

PISA: L'exception française

Michel Delord, 24/04/2014

[Introduction et débat sur MEDIAPART](#)

What if you learned that Pisa's comparisons are not based on a common test, but on different students answering different questions? And what if switching these questions around leads to huge variations in the all- important Pisa rankings, with the UK finishing anywhere between 14th and 30th and Denmark between fifth and 37th? What if these rankings – that so many reputations and billions of pounds depend on, that have so much impact on students and teachers around the world – are in fact “useless”?

[Is Pisa fundamentally flawed?](#) in *Times Educational Supplement*, 2013/07/26.

This lack of statistical transparency has also been a focal point of criticism leveled at Pisa and indeed the OECD. The last time they published a technical report to make their methods more open, they weren't able to make it any more concise than [419 pages](#).

[The Pisa methodology: do its education claims stack up?](#) in *The Guardian*, 2013/12/03

1

*
* *

[Réponses plausibles des élèves](#)

[Le modèle de Rasch](#)

[La terrible précision PISA : Groland est 30^{ème} sur 50 pays, avec une marge d'erreur de 200 places ...](#)
[L'exception française](#)

ANNEXES :

[Annexe 1 - Intermède : A l'école, se sentir « chez soi » ou se sentir « à sa place » ?](#)

[Annexe 2 - Quelques mots sur l'étude de l'APED « PISA 2012 sans fard et sans voile »](#)

[Annexe 3 - Pour la bonne bouche : les 'risk groups', une bonne raison sécuritaire de soutenir PISA](#)

[Annexe 4 - Bibliographie : quelques références critiques récentes sur PISA](#)

*
* *

Dans [Vaccination contre le PISA-Choc](#) de février 2014, je m'étais bien gardé d'évoquer les critiques portant sur la méthodologie statistique de PISA, me contentant d'évoquer d'autres aspects suffisamment nombreux à mon sens pour justifier une défiance majeure à la fois par rapport à la valeur de PISA et aussi par rapport à la manière dont ces résultats étaient utilisés médiatiquement et politiquement.

Je ne m'attendais certes pas à ce que les médias et l'éventail politique s'intéresse à mon point de vue qui met en cause la globalité de leur conception de l'école. Mais il se trouve que même certains auteurs qui montraient le plus souvent un esprit critique par rapport aux thèses de l'OCDE, comme Nico Hirtt, ne sont manifestement pas convaincus par mon analyse et continuent à considérer comme valides les données de cette enquête.

Suivent donc des éléments certes brefs mais fondamentaux à propos d'une catégorie de biais méthodologiques - *je souligne une car il y en a bien d'autres* - qui entachent fondamentalement la validité de l'ensemble des données de l'enquête PISA. Ces biais sont de plus extrêmement intéressants à connaître puisqu'ils peuvent entacher la valeur de ce qui est très à la mode, les méta-analyses, enquêtes qui ont le double avantage de sembler profondes « puisqu'elles reprennent tous les résultats connus sur un sujet » et de ne pas coûter cher puisqu'il n'y a pas d'échantillon de population à interviewer.

Bonne lecture

Barde, le 24 avril 2014

Michel Delord

Réponses plausibles des élèves

"Each PISA survey includes about seven hours of test material. From this, each student takes a two-hour test, with the actual combination of test materials different for every student."

Source : [OECD, About PISA, PISA FAQ, The Test and Questionnaires, How does PISA test this ?](#)

Il faut donc sept heures pour répondre à toutes les questions d'une évaluation PISA mais, pour diverses raisons, chaque élève n'en passe que deux heures. Tous les élèves testés ne répondent donc pas aux mêmes questions.

Mais comme les résultats de PISA portent sur toutes les questions, quelle réponse attribue-t-on à un élève à qui la question n'a pas été posée ? On attribue ce que l'on appelle une « **réponse plausible** ».

Ce cas est-il marginal ? Pas du tout. Dans l'enquête PISA 2006, il n'y a qu'en Sciences, majeure de PISA 2006, que tous les élèves ont été testés. Seule la moitié de l'effectif a été testée en *Culture mathématique*. Il en a été de même en *Compréhension de l'écrit*, domaine dans lequel de plus toutes les questions n'ont pas été posées à ceux qui étaient effectivement testés. Svend Kreiner, mathématicien spécialiste de ces questions -voir bibliographie [infra](#) - déclare : *il y avait au total 28 questions et approximativement seulement 10% des élèves qui ont pris part à PISA ont effectivement répondu à ces 28 questions.*

Le modèle de Rasch

Sans entrer dans des détails trop techniques, disons que pour attribuer ces « réponses plausibles » et pour rendre comparables les résultats des élèves testés sur un sujet et ceux des non testés, ce qui est la source de toutes les comparaisons possibles tirées de PISA, que ce soient les comparaisons entre pays dans une évaluation PISA ou l'évolution des performances d'un pays, on utilise un modèle statistique de mise à échelle, le [modèle de Rasch](#) du nom du mathématicien [Georg Rasch](#) qui l'a inventé.

Mais, comme toute théorie qui se respecte, ce modèle ne produit les résultats escomptés que s'il est appliqué dans des conditions pour lesquelles il a été prévu. Et c'est là que le bât blesse.

On peut légitimement s'appuyer sur ce qu'en dit *Svend Kreiner*, professeur émérite à l'Université de Copenhague puisque c'est un élève de Georg Rasch avec qui il a personnellement travaillé pendant quarante ans. Pour que l'on puisse appliquer le modèle de Rasch à PISA, il faut que toutes les questions utilisées fonctionnent de la même manière et qu'elles aient le même degré de difficulté pour tous ceux qui participent et en particulier pour tous les pays ; autrement dit en termes techniques lorsqu'il y a un *fonctionnement différentiel des items* [*differential item functioning*, DIF], on ne peut appliquer le modèle de Rasch. Et il existe bien un DIF pour PISA¹ ... et l'OCDE se sert cependant du modèle de Rasch : en effet, même s'il y a une certaine universalité en *Culture mathématique*, ce n'est pas le cas dans les autres domaines de PISA et de plus on ne peut affirmer que les différentes questions de mathématiques ne dépendent pas dans une certaine mesure de la langue dans laquelle la question est posée et de la civilisation concernée. En tout cas, si l'on quitte le terrain de la *Culture mathématique* qui est en d'autant moins universelle qu'elle est « pratique »

¹ J'exagère - mais si peu comme on va le voir - : il y a bien UN PAYS ET UN SEUL sur tous ceux qui sont testés par PISA pour lequel on peut appliquer le modèle de Rasch, c'est [roulements de tambours] ... *le Lichtenstein* comme le montre Svend Kreiner avec l'aide du Andersen's CLR test*, à la page 9 de [Model fit and robustness? - A critical look at the foundation of the PISA project](#).

* Andersen's CLR Test veut dire *Andersen's Conditional Likelihood Ratio Test* et l'on peut en trouver [ICI](#) une description faite par Erling Andersen lui-même.

et « conforme au contexte local » - ce qui est bien la revendication de PISA -, on est sûr qu'il y a de fortes différences dont PISA ne tient pas compte. Pour la *Compréhension de l'écrit*, un présupposé non explicité mais effectivement appliqué pour évaluer ce domaine est, par exemple, qu'un texte en chinois présente le même niveau de difficulté qu'un texte en espagnol. Ce qui se discute, c'est le moins que l'on puisse en dire.

NB : Je n'aborde ici que des *fautes méthodologiques générales graves* et non, par exemple, des fautes plus mineures comme des erreurs de traduction des questions, erreurs suffisamment importantes pour changer complètement le sens de la question. Mais Loys Bonod, qui a publié sur son site [La Vie Moderne.net](http://LaVieModerne.net) d'intéressantes critiques de PISA [[Nouvelle livraison de Pisa](#) , [Les résultats de PISA 2012](#) ...] en a relevé plusieurs. Vous en trouverez une [à la fin de ce texte](#).

La terrible précision PISA : Groland est 30^{ème} sur 50 pays, avec une marge d'erreur de 200 places ...

Pour donner un ordre grandeur de l'erreur engendrée par l'application induite du modèle de Rasch, prenons un exemple, le classement des pays. Souvenons-nous bien qu'il ne s'agit que d'un exemple d'erreur choisi ici parce que simple à traiter dans le cadre réduit de cet article ; mais ce ne sont pas simplement ces trois classements qui sont pour le moins critiquables mais toutes les données fournies par PISA qui s'appuient sur le modèle de Rasch.

En 2006, 57 pays passent les tests PISA ; suivant la manière dont on traite les données sur la compréhension de l'écrit, Svend Kreiner montre, par exemple,

- que le Japon, classé officiellement 24^{ème}, peut être à n'importe quelle place entre la 8^{ème} et la 40^{ème}.
- que le Canada, classé officiellement 14^{ème} aurait pu avoir n'importe quel classement entre la 2^{ème} et la 25^{ème} place.

Un peu bousculés pendant tout un temps, les responsables de l'OCDE ont fini par reconnaître - *discrètement certes sans le crier sur les toits, ce qui fait probablement que les journalistes français ne l'ont pas entendu* - qu'il y avait « une incertitude » sur le classement des pays et que « de grandes variations » pouvaient affecter le classement d'un pays.

Ils ont même précisé que, dans le cas de l'Angleterre en 2009,

- son classement pour la *Compréhension de l'écrit* annoncé officiellement comme 25^{ème} pouvait varier entre la 19^{ème} et la 27^{ème} place,
- son classement pour la Culture mathématique annoncé officiellement comme 28^{ème} se trouvait entre la 23^{ème} et la 31^{ème} place
- son classement pour les Sciences annoncé officiellement comme 16^{ème} pouvait varier de la 14^{ème} à la 19^{ème}.

L'OCDE reconnaît donc - même si elle les minimise car il n'y a, à ma connaissance, aucune raison de douter des calculs de Kreiner - qu'il y a des marges d'erreurs qui peuvent être considérables.

Mais, comme on peut le constater en consultant les résultats de PISA, l'OCDE, après avoir admis ces « quelques défauts » continue toujours à donner les résultats des classements de la même manière, c'est-à-dire sans indiquer la marge d'erreur (où si elle la donne, ce qui arrive, c'est dans une « annexe technique » ou un endroit quelconque aussi accessible au grand public que les conditions écrites en minuscules au dos de la 12^{ème} page d'un contrat). Un statisticien qui fait ça est *normalement* - à moins qu'on lui ait suggéré ou fortement recommandé de le faire - viré immédiatement.

Mais en France on ne s'inquiète pas de tout cela : lors des derniers résultats PISA, [Caroline Beyer du Figaro](#) s'inquiète du fait que « *la France perd deux places* ». [Stanislas Kraland du Huffington Post](#)² n'est pas en reste et titre « *Éducation : au classement PISA, la France perd deux places* », alors qu'il se peut tout à fait qu'une perte de deux places signifie, si l'on compile autrement les données, un gain de 10 places ou une perte de 15... Quant à [Mattea Battaglia et Aurélie Collas du Monde](#), elles se croient subtiles et progressistes en écrivant : « *Les statisticiens de l'OCDE situent la France à la 25e place sur 65 participants ... Si l'on ne tenait compte que des performances des élèves issus de milieux sociaux favorisés, la France se situerait dans le premier quart du classement (13e). En revanche, en ne comptabilisant que les élèves les plus défavorisés, elle perdrait vingt places, pour figurer au 33e rang.* » Remarquable arithmétique qui a un sens seulement si une 25^{ème} place officielle de la France ne signifie pas, comme pour le Japon classé 24^{ème} qu'on ne sait pas quelle est sa place, *favorisés et défavorisés compris*, entre la 8^{ème} et la 40^{ème}.

Finalement, l'appréciation de Svend Kreiner « *La chose la moins désagréable que l'on puisse dire sur les classements PISA, c'est qu'ils sont inutiles* » semble donc tout à fait mesurée.

Et celle du mathématicien Hugh Morrison de la *Queen's University* de Belfast, spécialisé dans les questions éducatives « *Le modèle utilisé dans PISA est « entièrement erroné » parce qu'il contient « une profonde erreur conceptuelle ». Pour cette raison, PISA n'est pas améliorable et ne fonctionnera jamais.* » mérite tout à fait d'être prise en compte et discutée.

Il faut également remarquer - je pense qu'il est utile d'insister sur la question - que, comme l'utilisation induite du modèle de Rasch perturbe au minimum et de manière non secondaire tout ce qui est comparaison entre pays, ce qui est sans véritable valeur est non seulement les classements sur les trois domaines *Culture scientifique*, *Culture mathématique* et *Compréhension de l'écrit* mais également toutes les autres comparaisons et par exemple celles sur le caractère discriminatoire des différentes écoles nationales : il est aussi ridicule de vouloir prouver à partir des tests PISA que la « France est la championne des inégalités scolaires » (Le Monde) ou que, au contraire la Finlande ou tout autre pays est le « champion de l'égalité ». Et il faudrait que l'APED avance beaucoup plus d'arguments pour montrer la validité de « son indice global d'inégalité sociale scolaire » apparemment tiré de données PISA qui s'appuient sur l'application induite du modèle de Rasch.

*Depuis le début de ce texte je ne me suis intéressé volontairement qu'à un seul biais négatif de l'enquête PISA, celui qui porte sur l'utilisation induite du modèle de Rasch. **Et je me suis seulement appuyé sur des articles provenant des médias** et surtout du plus complet d'entre eux, celui de*

² Loys Bonod me dit dans un courrier privé :

Remarque à propos de Stanislas Kraland du "Huffington Post" : techniquement, si l'on se fonde sur les résultats PISA 2012 dans les trois domaines, la France n'a pas perdu deux places pour arriver au 25^e rang : elle a au contraire gagné une place en se hissant au 24^e rang. M. Kraland a fait la même erreur que toute la presse : il a resservi servilement - visiblement sans comprendre le principe des trois domaines - les éléments du dossier de presse de PISA 2012, lequel ne concernait que les mathématiques... Bon ce classement n'a évidemment aucun sens, mais cette erreur dans l'erreur montre bien avec quel sérieux est traitée l'information...

William Stewart, *Is PISA fundamentally flawed ?* paru dans le *Times Educational Supplement* du 26 juillet 2013, qu'il est donc recommandé de lire d'abord entièrement.

Mais

- il y a d'autres biais méthodologiques quasiment aussi importants : ils sont présentés et critiqués dans les publications scientifiques de Svend Kreiner, Svein Sjøberg ou de Joachim Wuttke citées dans la bibliographie *infra*.

- il ne faut pas oublier non plus les critiques que j'ai faites dans mon texte précédent [*Vaccination contre le PISA-Choc*](#) dont celui-ci n'est qu'un complément. Je rappelle quelques unes de ces critiques :

- les questions posées dans les tests PISA sont des questions *explicitement* présentées comme non scolaires mais ... les réponses à ces questions, elles, servent à évaluer et classer les systèmes scolaires

- la *Culture mathématique* testée par PISA réduit les mathématiques à rien par deux biais

- en sont exclus *explicitement* « algèbre, calcul littéral, raisonnement déductif [...] objets géométriques » : c'est un peu comme si l'on continuait à appeler couteau l'objet obtenu en lui enlevant la lame et le manche mais dont on aurait gardé la moitié de la virole. Et tous ceux qui prétendent que les résultats des tests PISA ont une valeur quelconque ne commettent pas une erreur plus grande que ceux qui prétendent couper leur viande avec un couteau réduit à une demi-virole.

- les tests de *Culture mathématiques* sont définis comme ne posant pas de questions mathématiques - *c'est vrai* - mais des « problèmes de la vie courante et de la vie active » ce qui est faux car ce sont des caricatures de problèmes de la vie active imaginées par des gens qui n'ont aucune idée de ce qu'est un « problème pratique ».

- d'une manière plus générale les contenus traités par PISA sont en général des compétences et non des connaissances,

- la moitié des questions sont des QCM dont on sait qu'ils sont incapables de mesurer "certains types de compétences telles que l'aptitude à rédiger et à exprimer sa pensée" et favorisent - et valorisent donc puisque ce sont des tests d'évaluation - "une parcellarisation des connaissances".

- ...

L'exception française

Ceci dit, que ce soit dans les grands journaux anglais de droite, le *Times*, ou de gauche comme le *Guardian*, que ce soit en Allemagne, en Suède ou aux USA où une des références sur l'éducation, Diane Ravitch, en a parlé sur son blog et où la plus grande revue éducative, *Education week*, les a évoquées, on trouve dans les médias de ces pays des références aux graves critiques faites à PISA, critiques citées plus haut. Mais si l'on regarde ce qui se passe en France, on peut effectivement parler d'une *exception française* au pire sens du terme puisqu'aucun média n'a évoqué ces graves critiques que l'on peut faire à l'étude PISA, l'ensemble des médias et de l'éventail politique suivant aveuglement les conclusions de l'OCDE. On pourrait tout à fait comprendre que les médias et les chercheurs français ne partagent pas les critiques faites à PISA mais la réalité est qu'ils ne les mentionnent pas, soit par méconnaissance des faits soit par suivisme par rapport à l'OCDE ... soit pour les deux ? Sujet également intéressant à traiter ...

La seule référence en France à une vision non strictement OCDE de PISA dans les grands médias est un article intitulé [**PISA : un palmarès médiatisé mais discutable**](#) dû à Julien Grenet (Chercheur

à l'École d'économie de Paris et directeur adjoint de l'Institut des politiques publiques) paru dans Le Monde du 11 décembre 2013. Mais, à ma connaissance, il n'est ni repris ni même cité dans aucun article de ce même journal ni dans aucun autre média. Et pas plus chez les « chercheurs en éducation ».

Bonne lecture notamment des compléments bibliographiques.

Cabanac, le 23 avril 2014

Michel Delord

ANNEXES

Annexe 1- Intermède : A l'école, se sentir « chez soi » ou se sentir « à sa place » ?

Une erreur de traduction aux conséquences importantes révélée par Loys Bonod dans la note

PISA : à l'école comme "chez moi"

[...]

Cette donnée [sur le « *sentiment d'appartenance à l'école* »] procède pourtant pour la France d'une grossière erreur méthodologique de l'OCDE que nous avons déjà repérée et analysée dans un article de 2012 ([« La souffrance scolaire en France, mythe utile »](#)) : une traduction erronée dans une question posée aux élèves français par PISA 2003 faussait totalement en effet les conclusions qu'on en pouvait tirer sur l'école française. Pour résumer, on demandait aux élèves anglophones s'ils se sentaient « *à leur place* » à l'école (« *School is a place where I feel like I belong* »). Et aux élèves français, s'ils s'y sentaient « *chez eux* ». La question n'étant pas la même, la réponse ne pouvait pas l'être !

La France obtenait alors un score terrifiant, avec seulement 45% des élèves se sentant « *chez eux* » à l'école, quand en moyenne 81% des élèves de l'OCDE se sentaient « *à leur place* » à l'école ! Seule la Belgique, pays en grande partie francophone, descendait presque aussi bas, avec 56% des élèves se sentant « *chez eux* » à l'école.

Cet écart assez extraordinaire aurait dû mettre la puce à l'oreille des experts de l'OCDE. D'autant qu'il n'était absolument pas corroboré par les autres réponses des élèves français ou belges concernant le « *sentiment d'appartenance à l'école* ». Ainsi seuls 13% des élèves français et 16% des élèves belges considéraient que l'école était un endroit où ils se sentaient « *mal à l'aise et pas à [leur] place* » (contre 10% pour la moyenne de l'OCDE).

PISA 2012 : on ne change pas une mauvaise traduction

Hélas. Ce que l'on aurait pu espérer être une erreur ponctuelle vient pourtant de se révéler une erreur invariablement reconduite à chaque session de PISA, sans que personne ne s'inquiète de l'incohérence de ces données.

À l'occasion en effet du dernier « *choc PISA* » de la France en 2013, ce score a de nouveau été mis en exergue par les analystes de l'OCDE pour expliquer la performance médiocre de la France.

Le sentiment d'appartenance s'est quant à lui amélioré : la proportion d'élèves ayant déclaré se sentir comme des étrangers à l'école, mal à l'aise, pas à leur place, ou ne pas se sentir appréciés des autres élèves a diminué (tableau III.2.3f). Cependant, la France se situe toujours bien en dessous de la moyenne de l'OCDE, contrairement au Liechtenstein, à l'Autriche ou à la Suisse, où au moins quatre élèves sur cinq se sentent chez eux à l'école, contre moins d'un élève sur deux en France. La France, avec seulement 47 % des élèves déclarant se sentir chez eux à l'école, affiche la proportion la plus basse de tous les pays et économies ayant participé au cycle PISA 2012. Parmi les pays de l'OCDE, c'est en Corée et en France que le lien entre le sentiment d'appartenance et la performance est le plus fort.

Après vérification, on s'aperçoit que la question posée aux élèves en 2012 est toujours la même que neuf ans plus tôt. La dernière place de la France, de très loin, avec 47% des élèves se sentant « *chez eux* » alors qu'en moyenne 81% des élèves de l'OCDE se sentaient « *à leur place* » à l'école. Comme le reconnaît benoîtement la note de PISA pour la France, cette donnée contredit presque toutes les autres données sur le « *sentiment d'appartenance à l'école* » des élèves français, mais n'occasionne aucune circonspection.

[Suite](#)

*

* *

Annexe 2 - Quelques mots sur l'étude de l'APED : « [PISA 2012 sans fard et sans voile](#) »

Dans le cadre francophone, on peut s'intéresser aussi à cette étude sous-titrée *Pourquoi les systèmes éducatifs de Belgique et de France sont-ils les champions de l'inégalité sociale ?* Nico Hirtt, responsable du service d'étude de l'Apéd, la préface par un texte intitulé « *L'équité bafouée* ».

Lors de la publication de ce texte fin février, j'avais envoyé ce commentaire très court sur le site de l'APED

Salut Nico

Je crois que, ce coup-ci, je ne suis pas tout à fait d'accord avec toi. Va faire un tour sur mon texte « [Vaccination contre le PISA Choc](#) ».

Ça commence par

"L'OCDE, nous a fait cadeau de PISA en 2003 après son cadeau, 40 ans avant, des mathématiques modernes. PISA semble tellement important que certains, au Monde en particulier, voudraient créer un PISA-Choc. La médecine n'est-elle pas dangereuse ?"

Bonne lecture

Michel Delord

Ce qui m'a valu [la réponse suivante](#) de Nico Hirtt:

Cher Michel, ton discours est certes intéressant. Mais j'ai un peu de mal à voir en quoi il répond à notre analyse chiffrée, qui parle de ségrégation sociale et académique. L'as-tu lue ? Ou est-ce juste le mot PISA qui t'irrite ? Je suis bien le dernier à croire que la mesure des compétences par PISA constitue une vérité révélée. Et encore moins à nourrir la plus infime illusion quant aux objectifs de l'OCDE. Je suppose que je n'ai plus à démontrer cela.

Mais PISA n'en reste pas moins une base de données solide et statistiquement significative. Lorsqu'elle nous apprend qu'en France la séparation scolaire des élèves, selon leur origine sociale et leur statut d'immigration, est plus forte qu'ailleurs, ou que les taux de redoublement y sont très élevés et qu'ils sont grandement déterminés par le milieu social, ça n'a rien à voir avec l'idéologie de l'OCDE. Ce sont des faits. Point. Lorsque je montre que ces réalités sont statistiquement corrélées à l'inégalité sociale des performances "telles que mesurées par PISA dans le domaine des mathématiques". C'est encore un fait. Ensuite on peut évidemment analyser ces faits de différentes façons. Mais on ne peut certainement pas se mettre la tête dans le sable en les niant. Ton attitude me fait penser à Bellarmin devant Galilée : cachez cette vérité scientifique que je ne saurais voir....

Nico

8

Je me suis donc senti - en tant que Bellarmin supposé - un peu obligé de répondre et c'est la critique de Nico est une de raisons fondamentales qui m'ont à écrire ce texte car manifestement l'étude de l'APED utilise sans se poser de questions les données non fiables calculées à partir du modèle de Rasch, données fournies par PISA. Si je me trompe, ce qui est possible, il faut m'expliquer en quoi.

A) Je suis donc en désaccord avec ton affirmation « *PISA n'en reste pas moins une base de données solide et statistiquement significative* » puisque 1) Les données utilisées par l'APED ne sont pas « statistiquement significatives » puisqu'elles utilisent le modèle de Rasch 2) Même si on n'utilisait pas le modèle de Rasch mais d'autres modèles par exemple pour calculer les « réponses plausibles », il n'est pas prouvé que l'on arrive à un résultat satisfaisant (Voir les critiques de S. Kreiner, H. Morrison et J. Wuttke).

B) D'autre part, la méthodologie employée par l'APED est encore discutable sur un autre point : elle se sert de la « distance moyenne » entre le premier quartile et le dernier quartile des données pour calculer « *son indice global d'inégalité sociale scolaire* ». Or ce sont justement les valeurs extrêmes qui entraînent le maximum d'erreurs :

3.2 Unfounded models

In PISA, a probabilistic psychological model is used to calibrate item difficulties and to estimate student competences. This model, named after Georg Rasch, is the most elementary incarnation of item response theory. It assumes that the probability of a correct response depends only on the difference of the student's competence value and the item's difficulty value. Robert J. Mislevy³ calls this attempt to "explain problem-solving ability in terms of a single, continuous variable" a "caricature"]based in "19th century psychology". The model does not even admit the possibility that some items are easier

³ Norman Frederiksen, Robert J. Mislevy, Isaac I. Bejar, *Test Theory for A New generation of Tests*, Lawrence Erlbaum Associates, Hillsdale, NJ, page 25.

in one subpopulation than in another. The reason for its usage in PISA is neither theoretic nor empiric, but pragmatic: Only one-dimensional models yield unambiguous rankings.

Taking the Rasch model literally, there is no way to estimate the competence of students who solved all items or none: for them, the test has been too easy or too difficult, respectively. In PISA, this problem is circumvented by enhancing the probability of medium competences through a Bayesian prior, arbitrarily assumed to be a Gaussian. As distributions of achievement and psychometric measures are never Gaussian, this inappropriate prior causes bias in the competence estimates, especially at extreme values. This further undermines statements about "risk groups" with particularly low competence values.

Joachim Wuttke, [Uncertainties and Bias in PISA](#), 2008/07/08, pages 10/11.

C) J'en profite également pour faire remarquer que le raisonnement défendu par l'APED qui tend à opposer « de bonnes données » et « une mauvaise utilisation de ces données à des buts comparatifs » est particulièrement vain justement parce que le mode de calcul erroné de ces données a été justement choisi parce que c'est le seul qui permet des « classements sans équivoques » : « *The reason for Rasch model usage in PISA is neither theoretic nor empiric, but pragmatic: Only one-dimensional models yield unambiguous rankings.* »

D) Je n'ai jamais nié où que ce soit - que l'on me prouve le contraire - « l'existence d'une ségrégation sociale et académique » - et j'ai même plus qu'une forte tendance à y croire - mais je ne pense pas que l'on puisse prouver cette ségrégation à partir des données des enquêtes PISA. Et si l'on base donc une affirmation vraie dans un domaine fondamental sur des raisons fausses, le retour de bâton sera radical et mettra en pièces cette affirmation pourtant vraie et fondamentale.

E) PISA met en place un classement implicitement basé sur une hiérarchie des valeurs que, pour de multiples raisons, je ne partage pas et je ne peux donc pas, en toute cohérence, donner une valeur quelconque à ce classement. Je développerai cela ultérieurement mais, par exemple, accepter le type des problèmes posés dans PISA en *Culture mathématique* en utilisant des statistiques tirées des réponses à ces problèmes est bien s'appuyer sur la hiérarchie des valeurs de PISA et justifier publiquement cette hiérarchie.

F) Je ne pense pas du tout que « *Lorsque l'enquête PISA nous apprend qu'en France [...] les taux de redoublement sont très élevés et qu'ils sont grandement déterminés par le milieu social, ça n'a rien à voir avec l'idéologie de l'OCDE* » car cette affirmation revient à masquer le fait que l'OCDE pourfend les redoublements tout d'abord parce qu'ils coûtent cher, tout comme l'UNESCO pourtant *a priori* pourtant plus « culturelle » qui envisage **explicitement** le redoublement sous sa « *dimension économique tout d'abord [...], sa dimension pédagogique ensuite* » et ce dans des textes qui ne sont pas secondaires puisqu'il s'agit de rien de moins que des [Principes de la planification de l'éducation, Réduire les redoublements : problèmes et stratégie, Étude de l'UNESCO, 1998](#). Ceci était dit encore plus crûment par le gouvernement français du très progressiste premier ministre Pierre Messmer, -Ministre de l'éducation Joseph Fontanet -, qui, lui, n'évoquait même pas de prétextes pédagogiques. Dans le communiqué du conseil des ministres du 16 janvier 1974 intitulé « *Les principes directeurs de la réforme de l'enseignement du second degré* », le chapitre intitulé clairement « *Suppression des redoublements* » se bornait à :

Le recours abusif au redoublement sera énergiquement banni. La fréquence excessive des redoublements d'un taux exceptionnellement élevé en France, comparativement à d'autres pays, est une des plaies majeures de notre système éducatif. Elle provoque un alourdissement notable des effectifs scolaires et corrélativement des charges supplémentaires importantes.

Dans ce cadre, je suis en totale opposition avec le chapitre « Redoublements » de l'étude de l'APED.

J'y lis, par exemple :

On savait depuis longtemps que, du point de vue du progrès des apprentissages de chaque élève, le redoublement est plus souvent contre-productif que productif. Le seul argument en faveur du maintien d'une « menace de redoublement » est sa capacité supposée à « motiver » les élèves.

...

Le constat pour la Fédération Wallonie-Bruxelles s'aggrave encore lorsqu'on observe, sur le graphique suivant (n°23), à quel point les taux de redoublement sont étroitement liés à l'origine sociale des élèves. En FWB, un enfant issu du premier décile socio-économique (les 10% les plus pauvres) a sept chances sur dix d'avoir redoublé au moins une fois à l'âge de 15 ans.

Reprenons. « *On savait depuis longtemps que, du point de vue du progrès des apprentissages de chaque élève, le redoublement est plus souvent contre-productif que productif* »

Je n'ai jamais trouvé un seul texte sérieux prouvant cela. La première faute de raisonnement présente dans cette affirmation est de prétendre - implicitement certes - que le redoublement a le même rôle dans un système scolaire dégradé et dans un qu'il ne l'est pas.

Une deuxième faute est l'omission dans toutes les études sur le sujet de l'influence négative de la propagande contre le redoublement.

Une troisième faute, encore plus grave pour ceux qui prétendent combattre la mercantilisation de l'école et trouvent partout l'influence négative du marché, est de ne jamais mentionner toute forme d'argument tendant à prouver que la question des redoublements a une dimension mercantile. Et que, en fait, *ça coûte trop cher*.

Poursuivons : « *Le seul argument en faveur du maintien d'une « menace de redoublement » est sa capacité supposée à « motiver » les élèves.* ». C'est effectivement un très mauvais argument que je n'ai jamais défendu. Et avancer cet argument sert surtout à ne pas poser une question : *Faut-il faire passer dans la classe supérieure un élève qui ne peut pas suivre ?* Une réponse certes partielle mais en un certain sens vraie est : *Oui, si ça fait cinq ans qu'il a passé de classe en classe supérieure sans avoir les moyens de comprendre ce que l'on y enseignait. Puisque, en ce cas, c'est trop tard et que l'école n'a plus qu'à jouer son rôle fondamental dans la « théorie du camp »⁴.*

Je me permets de telles affirmations en liaison avec ce que j'expliquais déjà dans le texte précédant sur PISA :

Petite remarque sur le redoublement : Entendons-nous bien : le redoublement n'est pas une solution idéale. Mais ce que proposent les diverses autorités académiques pour qu'il n'y ait pas de redoublements - et ce depuis trente ans - est une panoplie de solutions toyotistes "zéro défaut" qui sont en fait pires que le redoublement. ***Et l'administration n'a d'ailleurs organisé aucune étude comparative entre les efficacités pédagogiques respectives du redoublement et des autres solutions : c'est compréhensible puisque elle a atteint son objectif lorsqu'elle a réussi à empêcher le redoublement.***

Poursuivons encore : « *Le constat pour la Fédération Wallonie-Bruxelles s'aggrave encore lorsqu'on observe, sur le graphique suivant (n°23), à quel point les taux de redoublement sont étroitement liés à l'origine sociale des élèves. En FWB, un enfant issu du premier décile socio-économique (les 10% les plus pauvres) a sept chances sur dix d'avoir redoublé au moins une fois à l'âge de 15 ans.* »

Supposons que les enfants des classes défavorisées soient défavorisés – c'est peut-être même leur seule *propriété* -et qu'ils aient par exemple plus de difficultés à faire leurs devoirs car il est peut-être plus difficile pour certains élèves de se concentrer intellectuellement à huit dans une pièce de 12 m² dans laquelle il fait 5°C qu'en étant seul dans des conditions agréables dans sa propre chambre.

Le raisonnement suivant est-il vraiment à bannir ? *Dans la mesure où ils sont défavorisés, les élèves qui ont du mal à se concentrer à 8 dans une pièce de 12 m² chauffée à 5°C et dont on peut penser que leurs résultats scolaires sont bas en liaison avec cette situation, méritent-ils qu'on considère que, tant que ces conditions*

⁴ Sur la « théorie du camp », lire [ICI](http://ici.be) ma réponse à la bonne question de Luc Cédelle : « *Mais pourquoi les gouvernements augmentent-ils la durée de la scolarisation si ce n'est pour instruire ?* »

ne sont pas changées, ils ont droit à autant d'années de redoublement qu'ils le souhaitent car c'est une dépense qui représente un temps supplémentaire que leur accorde l'école ?

*

* *

Annexe 3 - Pour la bonne bouche : les « at-risk groups », une bonne raison sécuritaire de soutenir les statistiques PISA

Revenons donc sur les “risk groups “ - évoqués dans le B) de l'annexe II - que PISA permet de définir grâce à une utilisation *subtile* de l'outil statistique, « at-risk groups » dont voici une première apparition en 2000 dans [PISA 2000 : Overview of the Study](#) :

In Germany, 13% of students only reach the lowest proficiency level and almost 10% are not even proficient at this level. This means that almost one-quarter of young people in Germany can only read at an elementary level (OECD average: 18%). In terms of independent reading and lifelong learning, these students must be regarded as a potentially at-risk group. In countries such as Korea, Finland, Canada, Australia and Sweden, this group is much smaller, at less than 15% of students.

Voici donc les trois pages saisissantes (pages 13 à 15) d'une présentation *Power Point* intitulée « [Pedagogical Issues and Mobile Learning](#) », œuvre de *Ben Bachmair* de l'université de Kassel (Allemagne).

3.1
The PISA deficit group in reading
PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT

Result of the PISA Study:

- There are risk groups
- * Boys
- * from migrant families,
- * from low levelled school
- * lower layer of social stratification and income

3.2 Habitus of a migrant male students

The case of Cyril, a 18 year old boy with a migrant background within the socio-cultural milieu of a hedonistic subculture (Sinus Sociovision):

- Self idealisation as an underdog, but better than the silly Germans;
- To investigate the real German underdogs and to have fun by risks;
- Conflict with the police.

4. The idea of assimilation:
Bringing outside in

Cyril's cultural practice outside

- Cyril uploaded mobile videos.
- Cyril is brought to court because of a video which shows - among other - the harassment of homeless men and of a video with Nazi symbols

On est passé *en trois pages* des caractéristiques des « groups à risque » déterminés par les merveilleuses statistiques PISA sur la compréhension de l'écrit - vous savez bien : justement le domaine où les données PISA incluent le plus de résultats non scientifiques ou proprement aberrants - aux représentants de ces groupes à risques qui - *surprise* - sont ceux qui, entre autres, ont un salaire faible, une famille immigrée et dont un exemple est Cyril qui a des conflits avec la police qui trouve qu'il a mis en ligne une vidéo avec des symboles nazis. Autrement dit les merveilleuses statistiques PISA ont permis de trouver en trois pages celui

qui, par excellence, a tout faux. Cela permet-il d'espérer que l'utilisation systématique de PISA permettra de trouver tous les ados qui ont tout faux ? Et peut-être même mieux, de classer en « groupes à risque » tous les élèves du primaire repérés par des tests qui prédisent le succès aux tests PISA, ce qui permettrait de regrouper tous ces élèves dans des classes à risque qui feraient partie d'écoles à risque que l'on placerait dans des quartiers à risques puisque les quartiers sans risques n'en voudront pas.

Et pendant que ces images radieuses de paix sociale nous envahissent agréablement, des grincheux irréalistes - et plus que probablement laxistes - comme moi voudraient supprimer un tel outil au prétexte qu'il n'est pas scientifique et qu'il ne permet pas de mesurer les connaissances mathématiques ?

Pourvu que Dieu, La République, La Morale, La Loi, AREVA et le PSG, c'est-à-dire **Toutes Les Valeurs qui Méritent des Majuscules** évitent à notre hexagonale Nation un aussi funeste sort.

*

* *

Annexe 5 - Bibliographie : quelques références critiques récentes sur PISA

I) Journaux, revues, blogs

The Times Educational Supplement, London, UK

[TES1] William Stewart, [Pisa rankings are 'utterly wrong'](#), published in TES magazine on 19 July, 2013.

[TES2] William Stewart, [Is PISA fundamentally flawed?](#), published in TESS on 26 July, 2013.

The Guardian, London, UK

[GDIAN] Juliette Jowit, [The Pisa methodology: do its education claims stack up?](#), The Guardian, 03 December, 2013.

BBC, London, UK

[BBC] Ruth Alexander, [How accurate is the Pisa test?](#), 10 December 2013.

Berlingske, Copenhagen, Denmark

[Berl] Ulla Gunge, [Flere lande mistænkes for Pisa-fusk](#), 14 November 2013.

Education Week, Maryland, USA

[EdW] Rick Hess, [7 Reasons I Don't Care About the PISA Results](#), Education Week, December 4, 2013.

Diane Ravitch, USA [Historian, Research Professor at NYU, U.S. Assistant Secretary of Education 1997/2004.]

[DRav] Diane Ravitch, [European Scholars : PISA Rankings are "Utterly Wrong" and "Meaningless"](#), December 17, 2013.

Lire aussi en français, c'était déjà recommandé dans mon précédent texte sur PISA : Diane Ravitch, [A quoi sert l'Education secondaire? Volte-face d'une ministre américaine](#), paru dans *The Nation* du 14 juin 2010 puis [traduit dans Le Monde Diplomatique](#) d'octobre 2010

II) Travaux de recherche récents

Bastian Braams, mathématicien du [Courant Institute of mathematical Sciences](#), New York University et de l'[Emory University](#)

[BB1] Bastian Braams, [Mathematics in the OECD PISA Assessment](#).

[BB1] Bastian Braams, OECD [PISA: Programme for International Student Assessment](#).

Prof. Dr. Elart von Collani, [Mathématiques et stochastiques](#), Julius-Maximilians-Universität Würzburg

[COLL] Elart von Collani, [OECD PISA: An example of Stochastic Illiteracy?](#), 2001

Harvey Goldstein est [professeur de statistiques à l'université de Bristol](#) (GB).

[HG] Harvey Goldstein, [Educational policy: what can PISA tell us?](#), 13 December 2013.

Julien Grenet, [Paris School of Economics](#), Deputy Director, Institut des Politiques Publiques

[JG1] ["Un palmarès médiatisé mais discutable"](#), *Le Monde*, December 11, 2013.

[JG2] “PISA : une enquête bancale ?”, *La Vie des Idées*, February 2008.

Svend Kreiner, disciple de Georg Rasch, est [Professeur émérite au Département de Biostatistiques de l'Université de Copenhague](#), Danemark.

[SK1] Svend Kreiner, [Model fit and robustness? Will the leaning tower collapse? A critical look at the foundation of the PISA project](#), novembre 2012.

[SK2] Svend Kreiner, [Is the foundation under PISA solid? A critical look at the scaling model underlying international comparisons of student attainment](#), 2012.

[SK3] Svend Kreiner, Karl Bang Christensen, [Analyses of Model Fit and Robustness. A New Look at the PISA Scaling Model Underlying Ranking of Countries According to Reading Literacy](#), Psychometrika, April 2014, Volume 79, Issue 2, pp 210-231.

Georges Malaty, Mathématiques, [University of Joensuu](#), Finland,

[Mal] Georges Malaty [PISA Results and School Mathematics in Finland: strengths, weaknesses and future](#), 2006

S. J. Prais

[P] S. J. Prais, [Cautions on OECD's Recent Educational Survey \(PISA\)](#), Oxford Review of Education, Vol. 29, 06/ 2003.

Svein Sjøberg est [Professeur de Sciences de l'Éducation à l'université d'Oslo](#), Norvège

[SvSj] Svein Sjøberg, [PISA: Politics, fundamental problems and intriguing results](#), Recherches en Education N° 14 - 2012.

Joachim Wuttke, professeur de *Computer Science*, Munich, Allemagne.

[JW] Joachim Wuttke, [Uncertainties and Bias in PISA](#), *PISA ACCORDING TO PISA. DOES PISA KEEP WHAT IT PROMISES*, Hopmann, Brinek, Retzl, eds., pp. 241-263, Wien, 2007.

Lire aussi [ICI](#) ses “Publications on educational testing”