texte a été conçu, à savoir « à des fins personnelles », « à usage public », « à usage éducatif » et « à usage professionnel ».

Évaluer ce que les concepteurs déclarent évaluer amène nécessairement à se poser toute une série de questions. Quelle est la pertinence des fondements didactiques à la base ? Sont-ils les mêmes dans les différents pays qui ont participé à l'évaluation ? Quels sont les choix jugés prioritaires dans certains pays et que l'on ne trouve guère dans le test ? Les questions traduisent-elles bien ce qu'elles sont censées évaluer. Sont-elles comprises de la même façon dans des pays de culture différente ? Les traductions d'une

langue à l'autre n'amènent-elles pas parfois à évaluer des choses différentes selon les pays ? À l'intérieur d'un même pays, les élèves de 15 ans répartis sur des années d'études différentes ou dans des orientations différentes (pas de tronc commun ou de « vrai » tronc commun) bénéficient-ils d'apprentissage reposant sur les mêmes fondements didactiques ?

Toutes ces questions ne sont pas innocentes et méritent une réflexion pour tenter d'approcher la juste relativité des choses. Donnons quelques exemples qui montreront que, malgré la volonté de sérieux et les compétences des responsables des enquêtes, certains résultats doivent être relativisés et interprétés en contexte, ce qui n'enlève pas le caractère intéressant des résultats fournis pour tenter d'améliorer chacun des systèmes éducatifs (voir 2^o partie).

En tant qu'expert étranger à la Commission nationale des Programmes du Ministère français de l'Éducation lorsque celle-ci existait encore, j'ai eu la chance de participer à certaines journées consacrées à la comparaison des programmes des différents pays européens. Si mon étonnement fut grand lorsque la journée consacrée aux programmes d'histoire montra des différences beaucoup plus grandes encore que ce que je pouvais imaginer (heureusement que les enquêtes PISA ne portent pas sur cette discipline!), mon étonnement le fut autant lors de la journée consacrée aux programmes de mathématiques. Alors que l'on pourrait penser au caractère invariant de la discipline, la comparaison des programmes montra des fondements didactiques très différents selon les pays du nord ou du sud de l'Europe. Au sud (dont la France), l'enseignement des mathématiques est tourné vers la production du savoir mathématique pour lui-même et est l'outil de sélection par excellence pour identifier ceux qui auront accès aux classes préparatoires, aux Grandes Ecoles et à l'élite des chercheurs. Au nord, les fondements didactiques sont nettement différents. Sur la base du constat que plus de 90% des élèves et des étudiants n'auront pas pour fonction de produire du savoir mathématique, mais auront besoin des outils mathématiques pour faire face aux problèmes de la vie professionnelle et vivre dans une société de plus en plus complexe, l'enseignement des mathématiques est avant tout une discipline outil (De Ketele, 2005). L'examen des manuels de mathématiques des pays du nord et du

sud ne fait que renforcer ce constat : plus inductifs et plus centrés sur la résolution de problèmes concrets au nord ; plus déductifs et plus abstraits au sud. Or nous savons que les enseignants sont particulièrement influencés dans leurs pratiques par les manuels, beaucoup plus encore que par les programmes officiels (Gérard & Roegiers, 2003). Notons qu'une même différence entre les pays du nord et du sud de l'Europe a été identifiée lors de la journée destinée à la comparaison des programmes d'enseignement des sciences.

Il est imprudent de tenir trop vite des propos comme « le système éducatif de tel pays est performant ou peu performant dans telle discipline ». En effet, ce n'est pas « toute » la discipline qui est évaluée, ni même tous les aspects de la discipline jugés prioritaires par les autorités éducatives ou par les enseignants eux-mêmes. Si on prend, à titre d'exemple, la maîtrise de la langue d'apprentissage (forcément variable d'un pays à l'autre), les enquêtes PISA évaluent seulement la compréhension de l'écrit. Elles n'évaluent pas, du moins directement, la maîtrise technique de la langue (orthographe, grammaire, conjugaison, analyse des composants de la phrase), ni l'expression orale, ni la production écrite. Le poids accordé à ces différents aspects peut être différent selon les pays, selon la nature de la langue, selon les choix didactiques, selon les pratiques dans les classes. Ainsi, dans les pays francophones, et plus largement latins, la plupart des enseignants consacrent un temps très important aux aspects techniques de la langue. Dans les pays scandinaves qui valorisent beaucoup le travail en groupes, le temps consacré au travail sur base de l'écrit est nettement plus important.

De nombreux débats ont porté sur la nature des items, sur la difficulté d'avoir, après traduction car la version originale est anglaise, des items équivalents, sur la qualité des adaptations contextuelles ... et même sur la qualité technique des consignes. Disons d'emblée que les responsables des enquêtes PISA ont fait des efforts considérables pour concevoir un dispositif qui assure la qualité des items et plus largement du dispositif de passation. Baye, Demonty, Fagnant, Lafontaine, Matoul et Monseur (2004, p.12) les synthétisent comme suit :

- ◆ La construction de l'échantillon d'écoles participant à l'épreuve doit se faire selon des règles

strictes, définies et vérifiées par un organisme dans ce domaine ;

- ◆ L'administration des épreuves doit respecter des règles précises, les administrateurs de tests étant des personnes extérieures à l'établissement, dûment formé à cette tâche ; des visites de contrôle de la qualité du déroulement des séances sont organisées (en Communauté française de Belgique, par exemple, elles ont été effectuées par les inspecteurs de l'enseignement) ;

- ◆ La correction des épreuves est réalisée par des correcteurs formés à cette tâche. De nombreux items sont codés plusieurs fois par différents correcteurs nationaux et un certain nombre d'entre eux sont à nouveau corrigés en dehors du pays, pour évaluer la fidélité des corrections au niveau international ;

- ◆ La traduction des textes et des questions et les adaptations nationales sont réalisées par des spécialistes des contenus pour s'assurer que les termes utilisés dans les épreuves sont bien ceux qui sont généralement employés dans le système scolaire de chaque pays participant.

Ces précautions ne rassurent cependant pas tous les experts. Parmi eux, Romainville (2002) qui a examiné les épreuves PISA 2000 est particulièrement sévère. « Bien que des mécanismes interactifs de discussion et d'approbation des tests aient été mis en place, le développement effectif des épreuves était de la responsabilité d'un consortium, à dominante anglophone, d'entreprises australienne, japonaise, américaines et hollandaise (dont certaines sont d'ailleurs privées), dirigé par un organisme australien » (p.88-89). L'auteur poursuit son argumentation en disant que

- ◆ 78% des questions provenaient de pays de langue anglaise ;

- ◆ La plupart des données soumises aux élèves dans les situations sont également originaires de ces pays ;

- ◆ Les contextualisations censées ancrer les questions de vie quotidienne ne sont guère significatives pour les élèves de nos pays (comme « ... de la compagnie Iran Air, en partant de Téhéran et à destination de Sydney.. . », ce qui ne dit pas grand chose aux élèves français ou belges en l'an 2000) ;

- ◆ La version originale des questions était en langue anglaise et les spécialistes connaissent les pièges des traductions, même pour les pays de même langue (exemple : l'expression « diagramme en camembert » utilisée dans la version française est

sans doute bien connue des élèves français, alors que les écoles belges utilisent l'expression « graphiques en portions de tarte »).

Romainville va encore plus loin dans sa critique. Il débusque de nombreuses ambivalences dans la formulation des questions et des consignes. Citons un des nombreux exemples qu'il donne :

« Une autre série de questions porte sur deux lettres d'élèves qui donnent leur avis sur les graffitis. Voici une de ces questions « On peut parler de ce que dit une lettre (son contenu) ; on peut parler de la façon dont une lettre est écrite (son style). En faisant abstraction de votre opinion, qui a écrit la meilleure lettre d'après vous ? Justifiez votre réponse en vous référant à la façon dont la lettre choisie est écrite ». Cette consigne est passablement alambiquée. Comment comprendre la juxtaposition des deux expressions « en faisant abstraction de votre opinion » et « d'après vous » ? Il s'agit bien sûr de faire abstraction de son opinion sur le contenu des lettres, mais cela ne va pas de soi. D'ailleurs la grille de correction indique que le crédit complet

sera accordé si l'élève « explique son opinion en faisant référence... ». On pourrait y perdre son latin. » (p. 93)

Sur la base de nombreux exemples, Romainville rejette donc ce qu'il appelle « le modèle anglo-saxon d'évaluation des compétences » et il partage la critique faite par le ministère français de l'éducation : « Les cadres d'analyse sont aussi influencés par les modèles sous-jacents : c'est d'ailleurs la raison pour laquelle le ministère français de l'Éducation, tout en trouvant l'exercice « intéressant », en a contesté certaines conclusions. Il s'est inquiété de l'idée selon laquelle seul le modèle anglo-saxon serait performant : en effet, il est frappant de constater que la plupart des pays qui arrivent en tête du classement sont anglo-saxons ; le premier pays francophone, la France, arrive péniblement à la quatorzième place, tout juste au-dessus de la Moyenne. » (p. 89-90)

Ces propos de Romainville portent sur l'enquête PSA 2000. Il est intéressant d'examiner s'ils s'avèrent encore exacts pour PISA 2003. Le tableau suivant répond à cette question.

Tableau 1 : Performances aux 4 tests PISA 2003 de 14 pays regroupés en 4 champs culturels : résultats et rangs

Regroupement	Pays	Culture mathématique	Compréhension de l'écrit	Culture scientifique	Résolution de problèmes
I Pays anglo-saxons	Australie	524 (3)	525 (3)	525 (2)	530 (3)
	Canada	532 (2)	528 (2)	519 (4)	529 (4)
	États-Unis	483 (11)	495 (8)	491 (7)	477 (11)
	N-Zélande	523 (4)	522 (4)	521 (3)	533 (2)
II Europe sous influence anglo-saxonne	Danemark	514 (5)	492 (9)	475 (3)	517 (6)
	Finlande	544 (1)	543 (1)	548 (1)	548 (1)
	Norvège	495 (9)	500 (6)	484 (10)	490 (9)
	Suède	509 (7)	514 (5)	506 (6)	509 (7)
III Europe francophone	France	511 (6)	496 (7)	511 (5)	519 (5)
	C.F. Belgique	498 (8)	477 (11.5)	483 (11)	496 (8)
IV Europe du sud	Espagne	485 (10)	479 ((10)	487 (8)	482 (10)
	Grèce	445(14)	472 (14)	481 (12)	449 (14)
	Italie	466 (12.5)	476 (13)	486 (9)	470 (12.5)
	Portugal	466 (12.5)	477 (11.5)	468 (14)	470 (12.5)
I	Moyenne	515.5	517.5	514	517.25
II	Moyenne	515.5	512.25	503.25	516
III	Moyenne	504.5	486.5	497	507.5
IV	Moyenne	465.5	476	480.5	467.75

Ces résultats n'infirmant pas la thèse de Romainville puisque les pays typiquement anglo-saxons ou sous forte influence anglo-saxonne ont globalement les meilleurs résultats et occupent pour la plupart les meilleurs rangs. Il est à noter que les Etats-Unis font exception. Certains attribueront cette faible performance au fait que la langue maternelle la plus parlée dans ce pays est l'espagnol et non l'anglais ; mais comment interpréter les très bonnes performances du Canada qui non seulement comprend une part non négligeable de francophones, mais aussi une communauté multiculturelle importante. Il faut aussi noter que la France occupe la meilleure place des pays non anglo-saxons (regroupements I et II) et devance même un certain nombre d'entre eux. Enfin, les pays de culture plus spécifiquement méditerranéenne ont systématiquement les moins bonnes performances.

Une autre façon de relativiser les résultats est de prendre en considération le fait que selon les pays, les élèves âgés de 15 ans se répartissent sur un nombre plus ou moins grand d'années d'études, voire de filières. C'est le cas des pays qui ne pratiquent pas un vrai tronc commun et où les élèves, suite à de nombreux redoublements, se retrouvent dans des années d'études nombreuses, aux programmes et aux fondements didactiques différents des programmes correspondant aux élèves de 15 ans n'ayant pas redoublé ou suivant la filière d'enseignement générale. C'est le cas de tous les pays des catégories III et IV du tableau 1 et c'est aussi le cas des Etats-Unis. À ces pays s'ajoutent des pays comme l'Allemagne et la Suisse qui privilégient une formation précoce technique ou professionnelle et non une filière d'enseignement général.

Tableau 2 : Performances aux 4 tests PISA 2003 des élèves de 15 ans de la Communauté française de Belgique selon l'année d'études (cfr. Baye et al., 2004)

Année d'études	Redoublements	Culture mathématique	Compréhension de l'écrit	Culture scientifique	Résolution de problèmes
2 ^o année	2	351	320	343	364
3 ^o année	1	443	422	429	445
4 ^o année	0	546	525	530	542
5 ^o année	0 (avance)	616	603	598	613

Pour bien lire ce tableau, il faut savoir qu'en Communauté française de Belgique il y a 6 ans de scolarité primaire et 6 ans de scolarité dite secondaire, que la 2^o année de l'enseignement secondaire correspond à la 4^o année du collège en France, mais qu'une partie non négligeable d'élèves se trouvent déjà dans des filières professionnelles. Ces résultats font dire aux auteurs :

« Les résultats des élèves « à l'heure » (en 4^o) sont relativement bons dans tous les domaines ; les scores des élèves « en avance » d'un an sur le programme sont encore plus élevés. À titre comparatif, on peut noter que le score des élèves de 4^o année approche, dans chacune des disciplines, le score moyen des pays en tête des différents classements. Par exemple, en mathématiques, les scores de nos élèves « à l'heure » est proche du score moyen des trois pays de tête : la Communauté flamande obtient un score moyen

de 553, Hong Kong (Chine) un score de 550 et la Finlande de 544. » (Baye & al., 2004, p. 102)

Que conclure de cette 1^o partie ? Certainement pas de « jeter le bébé avec l'eau du bain », car les résultats des enquêtes internationales sont toujours « intéressants » et méritent un examen pour leur donner un sens. L'hypothèse d'une « plus-value » due au fait pour un élève d'être éduqué dans une sphère d'influence anglo-saxonne n'est pas négligeable. Romainville (2002) fait d'ailleurs appel au travail de Fitoussi (2001) pour dire que cette plus-value existe aussi dans l'enseignement supérieur :

« Les chercheurs français, pour être reconnus, doivent travailler beaucoup plus que leurs collègues américains, sauf à renoncer à la fois à animer le débat public en économie dans leur pays et à alimenter les revues françaises » (Fitoussi, 2001, p. 114).

Mais il faut noter aussi que les pays francophones et latins ont leur part de responsabilité, comme le souligne le rapport de Bottani et Vrignaud (2005) commandité par le Haut Conseil de l'Évaluation. Ces auteurs notent que la France (comme d'autres pays) ont laissé la part belle aux experts « anglo-saxons », faute d'interlocuteurs maîtrisant les compétences nécessaires en éducativité pour discuter d'égal à égal avec les experts « anglo-saxons ». Lors de l'examen de ce rapport, les responsables de la DEP attiraient aussi l'attention sur la difficulté pour eux de trouver de bons experts dans ce domaine et soulignaient la quasi-absence de formation méthodologique à ce niveau dans les UFR de sciences de l'éducation.

L'impact des enquêtes internationales sur les systèmes éducatifs

Comme le soulignent les auteurs précités ou d'autres auxquels nous nous référerons (dont Dupriez, Dumay, Vandenberghe), il existe depuis un certain nombre d'années une pression évaluatrice de plus en plus forte, en grande partie due à la fois à la nécessité d'élever le niveau de l'éducation pour être compétitif

et au coût de plus en plus élevé de l'éducation. Les enquêtes nationales (en France : les travaux de la DEP, du Haut Conseil de l'évaluation qui a cédé actuellement sa place au Haut Conseil de l'Éducation, les nombreuses commissions dont la commission Thélot) et internationales (principalement de l'IEA et de l'OCDE) sont la manifestation la plus évidente de cette pression. Nul doute que ces travaux ont suscité de nombreux remous, pour ne pas dire des remous médiatiques importants dans certains pays. Il est d'ailleurs curieux de noter que plus les pays présentent une forte différenciation de leurs systèmes éducatifs, plus ils ont vécu et vivent encore de véritables « tempêtes médiatiques » (Buhlman, 2002 ; Demailly, 2006 a et b ; Mac Gaw, 2004 ; Mac Kenzie, 2004 ; PISA, 2006). C'est notamment le cas de l'Allemagne (avec l'autonomie des Lands), de la Suisse (avec ses nombreux cantons) et de la Belgique (avec ses trois communautés éducatives). Comment les journalistes, à la recherche de manchettes éditoriales fracassantes, ne réagiraient pas en découvrant des résultats comme ceux mentionnés dans le tableau suivant.

Dans tous les pays où existaient des systèmes éducatifs

Tableau 3 : Performances aux tests PISA 2003 pour les 3 communautés en Belgique

Pays	Communautés	Culture mathématique	Compréhension de l'écrit	Culture scientifique	Résolution de problèmes
Belgique	Flamande	553	530	529	547
	Germanophone	515	499	492	514
	Française	498	477	483	496

différenciés, des débats passionnés ont eu lieu. L'exemple de la Suisse, dont la réputation est d'éviter les confrontations, est éloquent. La consultation des articles publiés par ARLE (Association Refaire l'École) montre la violence des propos en opposant « l'école qui instruit » (position défendue) à « l'école qui milite » (position rejetée). À titre d'illustration, citons un florilège d'expressions utilisées par les auteurs des articles de cette association :

- ◆ « Le rapport PISA, qui a tant secoué la république genevoise, a introduit le vocable de « littératie » (qui dira la tristesse de certains mots?)»
- ◆ « J'ai de la peine à penser que toute

l'entreprise de démolition scolaire actuelle ne soit le fruit d'une erreur »

- ◆ « Ce refus d'être « éducasteur » a fait place à un discours que partout ailleurs on aurait trouvé hallucinant : puisqu'on ne parvient pas à donner à tous un diplôme, mettons les pupitres en rond, l'élève au centre, ... »
- ◆ « Il est temps de libérer l'école du carcan des pédagogistes ».

Heureusement, à côté de ses positions tranchées qui relèvent de l'opinion, d'autres associations cherchent à comprendre en s'appuyant sur des travaux qui s'appuient sur des faits vérifiés. C'est le cas, par exemple, de l'association «

Former sans Exclure » (2006) dont on citera par contraste quelques extraits:

- ◆ Face au redoublement défendu avec acharnement par certains, l'association cite une conclusion du Fonds national suisse de la recherche scientifique : « Les redoublements scolaires ne débouchent pas sur les progrès d'apprentissage escomptés. Une enquête menée pour la première fois au niveau suisse sur l'efficacité du redoublement montre que les élèves qui ont redoublé perdent cette avance au cours de l'année répétée et même qu'ils se retrouvent à nouveau en retard à la fin de l'année suivante ».
- ◆ Face à l'explication avancée par ARLE que les moins bons résultats de certains cantons s'expliquent par une plus forte population étrangère et à une moins forte intériorisation des valeurs traditionnelles (caractéristiques des cantons plus ruraux), l'association fait référence au Canada, pays présentant une population multiculturelle forte mais faisant partie des pays ayant les meilleures performances aux enquêtes PISA.

Les enquêtes internationales ont eu heureusement

un impact positif auprès d'organismes publics ou indépendants qui ont tenté de comprendre ce qui se cache derrière des résultats présentés trop souvent comme un palmarès. De nombreuses analyses complémentaires sur les données de départ (mises en accès public) ont été faites, non seulement par l'OCDE, mais aussi par des centres de recherche, comme l'IREDU en France, Le GIRSEF et le SPE en Belgique.

L'égalité ou l'inégalité des chances ou ce qui a été progressivement nommé sous le terme d'équité ou d'iniquité (Pourtois, 2005 ; Dupriez, 2005 et 2006)) sont les problématiques auxquelles les pays ont été les plus sensibles suite aux nombreuses ré-analyses des données des différentes enquêtes internationales. Zachary, Dupriez et Vandenberghe (2003) se sont penchés sur la question de savoir si le fait pour un élève d'être dans telle classe ou telle autre pouvait rendre compte d'une partie de la variabilité des résultats au test de compréhension en lecture (enquête IEA 1990) et au test de Mathématiques (TIMMS 1995). Nous avons recompilé leurs données (dans la mesure où elles étaient disponibles) en fonction de la classification établie dans le tableau 1.

Les résultats exprimés en pourcentages doivent être

Tableau 4 : Pourcentages de la variance des scores aux tests de lecture (1990) et de Mathématiques (1995) en fonction de 4 regroupements de pays.

Regroupement	Pays	Compréhension en lecture (1990)	Mathématiques (1995)
I Pays anglo-saxons	Australie		
	Canada		22,60 %
	États-Unis		48,31 %
	N-Zélande		
II Europe sous influence anglo-saxonne	Danemark	11,0 %	
	Finlande	7,7 %	
	Norvège	9,4 %	10,42 %
	Suède	12,0 %	
III Europe francophone	France	38,0 %	22,67 %
	C.F. Belgique	45,9 %	30,80 %
IV Europe du sud	Espagne	24,4 %	16,71 %
	Grèce	24,6 %	15,28 %
	Italie	31,1 %	13,25 %
	Portugal	33,7 %	



lus comme suit : plus le pourcentage est élevé, plus le fait d'appartenir à telle classe plutôt qu'à telle autre a un impact sur les performances des élèves : ou, en d'autres termes, changer un élève de classe aura d'autant moins d'impact sur ces résultats que le pourcentage de variance est faible. Les résultats montrent clairement plusieurs choses :

- ◆ Changer son enfant de classe pose beaucoup moins de problèmes dans les pays de la sphère anglo-saxonne que dans les pays francophones et méditerranéens, car les performances sont plus homogènes d'une classe à l'autre.

- ◆ Tout en conservant des différences de performances entre les classes fort élevées, les pays francophones et méditerranéens ont cependant réussi à diminuer quelque peu cette hétérogénéité des performances entre les classes entre 1990 et 1995 (la différence ne semble pas due à la différence de tests).

La plus grande homogénéité des performances (faible variance) entre les classes de certains pays signifie que ces classes ont une mixité sociale plus grande que dans celles des pays où la variance est plus grande (les élèves se retrouvent dans les classes en fonction de leur appartenance sociale et de leur parcours scolaire antérieur). En retravaillant les données des enquêtes OCDE 2000, Vandenberghe (2004) montre que le regroupement des élèves par classe de niveau d'aptitude homogène n'a guère d'impact sur les performances. Par contre, il met en évidence un « effet de pairs » en faveur des élèves

moins performants: le fait pour ces élèves d'être dans une classe avec des élèves plus performants leur fait faire de plus grands progrès que s'ils étaient dans des classes de même niveau d'aptitude. Ce type de résultats est de plus en plus connu par les autorités éducatives (par exemple en France et en Belgique) qui se demandent comment introduire plus de mixité sociale dans les classes.

Le problème de l'effet d'appartenance à telle classe plutôt qu'à telle autre est déjà un indicateur d'équité ou d'iniquité. Beaucoup d'analyses a posteriori des résultats des enquêtes ont tenté d'étudier cette problématique via des indicateurs variés. Dupriez et Vandenberghe (2004) ont calculé le pourcentage de variance des résultats des enquêtes PISA 2000 en prenant comme variables prédictrices le niveau d'éducation de la mère, un indice socio-économique des parents et le pays de naissance du père (indicateur 1). Baye et al. (2004) ont fait de même pour l'enquête PISA 2003 mais en se basant sur les indicateurs fournis par les chefs d'établissements (indicateur 2). Dans le tableau 5, nous présenterons ces résultats pour les performances en mathématiques pour les quatre regroupements de pays. À ces résultats, nous ajouterons deux autres indicateurs : les différences de résultats en mathématiques entre les élèves issus du quart des familles les moins favorisées et ceux des élèves issus du quart des familles les plus favorisées (indicateur 3) ; le pourcentage des élèves de 15 ans ayant au moins redoublé une fois dans leur scolarité (indicateur 4).

Tableau 5 : Influence de quelques indicateurs d'équité (iniquité) sur les performances en mathématiques.

Regroupement	Pays	Maths 2000 R ² Indicateur 1	Maths 2003 R ² Indicateur 2	Maths 2003 Différence Indicateur 3	Redoublements % (2003) Indicateur 4
I Pays anglo-saxons	Australie		13,7 %	77 points	9 %
	Canada		10,5 %	63 points	9,7 %
	États-Unis		19 %	82 points	11,3 %
	N-Zélande		16,8 %	79 points	4,5 %
II Europe sous influence anglo-saxonne	Danemark	9,9 %	17,6 %	73 points	3,4 %
	Finlande	4,9 %	10,8 %	61 points	2,8 %
	Norvège	6,8 %	14,1 %	70 points	0 %
	Suède	9,8 %	15,3 %	73 points	3,4 %
III Europe francophone	France	8,6 %	19,6 %	88 points	38,3 %
	C.F. Belgique	21,8 %	25,5 %	110 points	37,7 %
IV Europe du sud	Espagne	9,2 %	14 %	66 points	28,6 %
	Grèce	9 %	15,9 %	84 points	7 %
	Italie	4,6 %	13,6 %	72 points	15 %
	Portugal	10,6 %	17,5 %	80 points	29,5 %

Les résultats de ce tableau sont particulièrement éloquentes :

- ◆ Les pays francophones sont les champions (désolé pour le terme qui sent le palmarès!) du redoublement et ces redoublements s'accompagnent d'iniquité (indicateurs 2 et 3 + indicateur 1 pour la Communauté française de Belgique).
- ◆ Les pays scandinaves sont les champions du non-redoublement. Avec les pays anglo-saxons et les pays méditerranéens, ils n'accroissent pas de façon aussi forte l'iniquité due à la position sociale des parents.
- ◆ Les pays francophones ont accru au fil du temps l'iniquité de leur système.

Si on croise les indices d'efficacité du tableau 1 et les indices d'équité du tableau 5, on constate que tous les cas de figure existent. Ainsi, par exemple, le Canada et la Finlande sont très efficaces et parmi les plus

équitables ; la Communauté française de Belgique est peu efficace (contrairement à la Communauté flamande qui se trouve parmi les plus efficaces) et très inéquitable (au même titre que la Communauté flamande) ; la France est moyennement efficace (un peu supérieur à la moyenne), mais est un des pays les moins équitables ; les pays de l'Europe du sud sont très peu efficaces, mais moyennement équitables.

Les autorités des pays les moins équitables ont été touchées fortement par ces constats. De nombreuses hypothèses ont été formulées pour comprendre pourquoi. Parmi celles-ci, on trouve le plus fréquemment le taux d'élèves d'origine étrangère, l'importance de l'enseignement privé et la structure scolaire (tronc commun vs orientation précoce différenciée). Nous avons déjà mentionné que le Canada comprenant une population d'origine étrangère importante avait des indicateurs d'efficacité et d'équité parmi les plus favorables. Baye et al. (2004) montrent que sur la base des résultats PISA 2003, il est intéressant de distinguer le fait de n'être

pas né dans le pays et le fait de ne pas parler la langue de l'école. Parmi les cas de figure, distinguons :

- ◆ Une iniquité élevée tant pour les élèves non natifs que pour les élèves ne parlant pas la langue de l'école, comme c'est le cas pour la Communauté française et la Communauté flamande de Belgique, les Pays-Bas, la Suisse, le Japon et, curieusement, pour la Finlande et la Suède, qui cependant discriminent relativement peu les élèves en fonction de la classe sociale et pratiquent peu le redoublement ;
- ◆ Une iniquité élevée surtout pour les élèves non natifs, mais non en fonction du fait de ne pas parler à la maison la langue de l'école : c'est le cas du Portugal (on peut émettre l'hypothèse qu'un grand nombre d'élèves d'origine étrangère sont issus des anciennes colonies portugaises) ;
- ◆ Une iniquité élevée surtout pour les élèves qui ne parlent pas à la maison la langue de l'école : c'est le cas de la France, de la Corée et de la République slovaque.

Ces résultats nous conduisent à dire qu'une étude

de cas sur la politique canadienne mériterait d'être menée. En effet, ce pays a développé depuis plusieurs années une politique ambitieuse (à la fois en termes de qualité visée et de budgets débloqués) visant à apprendre la langue de l'Etat pour les étrangers qui s'installent dans le pays. S'inspirant de cet exemple, la Communauté flamande de Belgique est en train de mettre en place également une politique visant à intégrer les étrangers dans la culture flamande.

Vandenberghe (2003) et Vandenberghe & Robin (2003) ont comparé les performances des élèves de 15 ans selon le nombre d'élèves qui sont dans (I) le privé au sens strict du terme (gestion totalement indépendante du public), (II) le privé dépendant du gouvernement (écoles ne cherchant pas le profit et dépendant au moins pour 90% des fonds publics) et (III) le public. Dans le tableau 6 regroupant des pays européens pour lesquels les données sont disponibles, nous donnons le % d'élèves dans chacune des trois catégories d'écoles et les résultats obtenus au test de compréhension en lecture.

Un premier examen donne l'impression qu'à une exception près (le privé dépendant du gouvernement

Tableau 6 : Performances aux 4 tests PISA 2003 de 14 pays regroupés en 4 champs culturels : résultats et rangs

Regroupement	Pays	Pourcentages d'élèves			Compréhension en lecture		
		I privé	II privé G.	III public	I privé	II privé G.	III public
I Pays anglo-saxons	UK	5 %	62 %	95 %	609		517
	Irlande			38 %			
II Europe sous influence anglo-saxonne	Danemark		23 %	77 %		500	498
	Finlande		3 %	97 %			
	Suède		3 %	97 %			
III Europe francophone	France	8 %	14 %	78 %	519	495	500
	C.F. Belgique		70 %	30 %			
IV Europe du sud	Espagne	8 %	29 %	62 %	539	509	479
	Italie	4 %		96 %	515		488

en France), les élèves des échantillons du privé présentent des scores supérieurs aux élèves des échantillons du public. Il faut cependant relativiser ces résultats. En effet, Vandenberghe (2003) et Vandenberghe et Robin (2003) ont réexaminé ces résultats (plous ceux de 13 autres pays) en tentant de neutralisant les caractéristiques individuelles, sociales et culturelles (13 indicateurs). La conclusion de leurs analyses révèle alors une absence d'effets systématiques du privé sur les performances.

L'hypothèse de l'influence de la structure scolaire sur les performances des élèves a fait l'objet de plusieurs études, dont celles de Dupriez et Dumay (2005 et 2006). Les auteurs partent de la typologie développée par Mons (2004) dans sa thèse de doctorat. Cet auteur distingue quatre modèles de structures scolaires à travers quatre modes de gestion de l'hétérogénéité :

- ◆ Le « modèle de séparation » basé sur le principe de séparation des élèves et de sélection précoce. C'est le cas surtout des pays germaniques (Allemagne, Autriche, Hongrie, Suisse et Luxembourg), mais c'est aussi dans une certaine mesure le cas en Belgique et aux Pays-Bas.
- ◆ Le « modèle de l'intégration à la carte » qui maintient jusqu'à l'âge de 16 ans un programme relativement commun, mais gère l'hétérogénéité par des politiques de classes de niveau ou de

regroupement des élèves selon leurs résultats. C'est le cas de pays anglo-saxons comme les USA, le Canada, le Royaume-Uni et la Nouvelle-Zélande.

◆ Le « modèle d'intégration uniforme » qui repose sur un tronc commun et qui soumet tous les élèves aux mêmes conditions éducatives, mais gère l'hétérogénéité par une politique forte du redoublement. C'est surtout le cas de certains pays latins, comme la France, l'Espagne, le Portugal, l'Argentine, le Chili.

◆ Le « modèle d'intégration individualisée » qui repose sur un programme commun et le refus du redoublement mais qui développe de nombreuses formes de stratégies d'individualisation et de différenciation pédagogiques. C'est le cas des pays scandinaves, comme le Danemark, l'Islande, la Finlande et la Suède).

Dans leur étude, Dupriez et Dumay (2006) comparent 8 pays (2 par modèle). Dans le tableau 7, nous extrayons de cette étude les résultats obtenus au sein de l'école par les élèves « peu performants (PP) » et « très performants (TP) » (colonnes 2 et 3) en mathématiques au test PISA 2003, puis les pourcentages de variance expliquée par la composition académique au sein de l'école (c'est-à-dire la distribution des élèves peu et très performants), le niveau socioculturel moyen des élèves au sein de l'école, le climat de discipline et le concept de soi en mathématiques.

Tableau 7 : Performances des élèves peu (PP) et des élèves très (TP) performants au test de mathématiques PISA 2003 et pourcentages de variance des résultats expliquée par la composition académique (Comp), le niveau socioculturel (SES), le climat de discipline (climat) et le concept de soi en mathématiques (C.Soi).

Modèle	Pays	Scores des élèves		% variance expliquée			
		PP	TP	Comp.	SES	Climat	C.Soi
Intégration individualisée	Finlande	536	553	12%	2%	2%	49%
	Suède	491	525	18%	9%	4%	39%
Intégration uniforme	Écosse	505	541	25%	17%	18%	29%
	États-Unis	442	515	38%	26%	13%	25%
Intégration à la carte	France	446	566	69%	47%	4%	19%
	Espagne	454	512	35%	22%	7%	23%
Séparation	Belgique franc.	421	555	52%	48%	12%	7%
	Allemagne	416	576	70%	60%	10%	12%

Ce tableau nous permet de dégager des conclusions intéressantes pour les gestionnaires des systèmes éducatifs :

- ◆ Les différents modèles ne font guère de différence en ce qui concerne les résultats des élèves les plus performants.
- ◆ Mais le modèle de l'intégration individualisée présente des résultats particulièrement élevés pour les élèves peu performants, comparativement aux autres modèles (l'Écosse fait cependant exception).
- ◆ Le modèle de l'intégration fait également la différence pour la composition académique, le niveau socioculturel, le climat disciplinaire et le concept de soi. Ceci signifie que les écoles de ces pays présentent une plus grande mixité académique et sociale et des climats disciplinaires plus semblables (d'où les pourcentages de variance plus faibles). Par contre, le concept de soi joue un rôle plus important pour expliquer la variabilité des performances.
- ◆ Les modèles d'intégration à la carte et surtout le modèle de séparation présentent les résultats les plus faibles pour les élèves peu performants. Par ailleurs, la composition académique au sein de l'école, le niveau socioculturel et le climat disciplinaire sont les variables qui expliquent la plus grande part de la variance des performances.

Ce type de conclusions a suscité de nombreux commentaires et a fait l'objet de nombreux débats. Que l'on se réfère, par exemple, au rapport de la Commission sur l'avenir de l'École en France. En Belgique, les autorités politiques et éducatives ont mis en place une commission de pilotage et ont conçu un « Contrat pour l'École », dont la fonction est de lutter contre les trop fortes inégalités scolaires.

Conclusions

Au terme de ce parcours, nous revenons à l'idée de départ, à savoir la nécessité d'évaluer les évaluations internationales. Par là, nous soulignons le fait de donner du sens aux tableaux de bord fournis par les enquêtes internationales. Ceux-ci fournissent des résultats « intéressants », mais il est nécessaire de

les replacer dans leur contexte, de se souvenir qu'ils sont les fruits de nombreux facteurs sous-jacents et que les dispositifs qui les ont créés reposent sur des fondements plus ou moins explicites qui ne sont pas les seuls possibles ni nécessairement les plus souhaitables pour tous. À travers les analyses rapportées, nous montrons la prégnance d'un modèle anglo-saxon. On pourrait interroger le concept de compétence, tel qu'il est opérationnalisé dans les enquêtes PISA (De Ketele & Gérard, 2005 ; BIE, 2005). On peut aussi se demander s'il ne faut pas distinguer plus fondamentalement les socles incontournables de compétences à maîtriser par tous les élèves de 15 ans et différents autres niveaux de compétences qu'il est souhaitable de voir maîtriser de façon différentielle par les élèves, compte tenu de leur hétérogénéité. Ces questions sont actuellement d'une brûlante actualité, comme le montrent les travaux d'instance comme le Haut Conseil de l'Éducation en France et les institutions préoccupées par les problèmes d'analphabétisme fonctionnel (ANLCI, 2006).

Les analyses sur lesquelles nous nous sommes appuyé sont de nature quantitative. Bien évaluer les évaluations internationales supposerait de recourir aussi à des études de cas plus qualitatives qui donneraient de la chair aux froids constats des chiffres en prenant ensemble (comprendre = prendre ensemble) les différentes dimensions de tout phénomène. Nous avons déjà mentionné l'intérêt de comparer des pays à forte population étrangère et présentant des indicateurs d'iniquité inquiétants avec un pays comme le Canada (plus efficace et équitable). Comme on pourrait se pencher à l'intérieur d'un même pays sur ce qui permet de comprendre les différences entre deux Communautés ou deux Cantons ou deux Lands aux résultats particulièrement différenciés.

II. Regard à partir des classements internationaux

Les classements internationaux laissent peu de gens indifférents. Trop souvent, ce sont les faces les plus visibles qui sont mises en lumière par les différentes catégories d'acteurs qui en parlent en termes d'effet palmarès. Cependant des voix s'élèvent pour en révéler les aspects cachés : les voix les plus expertes attireront l'attention sur leurs avantages mais aussi leurs limites et leurs dangers ; animées par la déception ou le déni, d'autres voix tenteront même d'occulter leur indéniable valeur relative. En nous inspirant de notre expérience belge mais aussi étrangère (tant au Nord qu'au Sud), nous tenterons dans un premier temps d'analyser les classements internationaux en nous référant à 10 questions fondamentales : Qui ? Quoi ? Pour quoi ? Pour qui ? Comment ? Avec quel paradigme ? Avec quels effets directs ? Avec quels effets indirects ? Pour quel développement de la personne ? Pour quel développement de société ? Pour chacune de ces questions, nous tenterons de mettre en balance synthétiquement la face plus visible et la face plus cachée : ce que l'on voit facilement et ce qui se voit moins ; ce que l'on dit davantage et ce que l'on dit moins ; ce qui est davantage privilégié et ce qui l'est moins ; ce qui apparaît comme positif et ce qui peut se révéler négatif ou pervers ; ce qui est plus fondamentalement et explicitement visé et ce qui l'est moins ou pas du tout. Sur la base des réponses apportées à ces questions, nous tenterons dans un second temps de modéliser les inévitables tensions créées par le développement spectaculaire des évaluations internationales et des classements qu'elles génèrent. Un tel modèle, s'il est pertinent, permet de comprendre toute une série de phénomènes actuels et même d'anticiper des développements ultérieurs plus ou moins heureux. Notons que nos analyses portent sur les évaluations et classements internationaux ayant trait à l'enseignement de base. Il serait utile de les compléter, avec le même schéma, par des analyses portant sur l'enseignement supérieur avec la vogue des classements de Shangai, Times QS, Taiwan ou Webometrics. Si des différences importantes les distinguent (ex. : évaluation des performances

des élèves versus indicateurs des productions institutionnelles), nous pensons que nous devrions aboutir à une modélisation des tensions générées par les classements dans l'enseignement supérieur, dont certains éléments se rapprocheraient de la modélisation effectuée dans cet article à propos de l'enseignement de base.

1. Une analyse des faces visibles et cachées

Tout en sachant très bien que, sur le plan des prescriptions officielles, le concept d'enseignement obligatoire ne revêt pas la même réalité dans les différents pays, les instances internationales (UNESCO, OCDE, Union Européenne) tendent à inciter progressivement les pays à étendre l'obligation scolaire à 9 années (ce qui est souvent appelé « l'enseignement de base ») et même à certaines conditions à 12 années pour préparer les élèves à vivre et agir dans une société de plus en plus complexe. Ce n'est pas un hasard si la vogue des épreuves d'évaluations internationales (TIMMS, PISA, PASSEC, SACMEC) et les classements générés portent sur des années clés du primaire ou/et du secondaire. Même si ces épreuves sont loin d'être équivalentes sur tous les points, elles comportent bien des caractéristiques semblables sur leurs faces visibles et cachées. C'est ce qui sera surtout l'objet de notre attention.

Qui classe ? Dans tous les cas, ce sont des « consortiums internationaux » qui s'appuient sur des équipes internationales d'experts de la mesure et travaillent, dans chaque pays, avec une équipe nationale. Les équipes internationales sont composées essentiellement de grands spécialistes de l'évaluation (des experts de la conception des épreuves et des questionnaire) et de la mesure (des experts de la validation des items proposés et de l'élaboration des classements sur une échelle commune). Orientées vers l'évaluation et les classements des pays riches ou émergents, les épreuves PISA (actuellement les plus représentatives) sont le fruit d'un consortium indépendant, essentiellement anglo-saxon d'origine australienne et fortement soutenue par l'OCDE. Chargées d'évaluer les performances des systèmes éducatifs africains francophones, les épreuves PASSEC sont le fruit d'une décision des Ministres

de l'Éducation réunis au sein de la CONFEMEN; ici encore, le travail de conception, d'évaluation et de classement est confié à une équipe d'experts internationaux qui travaillent de façon assez comparable à ceux de PISA.

Qui ne classe pas ? Il y a eu des tentatives d'ONG internationales particulièrement sensibles aux besoins des publics défavorisés et aux enjeux sociopolitiques. L'expérience la plus connue est celle menée conjointement par l'UNESCO et l'UNICEF et centrée sur les compétences de base en mathématiques, dans la langue d'apprentissage et dans les « compétences de vie ». Ce n'est sans doute pas un hasard que cette initiative intéressante dans l'esprit ait été abandonnée faute de soutien (moyens insuffisants, expertise trop réduite, plaidoyer insuffisant). Il n'est cependant pas impossible que face aux insuffisances des initiatives actuelles, des voix s'élèvent et acquièrent suffisamment de force pour créer des dispositifs d'évaluation (mais pas nécessairement de classement) qui prennent en compte ce qui ne l'est guère jusqu'à présent.

Quoi ? Qu'est-ce qui est évalué et donc objet du classement ? Sur la base d'une analyse de différentes épreuves internationales comme PISA et le PASSEC, nous constatons que les performances évaluées se cantonnent aux « ressources », terme technique qui désigne à la fois les savoirs à restituer fidèlement (ex.: connaître une formule ou une règle), les savoir-faire ou applications de type scolaire (ex. : calculer l'aire d'une figure géométrique dont on donne les dimensions) et les savoir-faire ou applications habillées (ex. : calculer l'aire d'un champ qu'un propriétaire désire louer et dont on a le dessin et les dimensions requises). Comme cela a été montré ailleurs (De Ketele, 2006 ; Behrens, 2009), l'absence d'items relatifs à des niveaux de savoir plus complexe s'explique par leur rejet par le modèle mathématique de validation des épreuves utilisé par les experts internationaux. En effet, le TRI (modèle de la Théorie de Réponse à l'Item) suppose notamment l'unidimensionnalité de l'objet évalué ; or, les savoirs et savoir-faire complexes (qui dépassent donc le niveau des applications scolaires ou habillées) sont nécessairement multidimensionnels. Les experts essaient alors de décomposer des tâches complexes en leur composants unidimensionnels, mais on sait

que « le tout n'est pas la somme des parties ».

Quoi ? Qu'est-ce qui n'est donc pas évalué et donc objet de classement ? En liaison avec ce qui a été dit précédemment, les épreuves, et donc les classements, ne portent pas sur des performances liées à des savoirs et savoir faire complexe ; tout au plus, elles prennent en compte des applications habillées (l'énoncé présentant la situation et la tâche dit explicitement quel savoir ou savoir-faire il faut activer) et non des compétences au sens strict du terme (car, dans ce cas, l'énoncé présentant la situation et la tâche ne dit pas quels savoirs et savoir faire il faut mobiliser). Par ailleurs, pour des raisons facilement compréhensibles, les épreuves internationales ne fonctionnent que sur base de l'écrit ; tout ce qui relève de l'oral ou du comportemental n'est donc jamais pris en considération ; or le savoir communiquer dans diverses situations d'interaction ainsi que le savoir faire et le savoir être citoyen sont, sans aucun doute, des aspects qualitatifs très importants pour le développement personnel et citoyen.

Pour quoi ? Quelles sont les fonctions explicitement recherchées par les classements ? Puisque le produit est un palmarès qui range les pays en termes de performances sur des « ressources », qu'il fournit les écarts entre les catégories d'élèves les plus et les moins performants, qu'en outre il permet des analyses de seconde main par les chercheurs des différents pays, on peut donc dire avec les concepteurs que ces épreuves ont une fonction de signal ou d'alerte et même une fonction de régulation sur certains aspects, surtout si des études complémentaires sont commanditées par les responsables des pays (Rocher, 2009). Si cette régulation est conduite par les pays, de telles épreuves ont une fonction profonde d'homogénéisation et de standardisation, sans oublier que ceci agit essentiellement sur ce qui évalué (les ressources) au détriment de ce qui ne l'est pas.

Pour quoi ? Quelles sont donc les fonctions non recherchées par les classements et cependant agissantes ? Qui dit valorisation et homogénéisation vers un socle commun de ressources évaluables par des épreuves écrites entraîne de fait une certaine forme de minorisation ou de négligence envers ce qui n'est pas évalué. Certains chercheurs l'ont très bien montré. En se référant au Projet Harnos visant en Suisse à harmoniser les 24 systèmes éducatifs,

Behrens (2009) montre bien que l'essai de construire une évaluation standardisée commune n'a pas porté ses fruits, le modèle TRI de validation de l'épreuve rejetant trop d'items proposés suite aux marquages linguistiques et culturels. Plus largement, les épreuves standardisées et les classements internationaux entraînent donc de fait une décontextualisation et une déculturation des systèmes éducatifs, et donc une perte de la richesse liée à la diversité. Elles servent avant tout le processus de globalisation et non la prise en compte des aspects et besoins locaux.

Pour qui ? c'est-à-dire au profit de qui ? Officiellement, les évaluations et les classements sont une aide à la prise de décision des autorités politiques et éducatives des pays. Dans les faits, elles servent davantage soit au faire valoir des politiques et décisions prises par celles-ci, soit elles permettent à celles-ci de trouver des arguments pour faire passer des décisions qui sont déjà prises et dont les enjeux peuvent être bien différents que l'amélioration du système éducatif lui-même. Mais il est une autre catégorie d'acteurs qui sont les principaux bénéficiaires de ces classements abondamment diffusés et commentés par les médias ; il s'agit des parents des classes sociales culturellement et économiquement favorisées qui, soucieux d'offrir le meilleur à leurs enfants (souci compréhensible évidemment), vont adopter des comportements de consommateurs et rechercher les établissements qui par réputation sont censés être les meilleurs et recrutent les élèves potentiellement les plus performants. Les classements favorisent les comportements caractéristiques de ce que l'on appelle maintenant le « quasi-marché » et accentuent l'homogénéisation sociale.

Pour qui ? c'est-à-dire aussi au détriment de qui ? En liaison avec ce qui vient d'être dit, les classes sociales culturellement et économiquement les moins favorisées, et donc les moins informées et ne disposant pas des moyens nécessaires pour opérer certains choix, sont donc les grandes perdantes de l'opération, car elle ne pourront pas profiter des plus-values apportées par la composition des établissements et des groupes classes dits socialement diversifiés (Dumay et Dupriez, 2009). Le fossé grandit dans le pays entre les classes sociales ; dans les pays du Nord, un « quart-monde » se crée et se développe de plus

en plus ; dans les pays du Sud, les hommes puissants envoient leurs enfants faire leur scolarité dans le privé ou à l'étranger, laissant le public ou des organisations aux finalités parfois douteuses prendre en charge les enfants les plus démunis.

Comment ? Quelles sont les procédures mises en œuvre ? Comme il a été dit plus haut, les évaluations portent sur des performances facilement mesurables (les « ressources ») et les moins sujettes à des biais culturels et langagiers (effet du modèle de validation TRI). Tant dans les procédures de conception, que d'échantillonnage, que de passation, de traitement et d'établissement de la mesure, elles sont le fruit d'une très grande expertise ; elles ont donc le très grand avantage d'une standardisation très soignée et permettent d'établir des comparaisons valides, à condition de ne pas oublier ce qu'elles évaluent réellement et de ne pas généraliser abusivement à ce qu'elles n'évaluent pas.

Comment ? Quelles sont les procédures évitées ? PISA et le PASSEC ne recourent qu'aux modalités de réponse susceptibles d'obtenir les coefficients de fidélité les plus forts et donc à certaines formes de questions à choix multiples (seule procédé utilisé par le PASSEC) et aux questions de production à réponse courte (PISA utilise les deux procédés). Ces épreuves évitent donc les questions de production écrite longue et à plus forte raison les productions orales et les mises en situation. Or certaines compétences reconnues comme pertinentes pour la formation et l'éducation nécessitent de recourir à de tels dispositifs d'évaluation. Les épreuves internationales privilégient donc la fiabilité (la réduction des biais dans la mesure) aux dépens de la pertinence (ce qu'il importe d'évaluer) et parfois même de la validité (correspondance entre ce qui est évalué réellement et ce que l'on déclare évaluer). En effet, sur ce dernier point, les résultats sont souvent présentés comme évaluant des « compétences », alors qu'ils n'évaluent que des ressources et, parmi elles, tout au plus des applications habillées ou des sommes d'applications habillées correspondant effectivement à une compétence, mais non à la compétence elle-même.

Avec quel paradigme à l'œuvre ? On pourrait sans doute ressortir ici le titre judicieux utilisé par le CIEP pour le colloque organisé en mars 2009:

« Un seul monde ? Une seule école ? ». Il faut évidemment enlever le point d'interrogation pour caractériser le paradigme privilégié par le courant des évaluations et des classements internationaux. En effet, des instances internationales s'emparent de certains aspects fondamentaux de la gouvernance des systèmes éducatifs, notamment en fixant des standards internationaux facilement mesurables et mis au service du développement d'un citoyen susceptible de mobilité dans un monde global, en favorisant ainsi par la pression du palmarès la compétitivité mais aussi les fractures sociales et les inégalités scolaires, en considérant l'éducation comme un segment du marché scolaire (voir le N°52 de la Revue Internationale d'Éducation).

Avec quel paradigme mis en veilleuse ? Le revers du paradigme « un seul monde, une seule école » est la mise en veilleuse d'un paradigme que nous pourrions appeler « des écoles riches de leurs enracinements culturels, pensant certes globalement, mais agissant localement ». Si, pour prendre un exemple parmi d'autres, la compétence en lecture est un bien commun et un droit pour tout citoyen du monde, lire ne signifie pas la même chose pour des personnes de culture différente et vivant dans des contextes environnementaux non semblables (pensons à un enfant dans une école de brousse et un enfant dans une école urbaine richement dotée). Que peut faire une gouvernance globale, forcément anonyme et dépersonnalisante, face à une telle diversité de contextes et de richesses culturelles...si elle ne laisse pas la gouvernance à ceux qui, au niveau local, connaissent le contexte et les besoins liés à celui-ci et savent exploiter les richesses culturelles, humaines et environnementales propres à ces contextes ?

Avec quels effets indirects positifs ? À côté des effets déjà énoncés se développent toute une série d'effets indirects que l'on peut qualifier de positifs, comme par exemple : la prise de conscience par les pays que leur système éducatif n'est pas le seul possible (Roger-François Gauthier (RIE, n° 52) parle de « bigarrure du monde »), qu'il est sans doute intéressant d'aller voir ailleurs (que l'on pense aux nombreuses missions envoyées dans un pays comme la Finlande), qu'on comprend mieux son propre système éducatif en cherchant à en connaître d'autres (d'où la multiplication des échanges), mais que tout

système a une histoire propre. Et on pourrait citer d'autres effets positifs de cet ordre.

Avec quel effets indirects pervers ? L'analyse des facettes cachées ci-dessus a mis toute une série d'effets pervers potentiels : accentuation de la compétitivité plutôt que de la solidarité ; uniformisation vers une culture unique plutôt que l'exploitation de la diversité culturelle ; développement du mesurable au détriment du complexe ; des « ressources » au détriment des « compétences » ; du quantitatif au détriment du qualitatif... Mais on pourrait aussi mettre en évidence des effets pervers concrets plus subtils (voir la dernière enquête internationale de Mons, 2009), comme l'accent (et donc la pression) mis sur certaines disciplines (la langue, les mathématiques, les sciences) et la tendance à négliger d'autres disciplines (les « petites » disciplines comme elles sont souvent appelées) qui ont cependant une fonction d'éveil très importante pour l'esprit, le corps et le cœur.

Pour quel développement personnel privilégié ?

Compte tenu des analyses précédentes, on peut se demander si la pression des classements ne contribue pas davantage à un développement personnel centré sur « l'avoir », c'est-à-dire la capitalisation des ressources édictées par les standards internationaux. Plus la capitalisation sera élevée et plus la personne sera compétitive, plus elle aura des chances d'accéder aux bénéfices du village global (accès aux fonctions les plus recherchées, prestige, mobilité, privilège du choix...).

Pour quel développement personnel non privilégié ?

Si la pression des classements tend à privilégier l'avoir, c'est sans doute au détriment d'un développement centré sur « l'être ». Être une personne, c'est développer son identité propre, donc unique, riche de son unicité. Mais on sait aussi que le développement identitaire se fait à travers la rencontre avec ses parents et proches d'abord, avec son groupe social d'appartenance ensuite, grâce à son enracinement culturel mais aussi la rencontre d'autres cultures. Ce développement identitaire est d'autant plus riche qu'il se déroule à travers de « vraies rencontres » et non dans des environnements à la pensée unique et dépersonnalisant par des processus d'anonymisation. La pression de classements, si l'on y prend garde, peut conduire à de tels processus.

Pour privilégier quelle société ? De nombreux écrits (Zgaga, 2009) portent sur les phénomènes de globalisation de la société et l'émergence d'un « village global ». Les analyses ci-dessus montrent que la pression des classements y contribue, non seulement à travers leurs effets que l'on peut qualifier de positifs mais aussi à travers leurs effets pervers.

Pour retarder l'émergence de quelle société ? On sait que les sociétés ne sont pas stables et évoluent à travers les siècles par des processus kantieniens de balancements et de réactions à l'existant. Des analystes comme Bernard Charlot (RIE, N°52) mettent en évidence que les pressions de la globalisation induites par les évaluations et les classements internationaux contribuent à une différenciation interne des systèmes éducatifs, entre un réseau d'éducation défini par des standards globaux pour les plus riches et, pour les plus pauvres, des écoles fondées sur des valeurs nationales ou locales. On peut se demander si cette situation pourra encore tenir longtemps et si des mouvements forts dans leurs engagements (comme l'Histoire du monde en est riche) ne vont pas contribuer à faire émerger une société plus respectueuse de la diversité

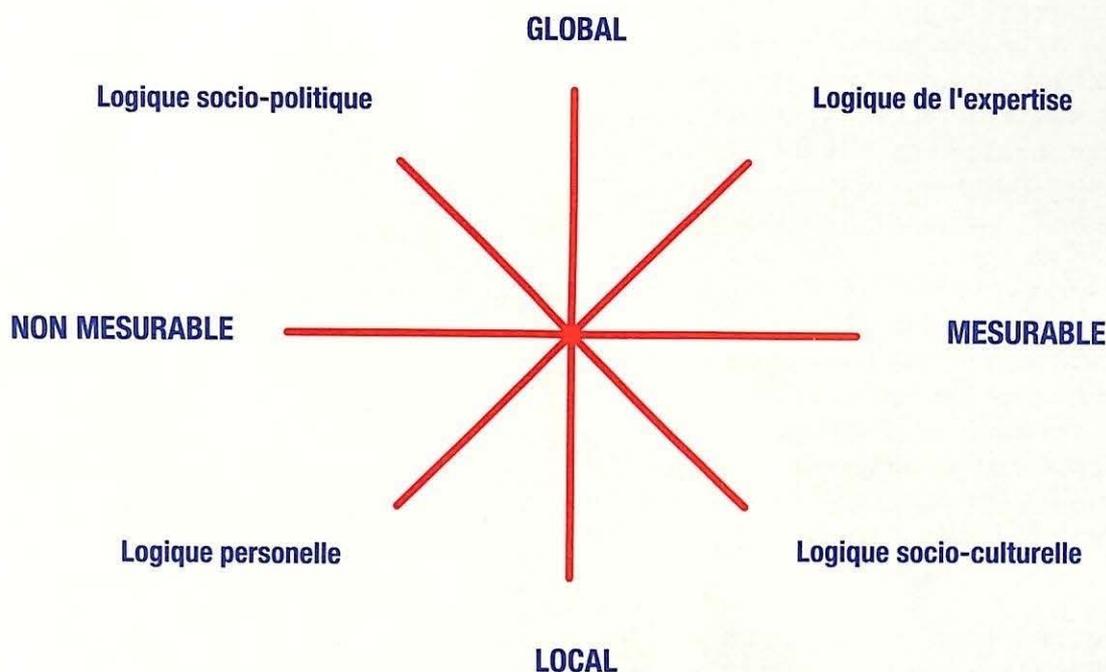
et de la richesse qu'elle représente, moins anonyme, et où le « penser global » et « l'agir local » se rencontrent et s'allient. Les systèmes éducatifs s'investiront-ils dans ce sens ?

2. Une tentative de modélisation des tensions

La prise de conscience que derrière les faces visibles existent nécessairement des faces cachées, que « la présence de » implique « l'absence de », que le « au profit de » sous-tend « au détriment de », que « privilégier telle chose » entraîne « négliger ou mettre en veilleuse telle autre chose »..., cette prise de conscience amène des tensions, surtout lorsque des acteurs différents n'ont pas les mêmes enjeux. Comment modéliser de telles tensions ?

À partir des analyses émises ci-dessus, il nous semble qu'il faille prendre plusieurs catégories de paramètres. Tout d'abord, il nous paraît important de croiser deux dimensions bipolaires : le « penser et l'agir global » versus « le penser et l'agir local » ; le « mesurable » versus le « non mesurable ». Le croisement opéré permet de générer les quadrants

MODÉLISATION DES TENSIONS DANS LES SYSTEMES ÉDUCATIFS PRIMAIRES ET SECONDAIRES



dans lesquels on peut situer quatre logiques en œuvre et, associés à chacune d'entre elles, des acteurs clés et des valeurs prioritaires. Le graphique suivant permet de visualiser les diverses tensions générées par les évaluations internationales et les classements.

Le premier quadrant (en haut à droite) permet de comprendre ce qui sous-tend le développement des évaluations internationales et des classements qui en découlent. Ceux-ci sont faits par des experts, avec une logique d'experts, avec un arsenal technique très sophistiqué (ex. : le modèle TRI, un des modèles mathématiques les plus avancés dans le domaine de la mesure), avec la recherche d'un accord sur des standards évaluables et mesurables, permettant aux gestionnaires politiques et des systèmes éducatifs de se situer par rapport aux autres pays et de prendre des décisions pour s'élever dans la compétition du prochain palmarès. Ainsi est créée une tension stimulante censée servir le développement socioéconomique et le processus de globalisation, grâce au tableau de bord offert par des indicateurs mesurables et supports de la prise de décision.

Le deuxième quadrant (en bas à droite) illustre bien un des effets d'une autre catégorie d'acteurs aux enjeux différents. Il s'agit d'acteurs locaux, c'est-à-dire des consommateurs vivant dans un contexte bien défini, qui cherchent à s'informer pour mieux consommer les « produits » ou les « bénéfiques supposés » des systèmes éducatifs. Ceux-ci contribuent à créer des « quasi-marchés » : les tableaux de bord offerts par les classements génèrent des processus de comparaison et entraînent au niveau local une recherche des meilleurs établissements par les consommateurs sur la base d'indices subjectifs liés à la réputation, puisque les tableaux de bord ne fournissent pas d'indicateurs objectifs locaux. Parmi ces « consommateurs », on trouve avant tout les parents qui ont la possibilité et la volonté de s'informer, appartenant pour la plupart aux milieux les plus favorisés. Dans les pays du Nord, cela se traduit pour ceux-ci par la recherche des établissements les plus réputés et par des stratégies pour contourner les dispositifs mis en place pour assurer une mixité sociale. Dans les pays du Sud, cela se traduit le plus souvent, pour l'enseignement de base, par l'inscription dans les écoles privées de renom même si elles sont éloignées du domicile en utilisant les possibilités offertes par des membres de la famille élargie ; les plus puissants s'offrent même souvent des établissements étrangers en profitant des possibilités offertes par la diaspora ou par leurs relations.

Le troisième quadrant (en bas à gauche) est assez représentatif des tensions vécues par des acteurs dont le pouvoir est finalement peu important par rapport aux gestionnaires et aux consommateurs les plus puissants. Parmi ceux-ci, les enseignants qui vivent souvent mal les prescriptions des gestionnaires, les standards qu'ils jugent insuffisants ou inadéquats dans leur contexte, les accusations générées par les classements ne correspondant pas aux attentes, la non prise en considération des difficultés liées au contexte, l'absence de balises concrètes, l'insuffisance des moyens... et qui revendiquent au contraire être les seuls à bien connaître le terrain et avoir une praxis que n'ont pas ceux qui décident (« qu'ils viennent ne fût-ce qu'une journée à notre place ! »). Parmi ceux-ci aussi, les élèves dont le souci principal est de trouver une place dans le milieu scolaire ne correspondant pas nécessairement à un vrai choix personnel et qui vont développer des logiques personnelles très pragmatiques, face aux pairs, aux enseignants et aux parents (plus ou moins présents ou absents selon les cas).

Le quatrième quadrant (en haut à gauche) est assez représentatif d'une logique que nous pourrions appeler sociopolitique, car tournée vers les problèmes de société dans toute leur complexité et orientée vers des valeurs très marquées éthiquement, comme la justice sociale, le développement humain durable, le respect des cultures et des différences. En ce sens, l'IDH est assez représentatif de cette logique, comme le sont aussi les défenseurs des « biens communs », comme l'eau et l'éducation (Petrella, in RIE, n°52). Dans les PVD, l'action de l'UNICEF pour intégrer dans les curriculums les « compétences de vie » correspond bien à cette logique. Comme le sont tout autant les actions de ceux qui prônent l'importance du rôle éducatif de l'école à travers le contenu de certaines disciplines. Dans tous les cas, il s'agit d'une vision à plus long terme axée sur le qualitatif, l'humain et une société plus juste.

Que peut-on conclure de cette tentative de modélisation des tensions générés par les effets visibles ou cachés générés par la pression des évaluations internationales et des classements ? Le fonctionnement du modèle est susceptible de rendre compte d'une très grande variabilité de situations actuelles ou à venir. Deux situations extrêmes, entre lesquelles de nombreux cas intermédiaires, peuvent être décrits et méritent réflexion.

La première correspond à une exaspération des tensions : les acteurs locaux rentrent de plus en plus en conflit avec une noosphère anonyme de

gestionnaires et d'experts considérés comme coupés de la réalité et se retranchant derrière un arsenal technique sophistiqué ; les logiques personnelles des enseignants et des élèves rentrent donc en conflit direct avec les logiques de l'expertise, ou pire finissent par ignorer dédaigneusement celles-ci ; la praxis, les savoirs d'expérience et les évaluations internes, très marqués par la contextualisation, rentrent en conflit ouvert avec l'expertise décontextualisée de la noosphère, leurs standards, leurs évaluations externes et leurs indices trop exclusivement quantitatifs. Ce premier type de tensions se trouve encore accentué par une autre source de tensions : les logiques socio-politiques, porteuses de valeurs centrées sur le bien commun et sur plus de justice, rentrent en conflit ouvert avec tous ceux (l'OMC en particulier) qui contribuent à faire de l'éducation de base un bien privé comme tout autre bien de consommation et du système éducatif un marché dit ouvert, c'est-à-dire fermé aux catégories sociales défavorisées. Le paroxysme serait atteint si se construisaient des réseaux de pays ou d'acteurs à l'intérieur des pays qui, forts de leur logique propre, se refermeraient sur eux-mêmes et établiraient des murailles symboliques les protégeant de toute contamination par les autres réseaux.

À l'autre extrême correspond une situation où les axes (le global et le local ; le mesurable et le non mesurable) ainsi que les logiques (de l'expertise et personnelle ; sociopolitique et socioculturelle) ne fonctionneraient pas en termes d'exclusion (ou bien ceci...ou bien cela) mais d'inclusion (et ceci...et cela). Ceci implique de la part des différentes catégories d'acteurs une posture de reconnaissance de l'autre au sens défini par Ricoeur (2004) : reconnaître l'autre, c'est d'abord connaître l'autre et donc le situer dans son contexte, dans son statut, dans ses enjeux ; reconnaître l'autre, c'est ensuite lui reconnaître une valeur (son expertise ou sa praxis ou son engagement pour une cause ou son souci pour l'avenir de ses enfants) ; reconnaître l'autre, c'est enfin être reconnaissant vis-à-vis de l'autre, car chacun apporte ou peut apporter quelque chose d'important, et même d'unique, au développement tant personnel que sociétal. Dans cette situation extrême et utopique, les logiques ne s'opposent plus mais se complètent et s'intègrent. Auguste Comte n'attribuait-il pas à l'utopie un rôle important pour le développement, tant des personnes que des institutions et même pour le développement scientifique ?

BIBLIOGRAPHIE

1^o PARTIE

ANLCI (2006). Illettrisme : des chiffres pour éclairer les décisions. Lyon : ANLCI.

Baye, A., Demonty, I., Fagnant, A., Lafontaine, D., Matoul, A. & Monseur, Ch. (2004). Les compétences des jeunes de 15 ans en Communauté française de Belgique en mathématiques, en lecture et en sciences. Les Cahiers du Service de Pédagogie Expérimentale, n°19 et 20, 1-224.

BIE (2005). Progrès de la communauté de pratique dans le développement curriculaire. Le débat sur les compétences. Information et Innovation en Éducation, N°121, 1-12.

Bottani, N. & Vrignaud, P. (2005). La France et les évaluations internationales. Paris : Haut Conseil de l'Évaluation de l'École, rapport n° 16.

Bulmahn, E. (2002). PISA : les conséquences pour l'Allemagne. L'observateur OCDE, juin, 33-34.

De Ketele, J.-M. (2005). L'évolution des systèmes éducatifs européens. Montréal : Hommage en l'honneur du Rapport Parent.

De Ketele, J.-M. & Gérard, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par compétences, Mesure et Evaluation en Education, vol. 28, n°3, 1-26.

Demaillly, L. (2006a). En Europe : l'évaluation contre la crise des systèmes scolaires, l'évaluation en crise. Education et Sociétés, n°17, 105-120.

Demaillly, L. (2006b). La nécessaire méta-évaluation. In : G. Figari & L. Mottier-Lopez (Eds). Recherche sur l'évaluation. Paris : L'Harmattan. 203-210.

Dupriez, V. & Vandenberghe, V. (2004). L'école en Communauté française de Belgique : de quelle inégalité parlons-nous. Les Cahiers de Recherche du GIRSEF, n°27, 3-28.

Dupriez, V. & Dumay, X. (2005). L'égalité des chances à l'école : analyse d'un effet spécifique de la structure scolaire. Les Cahiers de Recherche du GIRSEF, n°45, 1-17.

Dupriez, V. & Dumay, X. (2006). Elèves en difficulté d'apprentissage : parcours et environnements éducatifs différenciés en fonction des structures scolaires. Les Cahiers de Recherche du GIRSEF, n°51, 3-24.

- Fitoussi, J.-P. (2001).** L'enseignement supérieur des sciences économiques en question. Paris : Fayard.
- Fowler, F. & Poetter, Th. (2006).** Comment et pourquoi les Français réussissent en mathématiques : leur politique, leurs programmes et leur pédagogie. Education et Sociétés, n°17, 121-120-139.
- Gérard, F.-M. & Roegiers, X. (2003).** Des manuels scolaires pour apprendre. Bruxelles : Editions De Boeck.
- Le Soir (2006).** Sommes-nous si nuls en maths ? Supplément au Journal Le Soir, 24 août 2006, 48P.
- Mac Gaw, B. (2004). Vers une éducation de qualité. Peut-on toujours aller plus loin. L'observateur OCDE, site Web OCDE.
- Mac Kenzie, Ph. (2004).** Enseignement : redorer le blason. L'observateur OCDE, mai, 11-12.
- Mons, N. (2004).** De l'école unifiée aux écoles plurielles. Evaluation internationale des politiques de différenciation et de diversification de l'offre éducative. Dijon : Université de Bourgogne, thèse non publiée.
- PISA (2006).** Les conséquences pour l'Allemagne. <http://www.observateurocde.org/news/fullstory.php/aid/424>, consulté le 23 août 2006.
- Pourtois, H. (2005).** Pertinence et limites du principe d'égalité des chances en matière d'éducation scolaire. Louvain-la-Neuve : document inédit présenté lors d'un séminaire du GIRSEF le 19 janvier 2005.
- Vandenbergh, V. & Zachary, M.-D. (2000). Efficacité-équité dans l'enseignement secondaire de la Communauté Wallonie-Bruxelles : essai d'évaluation dans le cadre d'une comparaison internationale. Les Cahiers de Recherche du GIRSEF, n°7, 3-17.
- Vandenbergh, V. (200 ?).** Evaluating the Magnitude and the Stakes of Peer Effects analysing Science and Math Achievement across OECD.
- Vandenbergh, V. (2002).** Is There an Effectiveness-Equity Trade-off in Education ? A cross-country comparison using TIMSS test scores.
- Vandenbergh, V. (2003).** Private, Private Government-Dependent and Public schools. An International Effectiveness Analysis. Louvain-la-Neuve : IRES-ECON-UCL.
- Vandenbergh, V. (2004).** Achievement Effectiveness and equity. The role of Tracking, Grade Repetition and Inter-school Segregation. Les Cahiers de Recherche du GIRSEF, n°28, 3-28.
- Vandenbergh, V. (2005).** Educational Equity. Social interactions might matter. In : W. Hutmacher, D. Cochrane & N. Bottani : In Pursuit of Equity in Education. Using International Indicators to Compare Equity Policies. Dordrecht/Boston/London : Kluwer Academic Publishers, chap. 11, 267-275.
- Vandenbergh, V. & Robin, S., M.-D. (2006).** Evaluating the Effectiveness of Private Education Across Countries : a Comparison of Methods, Labour Economics.
- Zachary, M.-D., Dupriez, V. & Vandenbergh, V. (200 ?).** Comparaison internationale de l'efficacité et de l'équité des systèmes d'enseignement. In D. de la Croix, F. Docquier, Ch. Mainguet, S. Perelman & E. Wasmer (Ed.) : Capital humain et dualisme sur le marché du travail. Bruxelles : De Boeck, Chapitre 13, 259-277.

2° PARTIE

- BEHRENS M. (2009)** : « Un laboratoire de formation à petite échelle. Le cas de la Suisse », Revue internationale d'éducation de Sèvres [www.ciep.fr/ries/colloque-2009/contributions.php/atelier-B.php].
- CHARLOT B. (2009)** : « Convergence internationale et diversification interne des modèles scolaires », Revue internationale d'éducation de Sèvres, N°52, 129-138.
- CIEP (2009)** : « Dossier : Un seul monde, une seule école », Revue internationale d'éducation de Sèvres, n°52.
- DE KETELE J.-M. (2006)** : « Mais qu'évaluent donc les enquêtes internationales ? », Revue internationale d'éducation de Sèvres, N°43, 141-157.
- DE KETELE J.-M. & GERARD F.-M. (2005).** « La validation des épreuves d'évaluation selon l'approche par compétences », Mesure et évaluation en éducation, Vol.28, N°3, 1-26.
- DUMAY X. & DUPRIEZ V. (2009)** : L'efficacité dans l'enseignement. Bruxelles : De Boeck.
- GAUTHIER R.-F. & ROBINE F. (2009)** : « Contenus et valeurs. Bigarrure du monde, convergence de questions », Revue internationale d'éducation de Sèvres, N°52, 77-84.
- MONS N. (2009).** Effets théoriques et réels des politiques d'évaluation standardisée, Revue française de pédagogie, n°169, 99-140.
- PETRELLA R. (2009)** : « Eau et éducation : des biens communs », Revue internationale d'éducation de Sèvres, N°52, 123-128.
- RICOEUR P. (2004)** : « Parcours de la reconnaissance ». Paris : Seuil.
- ROCHER T. (2009)** : « Nécessité et complémentarité des évaluations internationales et nationales. Le cas de la France. », Revue internationale d'éducation de Sèvres [www.ciep.fr/ries/colloque-2009/contributions.php/atelier-B.php].
- ZGAGA P. (2009)** : « Éducation : du cosmopolitisme au globalisme », Revue internationale d'éducation de Sèvres, N°52, 39-50.