

PISA : politique, problèmes fondamentaux et résultats paradoxaux

PISA: politics, fundamental problems and intriguing results

Svein Sjøberg

Traducteur : Faouzia Kalali et Edgar W. Jenkins



Édition électronique

URL : <https://journals.openedition.org/ree/5855>

DOI : 10.4000/ree.5855

ISSN : 1954-3077

Éditeur

Université de Nantes

Référence électronique

Svein Sjøberg, « PISA : politique, problèmes fondamentaux et résultats paradoxaux », *Recherches en éducation* [En ligne], 14 | 2012, mis en ligne le 01 septembre 2012, consulté le 30 mai 2021. URL : <http://journals.openedition.org/ree/5855> ; DOI : <https://doi.org/10.4000/ree.5855>



Recherches en éducation est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

PISA : politique, problèmes fondamentaux et résultats paradoxaux

Svein Sjøberg¹

Traduction de l'anglais :
Faouzia Kalali et Edgar W. Jenkins

Résumé

Durant la dernière décennie, PISA a renforcé fortement son influence sur les politiques éducatives dans les pays participants. Les résultats de ces enquêtes ouvrent bien des débats publics sur la qualité de l'enseignement scolaire. En effet, les résultats concernant les tableaux de scores et les classements des pays, bien que souvent pris dans leur valeur nominale, sont exploités tant par les médias que par les politiques et les décideurs. Cet article fait état de certaines réserves quant à la validité de ces scores et de ces classements que nous exposerons plus loin. Mais il y a, tout d'abord, lieu de soulever certaines questions fondamentales sur les raisons du lancement du projet PISA par l'OCDE. Des raisons qui sont, bien évidemment, fortement liées aux objectifs politiques d'ensemble de l'OCDE et à son engagement sous-jacent dans une économie de marché mondiale et concurrentielle. Pour étayer ce propos, des exemples seront donnés pour expliquer l'influence qu'exerce PISA sur les politiques éducatives. Nous commencerons par émettre deux catégories de critiques qui seront discutées. La première a trait au projet PISA en tant que tel. Nous montrerons qu'il est impossible de construire d'une manière objective une enquête qui peut être utilisée pour des pays et des cultures différents. De même, l'exclusion du local et du spécifique va à l'encontre des préconisations de l'UNESCO, de celles des éducateurs et des programmes scolaires nationaux relatifs à l'éducation scientifique. La deuxième catégorie de critiques se rapporte à certains résultats paradoxaux. Il semble que les élèves dans les pays à hauts scores développent les attitudes les plus négatives envers la science. Il ressort également que les scores sont indépendants de certains facteurs comme le niveau des ressources pédagogiques, la taille des classes, etc. Enfin, les résultats sont corrélés négativement à l'utilisation des méthodes pédagogiques actives, des démarches d'investigation et des TIC. En tous les cas que l'on croie ou non au bien fondé du projet PISA, de tels résultats aussi surprenants soient-ils méritent d'être discutés.

Les enquêtes PISA ont permis à l'éducation de prendre une place importante à la fois dans les médias et dans les débats politiques, tant à l'échelle nationale qu'internationale. Si ses résultats semblent être acceptés dans leur valeur nominale, c'est surtout en raison du peu de critiques dont ils ont fait l'objet. L'objectif de cet article portera donc sur les aspects problématiques des enquêtes PISA. Depuis la première publication en 2001, basée sur le test effectué en 2000, les résultats sont devenus une sorte de « gold standard » de la qualité de l'éducation. Bien que l'impact politique et éducatif varie d'un pays à l'autre, les résultats préparent souvent le terrain à des débats publics sur la qualité de l'enseignement scolaire. Les tableaux des scores sont largement repris dans les médias, ils sont exploités également par les politiques et les responsables dans le domaine de l'éducation. Ces résultats semblent fonctionner comme une sorte de test de « QI » des systèmes scolaires. Une question aussi complexe que la question éducative est ainsi réduite à un simple ensemble de chiffres qui peuvent être ordonnés avec une grande précision. Mais, comme pour les résultats des tests de QI, il y a de sérieuses réserves quant à la validité des scores.

¹ Professeur en didactique des sciences à l'université d'Oslo (Norvège).

1. Que mesure PISA ?

L'image qui ressort est, à bien des égards, confuse. PISA prétend ne pas mesurer les connaissances scolaires ou les compétences acquises à l'école, l'ambition proclamée est d'évaluer réellement la qualité de l'éducation.

La page d'accueil du Programme international introduit PISA en ces termes : *«Les étudiants sont-ils bien préparés pour les futurs défis ? Peuvent-ils analyser, raisonner et communiquer efficacement ? Ont-ils la capacité de continuer à apprendre tout au long de la vie ? Le Programme International pour le suivi des acquis des élèves (PISA) de l'OCDE répond à ces questions et bien plus encore, grâce à ses enquêtes concernant les élèves de 15 ans dans les principaux pays industrialisés. Tous les trois ans, il évalue dans quelle mesure les élèves en fin de scolarité obligatoire ont acquis des connaissances et des compétences essentielles pour une pleine participation dans la société.»* (Citation complète sur <http://www.pisa.oecd.org>, consulté le 31 janvier 2012).

On peut difficilement s'opposer à une telle ambition. Toutefois, cette dernière est loin d'être réaliste quand on prétend répondre à une multitude de questions. Notre jugement ne devrait pas porter sur les ambitions affichées, mais sur ce que l'on évalue réellement. Dans la déclaration ci-dessus et dans d'autres rapports, l'OCDE ne prétend pas que les résultats du PISA mesurent la qualité des systèmes scolaires, néanmoins ils fournissent de multiples données en lien avec l'école, l'environnement familial et social. Le rapport principal du PISA 2006 l'exprime clairement : *«Il faut se garder de conclure d'emblée que le système d'éducation ou les établissements sont plus efficaces dans un pays que dans un autre sur la seule base d'un score significativement supérieur sur l'échelle de culture mathématique, de compréhension de l'écrit ou de culture scientifique. En revanche, il est légitime d'en déduire que l'impact cumulé des expériences d'apprentissage vécues entre la prime enfance et l'âge de 15 ans à la maison et à l'école y a généré de meilleurs résultats dans les domaines d'évaluation de l'enquête PISA.»* (OCDE, 2007a, p.375-376)

Ces rapports mettent l'accent sur le fait que les évaluations ne se réfèrent pas aux programmes d'études nationaux, mais prennent appui sur les définitions et le cadre théorique élaborés par les experts de l'OCDE impliqués dans PISA (OCDE, 2006). Le rapport technique (OCDE, 2009) indique clairement que les connaissances et les compétences testées ne sont pas *« définies en termes d'un dénominateur commun des programmes scolaires nationaux, mais en termes de compétences jugées essentielles pour la vie future »*. Il précise également que sont exclues les questions relatives aux programmes et aux contenus portant directement sur la science étudiée en classe.

Bien que PISA affirme ne pas évaluer les connaissances scolaires en fonction des programmes d'enseignement nationaux, les résultats sont interprétés, y compris dans les rapports de l'OCDE, comme des mesures fiables de la qualité des systèmes scolaires nationaux. Ils donnent même lieu à des recommandations de nature politique concernant les écoles (Loveless, 2009).

2. La politique du programme PISA

L'OCDE est l'Organisation des pays hautement industrialisés et développés sur le plan économique, son mandat se reflète dans son nom : Organisation de coopération et de développement économiques. Son but

est avant tout de promouvoir des politiques et de fournir des normes pour le développement économique dans une économie mondiale de libre marché. On devrait se rappeler que l'OCDE représente l'économie, non l'éducation. Cette dernière est certainement un atout majeur dans le développement de la compétitivité nationale, c'est pourquoi elle devient une préoccupation essentielle dans les conseils stratégiques de l'OCDE. Le mandat de l'OCDE explique l'intérêt porté aux domaines qu'évalue PISA : la lecture, les mathématiques et les sciences. Ces

domaines sont considérés comme des éléments clés qui sous-tendent la compétitivité dans une économie mondiale dominée par la science et le développement technologique. Mais derrière ce choix, on envoie un message implicite de ce qui est considéré comme important pour le développement des jeunes dans les écoles. Reste à signaler les domaines qui n'entrent pas dans la mesure de la qualité des écoles : sciences humaines, sciences sociales, langues étrangères, histoire, géographie, éducation physique, etc. Notons également que PISA n'aborde pas les aspects fondamentaux des finalités éducatives de beaucoup de pays, tels que l'équité, l'empathie, la solidarité, la curiosité et l'engagement de préserver l'environnement, etc. Les déclarations sur ces missions assignées au système scolaire (juridiquement contraignantes) sont souvent oubliées ou ignorées au cours de discussions publiques et de débats politiques.

L'OCDE est généralement très claire sur le but économique du PISA et la nature concurrentielle internationale de ses classements : « *dans une économie mondialisée, évaluer les progrès des systèmes d'éducation à l'aune de normes nationales ne suffit plus, il faut aussi les mesurer sous une perspective internationale. L'OCDE a relevé ce défi avec la mise en œuvre du Programme international pour le suivi des acquis des élèves (PISA), qui évalue la qualité, l'équité et l'efficacité des systèmes d'éducation dans quelque 70 pays qui, ensemble, représentent neuf dixièmes de l'économie mondiale* » (OCDE, 2010a, avant-propos).

Il semble qu'il y ait une contradiction. D'une part, on déclare que l'on ne cherche pas à évaluer des programmes scolaires, ou encore des connaissances acquises à l'école. D'autre part, on prétend que les tests évaluent « *la qualité, l'équité et l'efficacité des systèmes d'éducation* ». Il est également intéressant de noter que l'importance du PISA est définie par la part de l'économie mondiale représentée (90%), et non en termes de la population mondiale. Le point de vue de la concurrence économique est également mis en avant lorsque les résultats du PISA sont présentés au public. Le 4 décembre 2007, lors de la Conférence de publication des évaluations de 2006 à Washington, on a pu lire : « *Perdre notre avantage : les étudiants américains sont-ils mal préparés pour l'économie mondiale ? Les leçons tirées des résultats de PISA... peuvent et doivent être utilisés pour orienter la politique éducative américaine afin que nos étudiants diplômés... soient prêts à soutenir la concurrence, à prospérer, et porter l'économie globale du 21^{ème} siècle* ». (Extrait, pour un aperçu complet et les vidéos voir sur <http://www.all4ed.org/events/losingedge>).

L'usage politique, économique et normatif du PISA est très clair. L'OCDE édite régulièrement des rapports économiques dans de nombreux pays et fournit des conseils sur la politique à mener. La Norvège, mon pays, en est un exemple. Dans le rapport adressé à la Norvège en 2008, on peut lire l'avis donné par les experts de l'OCDE, juste avant la crise financière : le pays devrait augmenter les différences de salaires, réduire les dépenses publiques, augmenter le taux de chômage, réduire le niveau de paiement des congés maladie et réduire les pensions pour handicapés (OCDE, 2008).

Ce rapport, soumis par l'OCDE à la Norvège (2008), avait le système éducatif comme point de mire. En s'appuyant sur les données recueillies lors des évaluations PISA, les conseils de l'OCDE sur le plan éducatif suggèrent : les écoles norvégiennes peuvent s'améliorer par la fermeture des petites écoles, l'augmentation de la taille des classes, le recours à davantage de tests d'évaluation, la publication des résultats par école (par enseignant) et la révision du salaire des enseignants au vu des résultats des tests. Le rapport se termine par un avertissement clair : « *une augmentation des dépenses concernant les écoles n'aura aucun effet* » (OCDE, 2008).

En fait, ce que prône cet « avis d'expert » est clair : la Norvège devrait devenir un pays différent. Ces conseils ne sont ni objectifs, ni neutres, ni « scientifiques ». Car les meilleures écoles sont définies comme celles qui sont les « plus rentables ».

3. Politiques nationales basées sur PISA : l'exemple de la Norvège

L'attention portée aux résultats des évaluations PISA dans les médias nationaux varie selon les pays, mais dans la plupart d'entre eux elle demeure importante.

En Norvège, les résultats des évaluations PISA 2000 et 2003 ont soulevé des polémiques dans la plupart des journaux nationaux. Notre ministre de l'éducation de l'époque (2001-2005), Kristin Clemet (représentant le Høyre, Parti conservateur) a commenté dans ces termes les résultats PISA 2000, sortis quelques mois après qu'elle eut pris ses fonctions, à la suite d'un gouvernement travailliste : « *La Norvège est une école perdante, maintenant on est bien informé. C'est comme revenir des Jeux olympiques d'hiver sans une médaille d'or* ». Elle a même ajouté : « *Et cette fois, nous ne pouvons même pas prétendre que les participants finlandais ont été dopés* » ! (Ce qui était le cas dans le récent Championnat) (Aftenposten, janvier 2001).

L'image d'une école perdante est relayée à travers les titres des journaux. Mais la réalité est différente. La Norvège était proche de la moyenne de l'OCDE dans les trois domaines des tests PISA 2000 et 2003². Les résultats de PISA (et TIMSS) conduisent à dresser un portrait déformé du système éducatif norvégien, allant au-delà des aspects qui sont évalués. C'est devenu une « sagesse populaire » de dire que les écoles norvégiennes sont généralement de faible qualité, et que les classes norvégiennes sont parmi les plus bruyantes dans le monde (bien que cet aspect n'ait jamais été étudié). Le débat public, peu enclin à la critique, ainsi que le peu de recherches effectuées pour rappeler les limites de telles enquêtes, entraînent une adhésion des décideurs politiques et du public à ce constat d'échec. En somme, PISA a créé une image publique non avérée de la qualité du système scolaire norvégien. En plus de cette image, qui peut être considérée comme nuisible, les enquêtes effectuées auprès des enseignants font de ces tests une des causes principales des difficultés de leur travail au quotidien.

Mais PISA n'a pas seulement façonné l'image publique des écoles norvégiennes, il a également servi à légitimer les réformes dans le domaine de l'éducation. Sous Clemet, (2001-2005), une série de réformes éducatives ont été introduites en Norvège. Elles ont été généralement légitimées en se référant aux résultats des évaluations internationales, principalement PISA. En 2005, période du changement de gouvernement, Helge Ole Bergesen, le secrétaire d'État de Clemet, a publié un livre dans lequel il a présenté l'« Inside story » des réformes introduites alors qu'il était au pouvoir. Les principales caractéristiques de l'ouvrage sont les nombreuses références à des études d'évaluation à grande échelle. L'auteur déclare explicitement que ces études ont fourni les principaux arguments et constitué la base des réformes scolaires et parascolaires. Sous la rubrique *Le choc du PISA*, il a confirmé le rôle clé de ce dernier : « *Avec [la publication] des résultats du PISA, le décor a été planté pour engager une bataille nationale sur les connaissances dans nos écoles. [...] Pour ceux d'entre nous qui venaient d'assurer le pouvoir politique dans le ministère de l'éducation et la recherche, les résultats du PISA ont fourni un "flying start"* » (Bergesen, 2006, p.41-42, traduction de l'auteur).

Plus tard, lors de la campagne des élections législatives de 2009, le premier ministre, candidat pour le même parti (Høyre - conservateur/modéré), adressait le message suivant, avec même une signature personnelle dans les journaux : « *Je, Erna Solberg, garantis que si nous formons le gouvernement après les élections, nous pouvons promettre plus de points en ce qui concerne PISA* » (Aftenposten, 27 mars 2009, traduction de l'auteur).

² Sans raison particulière, les norvégiens avaient pensé et prévu d'être les premiers au classement – comme ils le sont souvent selon d'autres indicateurs, par exemple l'indice de développement humain du PNUD (PNUD, 1990-2012) et les sports d'hiver. Lorsque nous ne sommes pas les premiers, nous nous considérons comme étant des perdants.

Il est intéressant de voir que cette déclaration a été faite peu de temps après que le parlement eut adopté à l'unanimité de nouvelles lois sur les valeurs fondamentales des écoles norvégiennes, à un moment où les nouveaux programmes d'enseignement nationaux étaient dans une phase d'élaboration. Un nouveau livre blanc sur les établissements d'enseignement secondaire inférieurs (KD 2011) a été présenté au Parlement en 2011. Les mots clés de cette loi nouvellement adoptée sont à peine mentionnés, alors que sont cités à chaque page les évaluations PISA et les experts de l'OCDE. D'autres exemples montrent bien comment PISA redéfinit les programmes d'enseignement nationaux et va bien au-delà des buts assignés à l'école et des finalités de l'éducation, tels qu'ils ont été définis à l'échelle nationale. Cette ligne de conduite transcende les clivages politiques. Des exemples similaires pourraient être donnés pour le Danemark et la Suède.

4. PISA entre économie de marché et globalisation

Jusqu'ici le point de vue défendu dans cet article est que PISA devrait être vu et compris dans un contexte politique plus large. Les deux éléments clés sont les idées de l'économie de marché (*market thinking*) et la globalisation.

Le programme PISA peut être considéré comme faisant partie d'une tendance politique internationale actuelle, où les concepts et les idées de l'économie de marché sont déployés dans le secteur de l'éducation. Le *nouveau management public* est désigné pour décrire ce marché inspiré par une philosophie censée rendre le secteur public plus efficace. Des termes comme qualité, efficacité, transparence, accountability et « valeur de l'argent » sont utilisés (souvent avec une charge positive) pour des réformes de nature politique, dans de nombreux secteurs publics tels que les enseignements primaire, secondaire et supérieur, la culture, la santé et les soins. Quant aux services publics traditionnels, ils sont soumis à des offres où ils sont en concurrence avec les fournisseurs privés. Cette tendance semble caractériser le développement dans plusieurs pays.

L'autre perspective politique et économique est celle de la mondialisation, où les nations et les grandes multinationales se font concurrence, dans un marché commun de plus en plus mondialisé. Le besoin d'un marché du travail souple et mobile requiert en matière d'éducation des normes collectives, des systèmes communs pour les examens, les diplômes et les qualifications. Ces tendances opèrent au sein des organisations supranationales comme l'Union Européenne, comme ce fut le cas lors du processus de Bologne. L'OCDE joue un rôle clé dans la promotion de normes communes, d'indicateurs et d'évaluations. PISA n'affiche pas ces aspects politiques et économiques comme des objectifs explicites. Néanmoins, les résultats sont fortement utilisés par les experts de l'OCDE pour fournir des « expertises indépendantes » aux pays membres. L'avis donné a un poids politique considérable dans ces pays (et aussi dans plus de 30 autres pays participants au PISA qui ne sont pas membres de l'OCDE).

De par sa « philosophie », PISA propose des idéaux et des modèles pour les pays participants. Il influence également les nouvelles réformes de l'éducation, comme l'a montré ci-dessus l'exemple de la Norvège. Par cette influence, PISA exerce une pression politique afin d'harmoniser les institutions nationales comme les systèmes éducatifs des pays, et de promouvoir la concurrence sur la scène éducative mondiale. Alors que la plupart des éducateurs soutiennent l'idée d'un enseignement axé sur les contextes et les programmes locaux, la pression exercée par PISA pousse dans la direction opposée. Les réformes qui en découlent sont le résultat de l'utilisation de normes mesurables et quantifiables qui servent à des fins quantitatives. C'est l'objectif de PISA, que les chercheurs le veuillent ou non.

5. Problèmes fondamentaux inhérents à PISA

■ *Les indicateurs sont-ils valables pour tous ?*

Le problème fondamental des évaluations PISA est leur ambition à « mesurer » la qualité du système éducatif des pays à l'aide d'indicateurs communs, universels, indépendants des systèmes scolaires, des structures sociales, des traditions, de la culture, des modes de vie et des modes de production. En outre, les évaluations prétendent mesurer « comment la jeune génération est prête à relever les défis du monde de demain ». Une telle ambition suppose que ces défis soient plus ou moins identiques pour les jeunes, quels que soient les pays et les cultures³. Bien que la vie dans de nombreux pays présente des traits probablement similaires, on ne peut guère supposer que les jeunes de 15 ans, par exemple, au Japon, en Grèce, au Mexique et en Norvège, se préparent pour les mêmes défis, et ont besoin de compétences et d'aptitudes identiques. Car il est important de rappeler que PISA évalue une classe d'âge à la fin de ce qui correspond à la scolarité obligatoire dans la plupart des pays. La grande majorité de ces jeunes doit affronter les réalités locales et nationales. Seule une minorité s'insérera dans un marché global et international.

Dans les différents pays, l'éducation est fondée sur des décisions nationales, le plus souvent de la part de gouvernements démocratiquement élus. Les traditions nationales, et les grands débats qui ont lieu, conduisent à des lois concernant aussi bien les finalités générales de l'école que son fonctionnement, comme par exemple les horaires attribués aux matières d'enseignement, les objectifs et les programmes scolaires, la nature des examens... Ces « traditions » sont souvent au cœur de l'identité nationale. L'ensemble de ces lois et règlements représente le mandat qu'une société a donné à ses écoles, à ses enseignants et à tous ceux qui œuvrent pour améliorer la qualité de l'éducation. Même si on affirme n'évaluer aucun système scolaire national, ni aucune école de quelques pays que ce soit, en réalité, il se dégage une sorte d'universalité du curriculum, vraisemblablement exempte de culture décidée par l'OCDE et ses experts.

6. Le test PISA

Des grandes ambitions de l'OCDE aux évaluations PISA, nous assistons à plusieurs étapes dont chacune est ponctuée de sérieux obstacles et nécessite de nombreuses prises de décisions.

La première concerne le choix des domaines de connaissances. En effet, une enquête comme PISA ne peut pas englober tous les domaines possibles des savoirs scolaires. Mais en sélectionnant certains savoirs et en ignorant d'autres, on envoie implicitement un message au grand public et aux politiques concernant ce qui est important à apprendre dans les écoles pour la vie future. Le choix des trois domaines de lecture, des sciences et des mathématiques, reflète l'objectif fondamental de l'OCDE d'œuvrer pour la compétitivité dans une économie Hi-Tech de marché mondiale.

■ *Le cadre de PISA*

La deuxième étape est la réalisation d'un cadre, en réalité d'un programme d'étude où les experts entrent en jeu. Les personnes choisies à cet effet sont reconnues à l'échelle internationale dans leur domaine d'expertise. Des institutions (celles qui remportent le contrat PISA), des experts, des psychométriciens jouent un rôle clé dans l'ensemble du processus. Mais ils sont contraints de travailler en accord avec les grandes lignes qui dessinent les contours du

³ En fait, l'OCDE et PISA restreignent cet énoncé qui doit être valable pour les « pays industrialisés », c'est-à-dire les pays membres de l'OCDE, mais l'influence de PISA est tout aussi forte dans les 30 pays non membres.

programme, de s'exprimer aisément en anglais, la langue dominante dans toutes les délibérations et les documents de travail. Ces cadres élaborés par les groupes d'experts intéresseront probablement les chercheurs, car ils comportent des idées, des sujets et des perspectives de grande importance (voir par exemple OCDE, 2006). Ceux-ci ne sont pas des modèles à suivre, mais ils pourraient servir de source d'inspiration pour construire des programmes d'enseignement nationaux et stimuler le débat sur les priorités éducatives. Le problème est que, loin d'être incitatifs, ces cadres conduisent à une normativité au service d'une logique évaluative.

■ **Sélection des items et élaboration de l'enquête**

Cette étape concerne l'«opérationnalisation» du cadre, c'est-à-dire son utilisation pour l'élaboration de l'enquête dans son ensemble, mais également pour le développement et la sélection des questions (ou items) du test. Il n'est pas question ici d'entrer dans les détails et les subtilités de ce processus complexe, largement bien décrit dans les 400 pages du rapport technique (OCDE, 2009). Cependant, nous relaterons quelques éléments marquants. Chaque pays participant au PISA (pays de l'OCDE seulement) est invité à présenter des items qui sont fondés sur des situations de la vie réelle. Avant que la série finale des questions soit retenue, un processus complexe implique de nombreux organismes et sous-comités. On assiste à des négociations et des débats sur la présélection et sélection des unités, le pilotage national et international, les pré-tests et les tests sur le terrain, et sur l'analyse psychométrique. On peut mesurer la complexité de ce processus tout au long de la lecture du rapport technique concernant en particulier les principes régissant l'exclusion de certains éléments du test. Les principes de telles exclusions sont décrits comme suit : « *Ces analyses [...] incluent des éléments de l'outil standard ConQuest® (item fit, élément de discrimination, difficulté de l'élément, analyse de leurres, capacité moyenne et corrélations par catégorie, etc.), ainsi que des analyses des interactions en fonction du sexe et du pays. Sur la base de ces statistiques essentielles, environ 40 nouveaux éléments ont été retirés du pool des éléments qui seraient examinés pour l'étude principale. (OCDE, 2009, p 41). Les raisons principales d'évaluer des unités en tant que peu adaptées sont le manque de contexte, contexte inadéquat, biais culturel, conformité au programme d'études, à la science scolaire...* » (OCDE, 2009, p. 34).

L'exclusion de tout item non conforme est un argument manifeste d'un point de vue statistique et psychométrique. Il signifie également que dans la réalisation d'une enquête à grande échelle, sans parti pris, sont exclues des questions considérées comme «inappropriées » (controversées dans certains pays), ou ayant un « biais culturel » (positif ou négatif), ou sont proches des programmes scolaires (dans certains pays mais pas dans d'autres). Une fois de plus cela montre clairement que PISA ne mesure pas les acquis de l'école ou les questions liées aux programmes scolaires. Néanmoins, il semble quelque peu étrange que ce test soit utilisé pour juger la qualité de la science *enseignée* à l'école dans chaque pays. Par exemple, dans la forme finale du test PISA, les étudiants norvégiens ne trouveront rien sur les éléments clés de l'économie norvégienne. Ils ne trouveront pas non plus de questions sur les conditions d'exploitation pétrolière sur le plateau continental de l'Arctique, l'aquaculture et l'élevage de poissons, ou les centrales hydroélectriques. Trouveront-ils quoi que ce soit sur les questions d'actualité et les conflits politiques actuels au sujet de la sauvegarde de la nature (exemple de la polémique sur les loups sauvages et les troupeaux de moutons), sur la neige, le ski ou le patinage, sur les aurores boréales (qui étaient la fondation de l'université Norvégienne) ou sur les défis du climat arctique ? Dans d'autres pays, il est tout aussi peu probable pour les élèves de trouver des questions relatives à leurs propres cultures, environnement naturel, historique, ou aux défis nationaux actuels. En réalité, dans le test final, les questions sont décontextualisées. Ce n'est pas une conséquence du cadre de l'enquête, mais une nécessité d'ordre statistique. Cela va à l'encontre de toutes les recommandations des enseignants de science, ainsi que de nombreux programmes d'enseignement nationaux qui promeuvent un contenu scientifique pertinent, intéressant, contextualisé, au moins au niveau de la scolarité obligatoire.

■ **Formulation des questions et problèmes de traduction**

Un autre ensemble de problèmes porte sur les traductions et concerne la langue dans laquelle sont rédigés les textes et les questions. Il s'agit de baser la plupart des unités d'analyse sur des textes assez longs, qui constituent la trame appelée « stimulus ». L'intention est louable. Il s'agit de présenter des textes authentiques et des situations réelles. Mais ce format, en particulier la longueur du texte, constitue une difficulté supplémentaire. PISA diffère dans ce sens de la plupart des enquêtes couramment utilisées en mathématiques et en sciences (exemple de TIMSS). Bien sûr, c'est un choix délibéré des experts qui soulignent que l'on n'évalue pas vraiment les connaissances scolaires. On prétend souvent que de nombreuses unités des évaluations PISA sont, dans une grande mesure, des tests de compétences textuelles plutôt que des tests scientifiques et mathématiques. Les fortes corrélations entre les résultats des essais sur la lecture, les mathématiques et les sciences viennent appuyer cette position, comme cela est souvent mentionné dans plusieurs documents de recherche. Afin d'atténuer cet impact des compétences en lecture, les versions ultérieures des unités d'évaluation du PISA sont devenues plus courtes. Ce qui témoigne que ces critiques sont prises au sérieux.

Un des résultats fondamentaux du PISA, ainsi que d'autres évaluations internationales concernant la lecture, comme PIRLS (International Reading Literacy Study), est que les filles surclassent les garçons en lecture dans tous les pays. Le plus étonnant est que les différences entre les sexes sont plus en faveur des filles dans le domaine des sciences et des mathématiques que dans la plupart des autres types de tests. Ce modèle de genre peu courant peut, au moins partiellement, s'expliquer par la lourdeur des textes et donc par des capacités de lecture en faveur des filles. Il est également intéressant de noter que ce modèle de genre devient assez différent lorsque les élèves répondent à un questionnaire informatisé, comme cela a été fait dans l'évaluation appelée « Computer-Based assessment in Science » (CBAS). Dans ce test, les garçons surclassent réellement les filles en sciences (OCDE, 2010b). Il s'agit d'une indication claire que le contexte et le mode de collecte des données influent sur les résultats de façon significative.

Les textes « authentiques », qui constituent l'introduction dans chaque unité d'évaluation, ont été réalisés par un pays particulier de l'OCDE et, bien sûr, dans la langue de ce dernier. Ce texte est ensuite, comme mentionné ci-dessous, traduit dans les deux langues officielles du PISA avant d'être ensuite traduit dans la langue de chaque pays participant. Ce processus de traduction suit des règles très strictes qui sont énoncées dans des instructions détaillées (OCDE, 2009). Ce processus de traduction soulève de nombreuses questions. Un travail approfondi sur les tests PISA dans le domaine de la lecture a été effectué par Inga Arffman, dans le cadre de sa thèse de doctorat (2007). Elle fournit une étude textuelle analytique détaillée de la traduction de l'anglais au finnois de trois unités d'évaluation PISA. D'un point de vue linguistique, et en accord avec la théorie de la traduction, son étude révèle en détail plusieurs dimensions critiques dans ce processus. L'une de ses conclusions, fondée sur des éléments théoriques et sur un examen de certaines études empiriques, est que l'on ne peut jamais arriver à ce qu'on peut appeler une « équivalence de traduction ». Elle souligne également la rareté des recherches sur cette question très importante. Tous les traducteurs reconnaissent qu'aucun texte ne peut être traduit conformément à un ensemble formalisé de règles. Et même lorsque de bonnes règles ont été suivies, les traductions reflètent rarement le sens authentique du texte d'origine.

Il semble qu'il y ait un manque d'études empiriques qui se penchent sur cet aspect très important du PISA (et de TIMSS, PIRLS etc.). Le rôle clé joué par les textes dans les évaluations PISA fait d'un tel examen une priorité de recherche. Une vérification minutieuse et transnationale de la traduction exige une coopération de chercheurs de nombreux pays, possédant aussi bien des compétences linguistiques que des connaissances approfondies des domaines évalués. Cependant, certaines langues se prêtent à des comparaisons plutôt simples, même par des non spécialistes. Les trois langues scandinaves offrent de bons exemples. Le suédois, le danois et le norvégien sont des langues très proches, qui ressemblent plutôt à des dialectes, avec une tradition littéraire commune. A partir d'un exemple, nous présentons dans ce qui suit une comparaison simple.

Le clonage est probablement l'élément le plus connu des évaluations PISA. Le point sur le clonage de la brebis Dolly est paru en 2006. Le texte original est en annexe, la traduction en français est reproduite ci-après :

Version en français de l'item « Clonage » Dolly de PISA

(consulté en février 2010 à partir de

http://www.oecd.org/document/38/0,3746,en_32252351_32236173_34993126_1_1_1_1,00.html)

Lisez l'extrait de presse suivant et répondez aux questions qui l'accompagnent.

Une machine à copier les êtres vivants ?

Aucun doute : s'il y avait eu des élections pour désigner l'animal de l'année 1997, Dolly les aurait remportées haut la main ! Dolly est la brebis écossaise que vous voyez sur la photo. Cependant, Dolly n'est pas une brebis quelconque : elle est le clone d'une autre brebis. Un clone signifie une copie conforme. Cloner signifie « copier à partir d'un original unique ».

10 Les chercheurs ont réussi à créer une brebis (Dolly) identique à une autre brebis qui a servi d'« original ».

Le chercheur écossais Ian Wilmut a été le concepteur de ce mécanisme à copier les moutons. Il a prélevé un minuscule fragment de la mamelle d'une brebis adulte (brebis 1). De ce fragment, il a extrait

le noyau, ensuite il a transféré ce noyau à l'intérieur de l'ovule d'une autre brebis (brebis 2). Il avait préalablement retiré de cet ovule tous les éléments qui auraient contribué à donner les caractéristiques de la brebis 2 à l'agneau qui en serait né. Ensuite, Wilmut a implanté cet ovule manipulé de la brebis 2 dans une troisième brebis (brebis 3). La brebis 3 est devenue pleine et a donné le jour à un agneau : Dolly.

Certains savants pensent que, dans quelques années, il sera également possible de cloner des êtres humains. Cependant, de nombreux gouvernements ont déjà établi des lois qui interdisent le clonage des humains.

Question 1 - À quel mouton Dolly est-elle identique ?

- A. À la brebis 1.
- B. À la brebis 2.
- C. À la brebis 3.
- D. Au père de Dolly.

Question 2 - Les lignes 15-16 décrivent la partie de mamelle utilisée par le chercheur comme « un minuscule fragment ». Le contenu de l'article permet de comprendre ce que veut dire ce « minuscule fragment ». Le « minuscule fragment » est :

- A. Une cellule.
- B. Un gène.
- C. Le noyau d'une cellule.
- D. Un chromosome.

A partir de l'original (voir annexe), les trois textes scandinaves (maintenant disponibles sur le site web officiel du PISA) furent traduits, probablement selon les règles détaillées et les instructions établies par PISA. Il est important de voir que les trois versions qui en résultent sont assez différentes les unes des autres. Elles ont toutes changé l'original de manière assez spectaculaire. De ce fait, l'observation la plus frappante et immédiate est que la traduction a rendu chacun des trois textes scandinaves insolite et lourd. En bref,

- les textes danois et suédois ont changé la disposition et la mise en forme, transformant les lignes courtes du texte original en lignes plus longues. On a également supprimé les numéros des lignes ainsi que la référence au texte à la question 2 ;

- les textes suédois, danois et norvégien ont changé le mot « nucleus » en « noyau de cellules » et ont fourni ainsi l'indice que le « minuscule fragment » de la question 2 est en fait une cellule ;
- alors que les textes en anglais (et en suédois) déclarent qu'on enlève « *le matériel qui aurait contribué à donner les caractéristiques de la brebis 2 à l'agneau qui en serait né...* », dans le texte danois on déclare que l'on « *supprime le matériel génétique* », changeant ainsi le sens de la phrase, en introduisant une notion scientifique qui ne figure pas dans le texte original.
- Dans la version norvégienne, l'affirmation « tout le matériel est retiré de la cellule-œuf » rend la phrase plus ou moins dénuée de sens ;
- le texte danois modifie la question 1 en demandant « de quel mouton Dolly est-elle la copie ? » (probablement parce que l'on trouve que cela concorde mieux avec la problématique, ce qui est vrai en effet.) La version danoise est également plus en accord avec le titre de la rubrique du texte : « Une machine à copier les êtres vivants ». (Cette façon de parler et d'écrire sur le clonage n'est en fait utilisée dans aucune langue nordique, ni sans doute dans aucune autre langue).

Dans les rapports du PISA, on assure les lecteurs de la haute qualité des traductions ainsi que du reste du projet (OCDE, 2009). Cette « équivalence de traduction » dans tous les pays est également soigneusement discutée dans Grisay et al. (2007). L'exemple ci-dessus atteste des réserves que l'on peut avoir. Arffman (2010) note que les mauvaises traductions conduisent les lecteurs, dans le cadre des enquêtes, à ne plus manifester d'intérêt pour le texte et à perdre toute motivation à répondre aux questions. Il en résulte des conséquences très négatives sur les résultats. Nous affirmons que le problème est amplifié dans les pays où les étudiants sont critiques, indépendants, et refusent d'obéir à l'autorité des écoles et des enseignants. Ce point sur la motivation et la volonté de s'engager a été discuté ailleurs (Sjøberg, 2007).

■ **Formulation des questions en situations de « vie réelle »**

Nous l'avons déjà souligné, la revendication fondamentale du projet PISA est d'évaluer la façon dont les jeunes sont prêts à affronter les défis de l'avenir. Si cette ambition est louable, les modalités de mise en œuvre ne ressemblent en rien à des situations de la vie réelle. D'abord, les questions sont rhétoriques, les élèves sont testés individuellement pour 2 ½ heures sans accès aux sources d'information, en situation de résolution de problème papier/crayon. Comment peut-on concevoir que cette situation ressemble à celle de la « vie réelle » ? Comment trouver des liens avec les défis auxquels les jeunes peuvent faire face dans leur vie future en tant que citoyens, participants à la démocratie de demain, et en tant que futurs travailleurs qualifiés ? Ensuite, de telles situations où une personne se trouve seule face à une épreuve écrite, sans aucune autorisation à communiquer ou à utiliser les technologies modernes de l'information, sont des situations d'évaluation dans les écoles et les universités. Dans la vie réelle, en privé, durant les loisirs, aussi bien qu'au travail, les modalités sont différentes.

■ **Quelles compétences exigent les employeurs ?**

PISA ne reflète pas les objectifs politiques et économiques de l'OCDE en matière d'emploi. Car on devrait s'attendre, d'une organisation comme l'OCDE, qu'elle insiste sur les compétences requises par les grands acteurs internationaux sur un marché concurrentiel.

Beaucoup de sources indiquent en détail les compétences et les qualifications que les grands employeurs dans le secteur de la haute technologie exigent de leurs employés. Voici deux exemples.

Lors du séminaire officiel de Bologne en 2008 sur *L'employabilité : les perspectives des employeurs et ses implications*, Frank Stephan Becker, directeur des ressources humaines de Siemens AG, a fait un exposé sur son entreprise et a listé les compétences qu'il cherche chez les futurs employés :

- connaissance approfondie de l'objet ;
- capacité à mener un raisonnement analytique, structurer son travail, mener des recherches, évaluer les informations et identifier les problèmes ;
- capacité à voir au-delà de son domaine de compétences et à prendre en compte d'autres facteurs ;
- Indépendance, initiative, apprentissage autonome, techniques de travail, discipline, tolérance à la frustration, capacité d'établir des priorités ;
- compétences en relations interpersonnelles : communication, rétroaction, capacité de travailler en équipe...

Siemens AG est la plus importante compagnie européenne en électronique et génie électrique. Elle et ses filiales emploient environ 360 000 personnes dans près de 190 pays. Il est clair que la plupart des compétences de la liste ci-dessus ne sont pas abordées par le test du projet PISA.

Le second exemple est une enquête menée par l'« Abelia », *Business Association of Norwegian knowledge and technology based enterprises*. Basée sur une enquête auprès de 500 chefs de file dans le secteur le plus concurrentiel de l'économie norvégienne, elle a révélé le classement suivant des compétences souhaitées pour les futurs dirigeants (Abelia, 2008, traduction de l'auteur) :

- bonnes compétences en communication (66 %)
- sens de la réflexion stratégique (61 %)
- capacité à motiver (60 %)
- préoccupation pour le personnel et les collègues de travail (58 %)
- confiance en soi et auto-efficacité (56 %)
- formation solide (48 %)
- visionnaire (33 %)
- compréhension des chiffres et des mesures (24 %).

Comme on peut facilement le voir, les deux listes de compétences sont similaires, mais sont différentes de ce que PISA propose. Toutefois, il est intéressant de noter que les points de vue exprimés par l'industrie de la haute technologie coïncident, à bien des égards, avec les buts et les objectifs de la scolarité dans de nombreuses démocraties modernes, et sont également en accord sur plusieurs aspects avec la pédagogie actuelle. Ainsi, les recommandations fondées sur les résultats du projet PISA peuvent être contre-productives, même pour les sociétés qui opèrent sur le marché mondial.

■ **Problèmes de statistique**

Le projet PISA est une grande entreprise. Il présente bon nombre de caractéristiques de ce que l'on peut appeler la « Big science » et la « techno-science » : il est coûteux et exige la coopération de nombreux pays ; la logistique du projet est compliquée, il existe de nombreux documents, volumineux, fournissant des instructions détaillées pour les différents groupes nationaux dans les pays participants ; des centaines d'experts issus de plusieurs domaines d'expertise sont impliqués ; les contrats avec les sous-traitants sont attribués par appel d'offres ; des milliers d'écoles et d'enseignants sont visés ; près d'un demi-million d'élèves passe 2½ heures à répondre au test cognitif et au questionnaire ; les données sont codées avec soin par des milliers d'agents spécialement formés.

Nous avons vu que le test final de PISA se compose d'unités choisies et sélectionnées, mais les cahiers sur lesquels répondent les élèves ne sont pas identiques. Un système de « rotation » des unités signifie que différents élèves répondent à différents livrets. De cette façon, PISA peut inclure un plus grand nombre d'unités dans le test. Après le long et fastidieux processus d'entrée de données et de codage, ces derniers sont soumis à une analyse statistique complexe, basée sur la théorie de la réponse à l'Item (IRT) et la modélisation de Rasch. C'est bien connu, le score final est normalisé pour fournir une note moyenne internationale de 500 avec une déviation

standard (Ecart type) de 100 pour l'OCDE dans son ensemble. Les étapes qui conduisent des réponses réelles des élèves aux valeurs finales accessibles dans les tableaux de scores sont nombreuses et peu transparentes, même pour les statisticiens et les lecteurs bien informés. Les méthodes ont toutefois été critiquées par les experts statisticiens, y compris certains de ceux qui ont réellement travaillé sur les données du PISA. Svend Kreiner, professeur de statistiques biomédicales à l'Université de Copenhague, a fait valoir que le Danemark peut être placé dans une position élevée et être ainsi dans les premiers rangs de PISA, simplement en modifiant certains paramètres de l'analyse statistique complexe. Il a également souligné que les méthodes de calculs statistiques du PISA sont publiées uniquement sous une forme très générale, ce qui rend difficile une critique détaillée. (Informations du Journal, 14 avril 2011, <http://www.information.dk/265524>)

7. Résultats paradoxaux

Cet article porte principalement sur les aspects politico-économiques du PISA et les faiblesses fondamentales d'ensemble du projet. Cependant, ce sont d'autres sujets de préoccupation qui devraient retenir l'attention, surtout de la part de ceux qui sont en charge de PISA et qui croient dans la fiabilité des données recueillies sur la qualité de l'enseignement et la réussite des élèves. Voici un bref aperçu de ces sujets.

■ **Les ressources n'ont aucune influence ?**

En utilisant les données de la première série de tests PISA, l'OCDE a produit des graphiques et des indicateurs qui ont montré des corrélations faibles ou négligeables entre les scores des pays et leurs dépenses dans le domaine de l'éducation (OCDE, 2001). Bien sûr, ceci a conduit l'OCDE à préciser que l'augmentation des dépenses de l'éducation n'améliorera pas sa qualité. Il est particulièrement intéressant de noter que dans chacun des cinq pays nordiques la relation entre les dépenses publiques et les scores aux évaluations PISA est en réalité fortement négative. Ces résultats sont souvent utilisés au détriment des écoles et des enseignants. La Finlande, par exemple, située en haut du palmarès, a les dépenses les plus faibles. Cette situation a des conséquences d'ordre politique. Les enseignants finlandais ont du mal à demander des salaires plus élevés et à exiger davantage de moyens. La Norvège, autre exemple, est beaucoup moins bien située sur le classement du PISA, mais affiche plus de dépenses publiques en ce qui concerne les établissements scolaires. En se basant sur les résultats des évaluations, le ministre norvégien de l'éducation a fait valoir que l'augmentation des dépenses dans le domaine de l'éducation n'améliorerait pas la qualité de l'éducation, prenant à son compte les recommandations faites par l'OCDE (rapport économique, 2008). Les résultats du PISA, appuyés sur les coûts et les financements du secteur de l'éducation, comme on vient de le voir, sont fréquemment utilisés dans les publications influentes de l'OCDE. Exemple de « *l'éducation annuelle en un coup d'œil* » qui conclue qu'« *en moyenne dans les pays de l'OCDE, il est possible de réduire les apports de 30,7 % tandis que l'on maintient constant les outputs* ». (OCDE, 2007b)

■ **Scores élevés en science, faible intérêt et attitudes négatives**

Les résultats des évaluations PISA sont souvent présentés comme un classement entre les pays, avec des gagnants au sommet et des perdants à la base. Les tests PISA comportent également de nombreuses questions sur les attitudes des jeunes envers la science. C'est une composante importante de l'étude PISA 2006, quand la science était le domaine majeur. La définition de la culture scientifique inclut la « *volonté de s'engager en tant que citoyen réfléchi à propos de problèmes à caractère scientifique et touchant à des notions relatives aux sciences* » (OCDE, 2006, p.39). Cependant, dans les scores du PISA qui ont été utilisés pour les classements, les indices et les valeurs concernant les attitudes n'y figuraient pas. Un numéro spécial de *l'International Journal of Science Education* (2011, vol.33, n°1) présente plusieurs résultats intéressants des analyses fondées sur ces données. On y apprend que de nombreux pays, qui ont un haut score dans PISA en science, sont au bas de l'échelle en ce qui concerne l'intérêt des

élèves pour les sciences (Bybee & McRae, 2011). La Finlande et le Japon sont de parfaits exemples : bien qu'en haut de l'échelle des scores en compétences en science, ils sont en bas de l'échelle en ce qui concerne les indices tels que l'« intérêt pour les sciences », la « motivation orientée vers l'avenir pour apprendre les sciences » et l'« emploi futur en science », c'est-à-dire la propension à se voir à l'avenir dans des études et des carrières scientifiques. En fait, le score de la compétence en science du PISA est corrélé négativement avec l'orientation future en science ($r = -0.83$) et avec l'emploi futur en science ($r = 0.53$) (Kjærnsli & Lie, 2011).

De tels résultats sont dérangeants. Si les élèves les mieux classés aux évaluations PISA achèvent la scolarité obligatoire avec une attitude fortement négative envers la science, il faut prendre du recul et réfléchir sur les raisons de cette situation ainsi que sur les conséquences possibles. Bien sûr, il ne faut pas confondre corrélation et causalité. Mais on doit se donner un temps de réflexion avant d'utiliser ces pays comme des modèles éducatifs et des idéaux à suivre. Il est également intéressant de noter que bon nombre de pays « gagnants » dans les scores en ce qui concerne la science sont également les pays où existent les plus grandes différences entre les sexes. La Finlande est encore un bon exemple, car les filles surclassent les garçons sur l'ensemble des trois sujets testés par PISA. Concernant la lecture, la différence est d'environ 50% d'un écart-type. Encore une fois, ces conclusions doivent se traduire par une mise en garde contre la tentative de prendre les « gagnants du PISA » comme des modèles à suivre.

■ **Taille des classes et des écoles**

Un autre ensemble de résultats des évaluations PISA liés à l'utilisation des ressources concerne la taille des classes et des établissements. Les corrélations négatives observées pour de telles variables conduisent à des recommandations en termes de rentabilité, par l'augmentation de la taille des classes et des écoles. Naturellement, de tels calculs sont purement statistiques, et ne prennent pas en considération la répartition géographique de la population ou les préoccupations culturelles et nationales concernant les politiques sociales et régionales.

■ **Enseignement traditionnel, de meilleurs résultats ?**

Le questionnaire *Elèves de l'enquête PISA* comporte une série de questions sur les origines et l'environnement familiaux, culturels, etc. Il comprend également une série de questions sur les méthodes d'enseignement et les pratiques pédagogiques en salle de classe qui caractérisent les expériences scolaires des élèves. Lorsque l'on cherche des relations possibles entre ces variables et les données du PISA, on trouve des résultats surprenants qui requièrent l'attention de la part des chercheurs et des décideurs. Le plus étonnant concerne certains résultats qui sont à l'opposé des préconisations actuelles émanant de divers spécialistes de l'éducation scientifique, du curriculum ainsi que des décideurs, etc.

Une tendance actuelle dans l'enseignement des sciences est l'importance accordée aux méthodes actives d'apprentissage et d'enquête. Ces opinions sont exprimées par des groupes de spécialistes de l'enseignement des sciences (Osborne & Dillon, 2008) ainsi que dans les rapports officiels de l'OCDE (2006b) et le rapport de l'Union européenne (UE, 2007). Ces rapports ont influencé une grande partie des travaux de recherche et de développement soutenus par le programme cadre en cours (FP7) de l'Union Européenne, notamment ceux qui portent sur l'IBSE (Inquiry Based Science Education ou son équivalent français la méthode d'investigation). Les résultats des évaluations PISA semblent être en corrélation négative avec beaucoup de méthodes pédagogiques actives comme celles qui préconisent de « formuler ses propres problèmes et réponses... » et la pratique expérimentale (Kjærnsli & al., 2007). Ces résultats constituent un vrai paradoxe quand on sait que des initiatives sont considérées comme des sources d'inspiration pour des éducateurs en science ainsi que pour les décideurs nationaux. « La main à la pâte », réalisée sous l'égide de l'Académie des Sciences, inspirée de l'expérience américaine initiée par le prix Nobel Léon Lederman, qui met l'accent sur l'investigation autonome, en constitue un exemple.

Jenkins (2009) a étudié les réformes et les tendances actuelles de l'enseignement des sciences. Il en ressort quelques concepts clés dans la pensée actuelle en éducation scientifique que l'on peut énumérer : science en contexte, méthodes d'investigation (IBSE), la main à la pâte, méthodes actives d'apprentissage, nature de la science, questions socio-scientifiques, argumentation, STS (Science, technologie et société). Il semble qu'aucun résultat du PISA ne peut soutenir de telles tendances. Les résultats conduisent plutôt à des contre-évidences. Cette contradiction semble problématique et nécessite une enquête qui doit être menée par la communauté des chercheurs.

■ *L'utilisation des TIC permet-elle de baisser les scores ?*

Le test PISA, à travers deux constructs pertinents, évalue l'utilisation des technologies de l'Information et de la communication (TIC) dans les écoles. Le premier est un indice lié à l'utilisation d'internet dans les écoles, le second correspond à l'utilisation de logiciels et de programmes éducatifs. Dans une étude détaillée dans cinq pays nordiques, Kjærnsli et al. (2007) établissent clairement une relation négative entre l'utilisation des TIC et les évaluations PISA. Il est intéressant de noter que le « vainqueur » de ces évaluations, la Finlande, est réellement en dessous de la moyenne de l'OCDE concernant cet indicateur. De surcroît, c'est de loin le pays nordique qui utilise le moins les TIC. En revanche, la Norvège qui présente des résultats plutôt modestes, est située au dessus de la moyenne de l'OCDE dans tous les indicateurs sur l'utilisation des TIC dans les écoles. Néanmoins, les directives en matière de politique, en Norvège, sont d'accroître l'utilisation des TIC dans les écoles afin d'atteindre des notes supérieures dans PISA.

■ *Résultats surprenants du PISA : remarque finale*

Certains des résultats problématiques mentionnés ci-dessus ne sont pas difficiles à comprendre. Une épreuve écrite comme PISA peut difficilement mesurer les compétences et les connaissances acquises dans un laboratoire ou lors d'une excursion ; elle peut difficilement mesurer l'intérêt, la curiosité et l'enthousiasme que peuvent susciter l'argumentation, l'enquête et la recherche de réponses aux questions que les élèves ont eux-mêmes formulées. Si pour évaluer la qualité de l'éducation on s'appuie sur un test final écrit, il n'est pas surprenant que l'enseignement sera plus rentable si les élèves ne consacrent pas beaucoup de temps aux excursions, aux travaux pratiques ou à des discussions portant sur des questions socio-scientifiques. L'utilisation des données du PISA pour des recommandations politiques est, au mieux, très sélective. Si l'on souhaite en tirer le meilleur profit à des fins politiques, il est essentiel de prendre tous les résultats au sérieux, surtout ceux qui sont inattendus et en contradiction avec les politiques existantes.

Conclusion

Cet article a mis l'accent sur les aspects problématiques des évaluations PISA, principalement ceux qui sont négligés par la recherche, ainsi que dans les débats publics. Les points positifs ne doivent pas être ignorés, comme l'évaluation des compétences que les jeunes ont besoin de développer pour acquérir les différentes « littéracies » nécessaires dans leur vie future ainsi que pour le bon fonctionnement de leurs sociétés. PISA a conduit à un intérêt accru pour l'éducation grâce à la banque de données énorme générée par les cycles successifs. C'est certainement la source de données la plus importante et la plus professionnelle dans l'histoire des sciences sociales et de la recherche en éducation. Ces données sont disponibles pour les chercheurs.

Les comparaisons internationales en matière d'éducation sont importantes. Elles peuvent ouvrir de nouvelles perspectives, être une source d'inspiration et fournir des idées pour les éducateurs, les chercheurs et les décideurs. Cependant, elles peuvent être comprises et utilisées de manière opposée. Ces études peuvent révéler et mettre en valeur une grande variété entre les jeunes, les nations et les cultures sur les aspects de l'éducation, et être une source d'inspiration [de quoi ?].

Elles peuvent également être, d'un point de vue normatif, à l'origine des pressions exercées par des experts externes pour se conformer à l'universel et à des normes communes. L'auteur est inquiet de voir cette dernière perspective dominer, comme on peut le voir, dès la phase de conception des évaluations PISA, à la façon dont les données sont utilisées par l'OCDE dans le débat public et les politiques éducatives des différents pays.

On peut facilement être en accord avec les intentions officielles des évaluations PISA. Personne ne peut contester la nécessité de vérifier si les jeunes développent les connaissances et les compétences nécessaires pour relever les défis qu'ils rencontreront en tant que citoyens. Cependant, les difficultés inhérentes aux évaluations sont rarement discutées. De même, les motifs économique et politique sous-jacents au projet conduit par l'OCDE sont souvent ignorés de ceux mêmes qui sont impliqués, comme les chercheurs. Cet article a souligné certains des points problématiques : la sélection des sujets (et l'ignorance des autres), le développement d'un test qui, par nécessité, conduit à des questions sans contexte. Les traductions engendrent de nouvelles difficultés. Dans cet article et ailleurs (Sjøberg, 2007), j'ai soutenu que la réalisation d'un test international sans parti pris est en fait une « mission impossible ».

Cependant, le public, les médias et les décideurs politiques considèrent souvent les scores et les classements du PISA comme des faits. Ils font confiance aux experts et pensent que les chiffres sont des mesures objectives et neutres. Confiants en ces scores, ils sont persuadés qu'ils atteignent bien le but recherché. Aucun test ne vaut mieux que les unités dont il est construit. Bon nombre d'articles publiés ont suscité de sérieuses critiques, tant sur leur contenu que sur le langage utilisé, mettant ainsi en doute leur pertinence.

Examiner les questions entourant l'utilisation statistique des données nécessite davantage de recherches. Une recherche attentive a aussi besoin d'être dirigée vers certains résultats apparemment paradoxaux concernant, par exemple, les scores du PISA et le niveau de ressources ou la taille des classes. Il y a aussi un ordre du jour pour les décideurs. L'influence du PISA sur la politique éducative doit être analysée de façon plus approfondie qu'elle ne l'a été jusqu'à présent de la part des responsables politiques, des enseignants et du grand public. Ceux-ci doivent porter beaucoup plus d'attention aux critiques formulées précédemment. En outre, ils doivent être conscients des dangers de l'utilisation de données pour justifier et légitimer des réformes de l'éducation, surtout lorsque de telles réformes vont à l'encontre de ce qui est largement admis par ailleurs, comme les pratiques pédagogiques ayant fait leurs preuves. PISA est essentiellement un projet politique et non une voix critique. S'appuyer sur une recherche indépendante est un besoin urgent.

Bibliographie

ARFFMAN I. (2007), *The problem of equivalence in translating texts in international reading literacy studies. A text analytic study of three English and Finnish texts used in the PISA 2000 reading test* (Research Reports 21), Jyväskylän, University of Jyväskylä, Institute for Educational Research.

ARFFMAN I. (2010), « Equivalence of Translations in International Reading Literacy Studies », *Scandinavian Journal of Educational Research*, vol.54, n°1, February 2010, p.37-59.

BERGESEN O.H. (2006), *Kampen om kunnskapsskolen* (The fight for a knowledge-based school), Oslo, Universitetsforlaget.

BYBEE R. & McCRAE B.J. (2011), « Scientific Literacy and Student Attitudes : perspectives from PISA 2006 science », *International Journal of Science Education*, n°33(1), p.7-26.

EU (2007), *Science Education Now : a renewed pedagogy for the future of Europe*, (The Rocard report), European Commission, Brussels, EC.

GRISAY A., DE JONG J.H., GEBHARDT E., BERZNER A., HALLEUX-MONSEUR B. (2007), « Translation equivalence across PISA countries », *Journal of Applied Measurement*, 2007-8(3), p.249-266.

JENKINS E.W. (2009), « Reforming school science : a commentary on selected reports and policy documents », *Studies in Science Education*, vol.45, n°1, p.65-92.

KD (2011), [The Ministry of Education and Research], *Motivasjon – Mestring – Muligheter* White paper 22 to the Parliament.

KJAERNSLI M. & LIE S. (2011), « Students' preference for science careers : international comparisons based on PISA 2006 », *International Journal of Science Education*, n°33(1), p.121-144.

KJAERNSLI M., LIE S., OLSEN R.V. & ROE A. (2007), Tid for tunge løft. Norske elevers kompetanse i naturfag, lesing og matematikk i PISA 2006. [Norwegian PISA report], Oslo, Universitetsforlaget.

LOVELESS T. (2009), « How well are American students learning ? », *The Brown Center Report on American Education*, vol. II, n°3.

OCDE (2001), *Knowledge and skills for life. First results from PISA 2000*, Paris, OCDE Publications.

OCDE (2006a), *Assessing Scientific, Reading and Mathematical Literacy, A framework for PISA 2006*, Paris, OCDE Publications.

OCDE (2006b), *Evolution of Student Interest in Science and Technology Studies*, Policy Report from OCDE Global science forum <http://www.oecd.org/dataoecd/16/30/36645825.pdf>

OCDE (2007a), *PISA 2006 : science competencies for tomorrow's world. Volume 1 - Analysis*, Paris, OCDE Publications.

OCDE (2007b), *Education at a Glance*, Paris, OCDE Publications.

OCDE (2008), *OECD Economic Surveys, Norway*, vol. 2008/13, Paris, OCDE Publications.

OCDE (2009), *PISA 2006 Technical Report Organization* for Economic Co-operation and Development. <http://www.oecd.org/dataoecd/0/47/42025182.pdf>

OCDE (2010a), *PISA 2009 Results : what Makes a School Successful ?*, vol. IV, Paris, OCDE Publications.

OCDE (2010b), *PISA computer-based assessment of student skills in science*, Paris, OCDE Publications.

OLSEN R.V. & LIE S. (2011), « Profiles of students' interest in science issues around the world : analysis of data from PISA 2006 », *International Journal of Science Education*, n°33(1), p.97-120.

OSBORNE J. & DILLON J. (éd.) (2008), *Science Education in Europe: Critical Reflections*, A Report to the Nuffield Foundation, King's College London.

SJOBERG S. (2007), « PISA and "Real Life Challenges" : Mission Impossible ? », *PISA zufolge PISA - PISA According to PISA*, S.T. Hopmann, G. Brinek & M. Retzl (éd.), LIT Verlag, Berlin.

UNDP (1990-2012), *Human Development Report*, New York and London, Oxford University Press.

Annexe

Version originale de l'unité sur le clonage (PISA 2006)

Read the newspaper article and answer the questions that follow.

A copying machine for living beings?
Without any doubt, if there had been elections for the animal of the year 1997, Dolly would have been the winner! Dolly is a Scottish sheep that you see in the photo. But Dolly is not just a simple sheep. She is a clone of another sheep. A clone means: a copy. Cloning means copying 'from a single master copy'. Scientists succeeded in creating a sheep (Dolly) that is identical to a sheep that functioned as a 'master copy'.
It was the Scottish scientist Ian Wilmut who designed the 'copying machine' for sheep. He took a very small piece from the udder of an adult sheep (sheep 1).

From that small piece he removed the nucleus, then he transferred the nucleus into the egg-cell of another (female) sheep (sheep 2). But first he removed from that egg-cell all the material that would have determined sheep 2 characteristics in a lamb produced from that egg-cell. Ian Wilmut implanted the manipulated eggcell of sheep 2 into yet another (female) sheep (sheep 3). Sheep 3 became pregnant and had a lamb: Dolly.
Some scientists think that within a few years it will be possible to clone people as well. But many governments have already decided to forbid cloning of people by law.

Question 1: CLONING S128Q01

Which sheep is Dolly identical to?

- A Sheep 1
- B Sheep 2
- C Sheep 3
- D Dolly's father

Question 2: CLONING S128Q02

In line 14 the part of the udder that was used is described as "a very small piece". From the article text you can work out what is meant by "a very small piece". That "very small piece" is

- A a cell.
- B a gene.
- C a cell nucleus.
- D a chromosome.