

# La formation tout au long de la vie et son double

## Contribution à une critique de l'économie politique de l'efficacité dans l'éducation

---

ROMUALD NORMAND  
UMR Éducation & Politiques (INRP-Lyon 2)  
19 Mail de Fontenay  
BP 17424, 69347 Lyon Cedex 07, France

Depuis quelques années, les pays anglo-saxons sont engagés dans un mouvement profond de restructuration de leurs systèmes éducatifs au nom d'une recherche toujours plus grande de l'efficacité et de la qualité (Normand 2003). L'élévation des standards est devenue un mot d'ordre lancé par les gouvernements conservateurs, avec pour effet un accroissement des exigences d'évaluation vis-à-vis des écoles et l'installation progressive d'un management de la performance (Gleeson & Husbands 2001). Ces conceptions sont reprises aujourd'hui dans les recommandations adressées par la Commission européenne aux différents États membres afin de promouvoir une formation tout au long de la vie (LifeLong Learning). Les instruments d'évaluation auxquels il est fait référence dans les rapports successifs de la Commission possèdent une forte normativité mais ils sont également le produit de connaissances objectivées et transportées dont il paraît intéressant de retracer ici la genèse.

D'abord s'est constituée une science de la mesure de l'efficacité de l'école, dont on trouve les premières traces au XIX<sup>e</sup> siècle, au moment où naissent la psychologie expérimentale et la statistique progressivement appliquées à l'éducation. Par la suite, cette science a contribué à légitimer des classements sur les élèves à partir desquels ont été élaborés des standards au niveau national et international. En France, par exemple, une standardisation et des classifications uniformes ont constitué le dispositif métrologique de l'État centralisé visant à unifier le territoire national et à réduire l'inégalité des chances en éducation (Derouet 1992, 2000; Desrosières 1993). La statistique publique s'enracine également dans des préoccupations de santé

publique et de construction de la citoyenneté politique où les grands corps de l'État ont joué un rôle important. Il en va différemment des pays anglo-saxons où les développements de l'appareillage statistique ont plutôt reposé sur des associations volontaires (Hacking 1990, Porter 1995). Toutefois, cette métrologie est aujourd'hui remise en cause sous l'effet d'un libéralisme normalisateur qui tend à déconstruire les anciennes formes d'équivalence pour en créer de nouvelles à partir de standards donnant la priorité à la mesure de la performance individuelle (Thévenot 1997).

Au-delà de la recherche de la qualité ou de l'efficacité des systèmes éducatifs, les aspects politiques de cette normalisation obligent à considérer la nature même de l'information générée par la mise en œuvre de ces standards au niveau européen. En effet, cette normalisation par les qualités tend à une mise en valeur des activités humaines dans une "cité de l'information" prenant appui sur les nouvelles technologies, des modèles de connaissance distribuée, des figures du réseau et du management qui instituent de nouvelles mises à l'épreuve pour les individus (Boltanski & Chiapello 1999, Thévenot 2003). Les implications d'un tel gouvernement par les normes appellent à un travail de sociologie critique pour rendre compte de l'unité et de la cohérence d'un mouvement mais cette critique doit également se déjouer d'une reprise de ses arguments par un management soucieux de les re-problématiser en nouveaux standards (Derouet 2003, Derouet-Besson 2003, Van Haecht 2003). L'intérêt d'un détour par l'histoire est qu'il permet de situer des concepts, comme efficacité et qualité, dans un champ d'expériences sociales tout en précisant leur impact dans les changements structurels à long terme et dans la construction d'une politique, ses glissements et ses ruptures (Koselleck 2000). De ce point de vue, les États-Unis constituent un terrain particulièrement fertile en raison du leadership qu'ils ont exercé sur les politiques d'éducation à l'échelle mondiale, notamment par leur influence au sein des organisations internationales, mais aussi parce qu'ils offrent un des meilleurs exemples de combinaison entre science, management et politique consacrée à la mesure l'efficacité de l'école.

## Science, management et politique : l'instrumentation de la mesure de l'efficacité de l'école

C'est au début du XX<sup>e</sup> siècle que le test de Binet-Simon devint rapidement une référence dans le domaine de l'éducation même si son succès fut plus important dans le monde anglo-saxon qu'en France. Sa réception fut particulièrement décisive aux États-Unis où le mouvement du "mental testing" contribua à donner des instruments de mesure de l'efficacité de l'école aux responsables de l'éducation, très influencés par la diffusion des thèses du

management scientifique dans l'opinion publique (Callahan 1962). C'est là que s'est opérée une première forme d'alliance entre science, management et politique avec d'un côté le souci des hommes politiques et des administrateurs de collecter des données empiriques sur l'éducation afin de mieux planifier et gérer les écoles américaines et, de l'autre, des chercheurs et des experts intéressés par le développement d'études et l'administration de tests s'appuyant sur les développements de la psychologie scientifique (Travers 1983). Les présupposés sur les races et l'hérédité de cette métrologie par les tests d'intelligence ont fait suffisamment l'objet de discussions au sein de la communauté scientifique pour qu'il soit utile d'y revenir ici, même si cette forme de controverse a ressurgi récemment aux États-Unis (Hermstein & Murray 2002, Gould 1983, Kamin 1974, Sternberg 1985). Il n'empêche que ces tests ont servi de point d'appui à la décision politique et que leur technologie a été utilisée à plusieurs reprises pour résoudre les controverses publiques marquant l'histoire des réformes de l'éducation aux États-Unis (Cremin 1961, Ravitch 2000).

L'une des plus importantes concerne les innovations pédagogiques introduites à Gary, ville industrielle de l'Indiana, dont Dewey a vanté les mérites dans son livre *Les Écoles de demain* (1915). Aucun système d'éducation n'a suscité autant d'intérêt aux États-Unis, plus d'ailleurs pour les économies substantielles qu'il générerait dans l'administration des écoles, que pour la nature nouvelle des relations entre les élèves et les enseignants ou les caractéristiques novatrices des pédagogies mises en œuvre. Toutefois, l'enthousiasme pour ce type d'organisation scolaire retomba fortement quand des techniques de mesure furent mises en œuvre pour évaluer l'efficacité de ces écoles. Les responsables de cette étude ambitieuse, consignée en 1918 dans une série de huit volumes de plus de quatre cents pages chacun, conclurent que les élèves de Gary, d'après leurs résultats aux tests, possédaient des compétences académiques inférieures à la moyenne des élèves américains. Les conclusions du rapport, d'une grande partialité, furent singulièrement biaisées par la façon dont furent évalués les élèves et le type de comparaison conduite à l'époque. Depuis, il a été montré que cette évaluation comportait de fortes résonances politiques et que les évaluateurs, responsables de l'étude, partageaient une même hostilité conservatrice vis-à-vis des innovations pédagogiques. Malgré cela, ce type d'évaluation a servi ensuite de référence pour les écoles américaines.

L'autre controverse est liée aux relations difficiles qu'entretenaient à la fin des années 1920 certaines écoles secondaires, développant des programmes éducatifs relevant de l'Éducation Nouvelle, et les universités américaines soucieuses de maintenir un haut niveau de sélection des étudiants. En avril 1930, la *Progressive Education Association*, mouvement de l'Éducation Nouvelle très influent aux États-Unis, constitua une commission qui lança elle-même un

projet d'évaluation des écoles nouvelles. Une équipe d'experts, dirigée par Ralph Winfred Tyler, conduisit cette évaluation connue aujourd'hui sous le nom célèbre de *Eight-Year-Study*. Cette étude survenait dans la conjoncture de la grande dépression américaine, un temps où l'opinion publique critiquait fortement les banques, le gouvernement américain et les écoles accusées d'une baisse du niveau des élèves. Alors qu'il ruinait les espoirs d'une diversification du curriculum et des pratiques pédagogiques portés par les défenseurs de l'Éducation Nouvelle, la *Eight-Year-Study* eut deux effets majeurs. D'abord, elle laissa en héritage un certain nombre d'instruments qui furent réactualisés et réutilisés au lendemain de la seconde guerre mondiale par les tenants de la pédagogie de la maîtrise dont le chef de file est Bloom. Ensuite, des programmes de tests furent entrepris progressivement par les États américains sur la base du modèle de Tyler qui participa lui-même au développement du *National Assessment of Educational Progress* (NAEP), un système fédéral d'évaluation des compétences des élèves et de comparaison entre États.

## Nouveaux standards et politique de l'efficacité : un modèle américain

Dans les années 1950-1960, en réponse aux suggestions de Tyler, Bloom et son équipe ont produit un système de classement des objectifs pédagogiques couvrant une gamme étendue d'apprentissages et une méthodologie visant à améliorer l'efficacité pédagogique des enseignants (Bloom & al. 1956). Cette théorie, baptisée par la suite "pédagogie de la maîtrise", a constitué le modèle d'évaluation des apprentissages utilisé par l'IEA (*International Association for the Evaluation of Educational Achievement*) pour établir des comparaisons internationales de résultats entre différents pays industrialisés dans les disciplines fondamentales, notamment les mathématiques. Le temps consacré aux apprentissages devenait une source d'investigation en même temps que, sous l'influence de la théorie économique, l'étude entendait se consacrer à une mesure des inputs et des outputs des systèmes éducatifs permettant d'expliquer les différences de performance entre les écoles et les élèves. Les deux économistes les plus influents en ce domaine, Theodore Schultz et Gary Becker, développèrent à la même époque leur théorie du capital humain. Schultz considérait que la théorie économique pouvait mesurer l'investissement et la rentabilité en éducation. À partir d'une étude sur les États-Unis couvrant la période 1900-1956, il montrait qu'il existait un facteur résiduel autre que le travail et le capital pour expliquer la croissance économique : l'amélioration de la qualité de l'éducation et de la recherche. Becker, en liant les revenus des individus à différents niveaux d'éducation

justifiait le fait que l'on pouvait estimer la rentabilité de l'éducation au cours d'une vie entière.

À cette époque, les prévisions d'une explosion des coûts de l'éducation dans les pays industrialisés et les pays en voie de développement s'étaient accompagnées, sous l'auspice des grandes organisations internationales, de l'adoption de stratégies de planification visant à établir des priorités en tenant compte, comme l'affirmaient les discours officiels, des ressources limitées à disposition. Les notions de mesure du rendement constituaient alors une notion commune qui pouvait satisfaire aussi bien les économistes, les psychologues de l'apprentissage, les sociologues, les spécialistes de l'évaluation, que les administrateurs et planificateurs de l'éducation. D'une certaine manière, on peut considérer que la ville de Chicago a constitué le lieu d'une "évaluation-monde", par comparaison avec l'expression d'"économie-monde" forgée par l'historien Fernand Braudel lorsqu'il montre que le bassin méditerranéen a constitué le berceau du capitalisme mondial. En effet, c'est à partir de cette ville américaine que se constitue progressivement un espace d'intéressement, une connaissance commune, des concepts et des instruments partagés dans le champ de l'évaluation de l'éducation, même si des universités comme Harvard, Princeton, ou Berkeley ont pu également jouer un rôle important dans la promotion de la politique des tests aux États-Unis (Leman 1999). Théorie béhavioriste des apprentissages, analyse de la mobilité sociale et de l'orientation professionnelle par les tests, conception réductionniste et néo-classique de l'économie de l'éducation, centration du curriculum sur les résultats scolaires des élèves, constituent les ingrédients à partir desquels se sont élaborées les politiques des trente dernières années dans le monde anglo-saxon.

Dans les années 1960-1970, de nouvelles préoccupations conduisent à faire émerger de nouvelles approches de l'évaluation par les tests aux États-Unis. Deux dates expliquent le développement des programmes d'évaluation de l'éducation : la première est le 4 octobre 1957 quand les Russes lancèrent leur Sputnik et que l'Amérique tout entière accusa le coup de son retard technologique ; la seconde est 1965 et correspond au lancement du Titre 1 de l'*Elementary and Secondary Education Act*, un programme de discrimination positive destiné aux enfants les plus défavorisés. Alors que le combat contre la pauvreté, l'injustice raciale et l'inégalité des chances recevait une résonance politique, l'administration Johnson et le Congrès américain chargèrent James Coleman d'une enquête nationale sur l'égalité des chances entre élèves. En utilisant des tests standardisés dans le domaine de la lecture, des mathématiques, et de l'évaluation des compétences verbales et non-verbales, le rapport montra des différences importantes entre les élèves blancs et ceux d'autres minorités ethniques ou raciales. Mais ce qui surprit le plus les commentateurs fut que la qualité des écoles, souvent évaluée en termes d'accès aux ressources et aux équipements, n'entretenait aucune relation

avec les résultats des élèves mesurés par des tests. Le rapport Coleman eut dans un premier temps un impact limité et il serait resté dans l'ombre sans une série d'événements qui a focalisé l'attention sur ces conclusions quelques années après. En janvier 1969, en effet, une administration beaucoup plus conservatrice arrive au pouvoir avec de nouvelles idées sur les réformes jugées nécessaires en éducation. Elle reprend à son compte les conclusions du rapport Coleman et le président Nixon lui-même les utilise en mars 1970 pour annoncer son programme pour l'éducation. C'est une période de crise financière où l'administration américaine doit arbitrer entre les engagements de la Grande Société et les pressions dues à une guerre au Vietnam de plus en plus coûteuse. Celle-ci voit dans l'éducation un domaine où elle peut réduire les dépenses publiques, d'autant plus que le rapport Coleman vient conforter des évaluations négatives des programmes de discrimination positive.

La critique ressurgit quelques années plus tard à l'occasion de la parution du rapport *A Nation At Risk* de la *National Commission on Excellence in Education* concluant que le système éducatif américain est devenu médiocre et qu'il est responsable des bas niveaux de compétences des élèves dans les disciplines fondamentales (*A Nation at Risk* 1983). Immédiatement après sa publication, un ensemble de décisions politiques a été pris à l'échelle des États américains, particulièrement ceux du Sud, en référence à un "mouvement pour l'excellence" destiné à élever les standards de performance pour les élèves et les enseignants. Ce mouvement, partisan d'une restructuration des écoles pour améliorer leur efficacité, impliquait fortement des associations et des éducateurs à l'échelle des districts, particulièrement pour proposer de nouvelles formes de planification et de management des écoles. L'idée était d'accorder une plus grande autonomie aux établissements, en contrepartie d'une exigence plus forte en termes de résultats (accountability). À cette époque, le management des entreprises s'orientait vers davantage de flexibilité, vers le travail en équipe autour de projets, vers l'adoption de procédures d'auto-évaluation. Ces conceptions furent transposées dans le management des écoles en même temps que se mettait en place une politique de l'excellence visant à définir des standards pour les contenus enseignés (content standards) et les résultats des élèves (performance standards). Ces standards reposent principalement sur l'administration de tests commerciaux où les questions à choix multiple prédominent. La plupart des États américains ont adopté ces standards de contenus et de performances, conçus par les associations disciplinaires dans différentes disciplines et, parfois, fortement liés à la certification et à la rémunération des enseignants.

Des recherches ont pu montrer que le fait d'attacher trop d'importance aux tests pouvait conduire à des effets pervers (Mc Neil 2000, Orfield & Kornhaber 2001, Sacks 1999). Pourtant cela n'a pas empêché le président Bush de mettre en place son projet "No Child Left Behind" en exigeant

un renforcement des procédures de tests et une élévation des standards en mathématiques et en lecture. Cette politique a reçu le soutien des républicains comme des démocrates selon une même conviction : les récompenses et les pénalités aux tests peuvent contribuer à une amélioration de l'éducation. Des travaux assez critiques ont cependant contesté l'idée que cette politique pouvait améliorer la qualité des écoles et montré que les États ayant mis en œuvre des programmes de tests faisaient parfois pire que ceux n'ayant pas développé de tels programmes (Haney 2000, Sacks 1999). Les effets de ce renforcement du testing semblent aussi particulièrement négatifs sur les élèves des minorités ethniques et ceux des milieux populaires (Goodson & Foote 2001, Haney 2000, Heubert & Hauser 1999, Klein Hamilton, McCaffrey & Stecher 2000, Kohn 2000, Sacks 1999). De même, le rétrécissement du curriculum tout comme la réduction des apprentissages et de l'éducation pour un entraînement des élèves aux tests peut avoir des effets négatifs, notamment pour les élèves les plus déshérités (Berliner & Amrein 2002).

## **De la planification de l'éducation aux indicateurs de comparaison des systèmes éducatifs : l'internationalisation du modèle de l'efficacité**

On a vu précédemment comment les recherches sur l'efficacité de l'école, en mobilisant la psychologie expérimentale et les instruments de la statistique, étaient progressivement parvenues à imposer une définition légitime de la qualité de l'école susceptible d'orienter les politiques d'éducation aux États-Unis comme dans d'autres pays anglo-saxons. Ces modèles de l'efficacité présupposent que la qualité est mesurable à partir de différentes variables et d'une collecte de données exprimées ensuite sous la forme d'inputs et d'outputs que les techniques modernes de l'économétrie permettent de faire émerger (Angus 1993). Cette modélisation est au cœur de la construction sociale et politique d'indicateurs de comparaison des résultats scolaires aux échelles nationale et internationale ainsi que des classifications des compétences et des savoirs qui en résultent (Nuttall 1994). La connaissance produite n'est pas neutre au sens où elle oriente de manière décisive les politiques éducatives adoptées dans chaque pays et qu'elle génère un modèle standard auquel les États cherchent à se conformer, avec plus ou moins de succès. Même si ces comparaisons internationales sont critiquables, elles véhiculent l'idée que la qualité des systèmes éducatifs peut s'évaluer en fonction des ressemblances et des différences de réussite scolaire des élèves, indépendamment du contexte social et politique dans lequel l'enseignement est dispensé. Leur finalité politique est pourtant tout autre puisque ces indicateurs constituent

avant tout des repères de qualité servant à appuyer les décisions politiques dans une gestion rationnelle et maîtrisée de la dépense publique d'éducation. D'une certaine manière, la justification des indicateurs internationaux tend à reprendre les arguments planistes des années 1950-1960, époque où on affirmait que la planification de l'enseignement, conjuguée à la maîtrise d'indicateurs sociaux, permettrait d'accéder au bonheur et à la qualité de la vie pour le plus grand nombre (OCDE 1973).

Toujours à la fin des années 1950, des chercheurs provenant de dix pays se rencontraient à l'Institut de l'Éducation de l'UNESCO à Hambourg pour échanger des informations et leur expérience sur l'échec scolaire, les examens et les méthodes d'évaluation. L'absence d'instruments d'évaluation à l'échelle internationale déboucha sur une réflexion concernant la définition nécessaire de critères internationaux de mesure permettant d'évaluer de manière uniforme les pratiques d'éducation et les standards des différents pays. À partir d'une étude de faisabilité conduite de 1959 à 1961, fut mis en œuvre "l'International Project for the Evaluation of Educational Achievement" (IEA) qui lança les premières comparaisons internationales de réussite dans le domaine des mathématiques. Parallèlement, l'OCDE (à l'époque OECE) lançait ses premières tentatives de comparaison internationale fondée sur des méthodes de prévision des besoins en personnel scientifique et technique dans le cadre de la planification de l'éducation (Padadhopoulos 1994). L'obtention d'un tableau comparé du fonctionnement des systèmes éducatifs a jeté les bases de travaux plus systématiques menés ultérieurement par l'OCDE. Au début des années 1970, la planification de l'éducation était envisagée en lien avec la mise en œuvre des politiques éducatives, notamment par le développement d'indicateurs sociaux permettant de contrôler le fonctionnement des systèmes scolaires, de procéder à des analyses coût/efficacité, d'évaluer les résultats de l'enseignement et de l'apprentissage. Cet intérêt pour les indicateurs sociaux relevait de travaux conduits dans le cadre de la comptabilité nationale et du PPBS (Programming Planning Budgeting System), principaux instruments de la planification. Au sein du Comité de l'éducation de l'OCDE, un groupe de travail sur les statistiques et les indicateurs était parvenu à présenter un cadre statistique permettant de mesurer l'incidence de l'enseignement sur la société, à travers plusieurs indicateurs de performance concernant aussi bien la transmission du savoir que l'égalité des chances et la mobilité sociale, la satisfaction des besoins de l'économie, le développement individuel et la transmission des valeurs. Toutefois, il a fallu attendre une dizaine d'années pour concrétiser ce projet parce que les pays ne s'accordaient pas sur les objectifs à atteindre et ne disposaient pas d'un système statistique permettant de produire ce type de mesure.

Au commencement des années 1980, les principes fondamentaux des politiques d'éducation n'étaient plus les mêmes en raison de la crise éco-



nomique mettant à mal le système de planification lentement élaboré par l'OCDE. Le gouvernement des États-Unis exerçait alors une forte pression pour obtenir que la qualité devienne un thème prioritaire dans l'activité de l'organisme après la publication du rapport *A Nation at Risk*. Les responsables américains souhaitaient aussi que l'on mette en œuvre une statistique s'inspirant de l'expérience du "National Assessment of Educational Progress" (NAEP). Parallèlement, cette pression était liée au développement des comparaisons internationales dans un climat marqué par la concurrence où l'IEA jouait un grand rôle. La définition de la qualité donna lieu à des débats virulents entre les ministres de l'éducation des pays membres de l'OCDE. Le clan des conservateurs, dans lequel on trouvait la Grande-Bretagne et les États-Unis, souhaitait renforcer les procédures d'évaluation des élèves et des enseignants, en se rapprochant d'un modèle concurrentiel jugé plus proche des consommateurs. Les opposants défendaient l'idée que tous les objectifs de l'enseignement n'étaient pas quantifiables et que les examens ne devaient pas jouer un rôle excessif au détriment d'autres aspects de l'apprentissage scolaire. Malgré ces divergences, les ministres finirent par se mettre d'accord sur l'idée que le thème de la qualité de l'enseignement devait être l'une des priorités dans l'activité ultérieure de l'OCDE, conformément à la thèse américaine. À partir de là, l'organisme a procédé à un travail de clarification du concept de qualité notamment à travers la mise en place d'indicateurs. En 1992, l'OCDE publiait son premier ensemble d'indicateurs de l'enseignement destiné à orienter les politiques éducatives des pays membres et à fournir une base de comparaisons internationales dans le sens d'une plus grande qualité et efficacité des systèmes éducatifs.

## Harmonisation des statistiques et construction d'indicateurs internationaux : vers une convergence des politiques d'éducation en Europe ?

À la suite de la publication de *Regards sur l'Éducation*, l'OCDE a proposé de lancer un projet plus ambitieux d'indicateurs internationaux de l'enseignement afin non seulement de continuer le travail de comparaison mais aussi pour orienter les politiques éducatives des États membres. Le NCES américain (National Center for Education Statistics) a joué un rôle actif aux côtés de l'OCDE pour mettre en œuvre ce projet visant à améliorer l'efficacité des systèmes éducatifs (Bottani & Tuijnman 1994). Le projet INES vise à développer un système d'indicateurs de comparaisons internationales permettant la coopération et l'échange d'informations entre les responsables des politiques d'éducation et la promotion des méthodologies de l'évaluation

fondées sur des critères objectifs et valides (Hutmacher, Cochrane & Bottani 2001). La philosophie sociale du projet INES est claire : elle vise selon ses promoteurs à mesurer les inégalités et à aider les citoyens à juger par eux-mêmes de la qualité de leur système éducatif notamment en éclairant l'opinion publique sur la base de critères objectifs fournis par les comparaisons internationales. Toutefois, l'utilisation des comparaisons internationales de résultats en éducation soulève bon nombre de débats parmi les spécialistes (Broadfoot & al. 2000, Mc Lean 1992). En effet, l'approche des comparaisons internationales de résultats présuppose qu'il existe une explication unilatérale des problèmes alors que les attentes en termes de réussite scolaire et de participation à l'enseignement doivent nécessairement être mises en relation avec les conceptions de l'égalité des chances et de la mobilité sociale propre à chaque pays, la nature de la sélection des élites, la conception égalitariste ou individualiste de l'éducation, la nature des contenus enseignés, les méthodes pédagogiques, etc. Ces comparaisons internationales de résultats soulèvent donc des problèmes de justesse et de justice.

De justesse parce qu'il est souvent difficile de distinguer les différences réelles de performance entre pays et celles qui sont dues aux faiblesses de la méthodologie utilisée. On rencontre inévitablement des problèmes de comparabilité des échantillons, de conception des tests, d'analyse des données et de situations d'apprentissage. La façon dont on peut défendre les comparaisons internationales au regard de ces critères détermine la validité et l'utilité de ces données. Dans beaucoup d'études, en dépit des efforts pour faire correspondre les échantillons d'élèves en termes d'âge, de sexe et de caractéristiques socio-économiques, des différences significatives demeurent comme, par exemple, le nombre d'élèves ayant redoublé dans les différents pays ou ceux relevant de l'éducation spécialisée qui sont inclus ou non dans l'échantillon. Stedman (1997) a montré que la réussite scolaire des élèves aux États-Unis était représentée de manière injuste dans les comparaisons internationales parce que les résultats pour l'ensemble d'une cohorte étaient comparés à ceux d'une petite élite dans d'autres pays. La motivation des élèves et leur attitude envers les tests ont été identifiées comme un facteur important des différences de performance entre pays. Des problèmes de biais culturels se retrouvent dans des items de toutes sortes. Ils sont maintenant bien connus dans la littérature sur le testing mais n'ont pas été pris en compte dans les comparaisons internationales de résultats. Ainsi à l'apparente solidité de la mesure des résultats ne correspond pas toujours une même solidité dans la signification des tests utilisés.

Plus encore, il est démontré que de mauvaises interprétations des comparaisons internationales pouvaient avoir des effets politiques redoutables comme l'ont illustré les débats en Grande-Bretagne suite à la publication de l'étude TIMSS quand il a été démontré que les élèves britanniques réus-

issaient moins bien en mathématiques que leurs pairs asiatiques ou de l'Est de l'Europe. Pourtant, l'étude de O'Leary, Madaus & Kellaghan (citée par Broadfoot & al. 2000), consistant à comparer les résultats de deux études internationales en mathématiques et en sciences, montre des différences significatives dans les rangs obtenus par certains pays aux deux évaluations. Non seulement ces différences rendent difficiles la mise en œuvre d'actions spécifiques pour les politiques d'éducation, mais elles interrogent sur la fiabilité des données elles-mêmes et l'utilité d'un tel exercice. L'étude soulève des questions sérieuses sur l'équité et l'utilité des comparaisons internationales des standards éducatifs, notamment quand elles sont fondées sur la comparaison de réussite à des tests (Brown 1998, Goldstein 1995). Enfin, la persistance d'une inégalité des chances pose un problème important aux promoteurs des procédures d'évaluation par les tests de performance, particulièrement quand ils cherchent à promouvoir une égalité de résultats. De nombreuses recherches ont montré que ces évaluations défavorisaient les élèves des groupes minoritaires mais aussi les filles par rapport aux garçons. Par conséquent, ces procédures tendent davantage à rationaliser les inégalités qu'à les soulager. De fait, certains groupes sociaux sont désavantagés par les évaluations nationales ou internationales parce qu'ils ont des capacités différentes devant les épreuves imposées. C'est pourquoi celles-ci sont productrices d'inégalités aussi bien avant qu'après la passation du test.

Les données statistiques lentement construites par les comparaisons internationales sur le modèle de la "school effectiveness" constituent aujourd'hui l'architecture selon laquelle s'élaborent de nouvelles formes de classification standardisées des activités d'éducation et de formation en Europe. L'établissement de nouvelles catégories, de classements, de nouveaux standards n'est pas neutre puisqu'il contribue à construire des informations de portée générale possédant des effets politiques non négligeables (Bowker & Star 1994). Cette information statistique utilise les acquis des recherches sur l'efficacité de l'école et circule le long de réseaux d'expertise participant à l'instauration d'une forme puissante d'objectivité donnant ainsi une assise certaine à une logique d'inspiration néo-libérale. Ces agrégats statistiques, autrefois instruments de mesure des inégalités, s'orientent aujourd'hui vers la prise en compte de situations individuelles non plus selon une logique administrative mais selon un souci de rentabilité, de flexibilité et de maîtrise des coûts de la formation. Cette normalisation statistique participe également d'une certaine forme de convergence des systèmes et des politiques d'éducation et de formation au niveau européen, elles-mêmes nécessaires à l'alimentation d'un vaste marché du travail où l'accroissement de la mobilité de la main d'œuvre constitue un des objectifs prioritaires de la Commission européenne. Cette dernière juge indispensable le développement d'informations statistiques comparables pour la mise en œuvre des politiques d'éducation et

de formation tout au long de la vie. Pour cela, elle a contribué à promouvoir des indicateurs et des objectifs quantifiés en s'appuyant sur l'expertise conjointe de l'UNESCO, de l'OCDE et d'EUROSTAT (Commission européenne 2000a). Une "task force" a été chargée de concevoir des indicateurs communs aux pays européens concernant la qualité de l'éducation et de la formation tout au long de la vie.

Ces statistiques et ces indicateurs s'intègrent à la méthode ouverte de coordination prévue lors de la conférence de Lisbonne à la suite de laquelle la Commission européenne a présenté son Mémorandum sur l'éducation et la formation tout au long de la vie (Commission européenne 2000b). Ce Mémorandum s'est accompagné d'un rapport sur des "objectifs concrets" invitant tous les États membres à améliorer la qualité de leur système d'éducation et de formation dans le domaine des apprentissages et de la pédagogie (Commission européenne 2001a). L'élévation des standards est présentée comme déterminante pour l'Europe si elle veut devenir une société plus compétitive et plus dynamique. D'après la Commission européenne, cette recherche de l'efficacité doit permettre aux citoyens de mieux développer leurs capacités et leurs compétences. Ces indicateurs d'efficacité et de qualité doivent faciliter une meilleure compréhension des systèmes nationaux et locaux ainsi que leur comparaison, l'identification des politiques "performantes", ainsi que l'échange, la connaissance et la diffusion des "bonnes pratiques" (Commission européenne, 2001b). Ils doivent permettre d'assurer un meilleur contrôle de l'investissement en éducation et de l'utilisation des ressources en termes de coût-efficacité des politiques éducatives, leur construction étant directement inspirée des grandes comparaisons internationales de résultats conduites par l'IEA. Les experts travaillant à leur élaboration appartiennent aux services statistiques de l'OCDE ou sont impliqués dans le projet INES.

## Conclusion

Le nouveau système de collecte de l'information statistique à l'échelle européenne est accompagné par un mouvement conjoint de normalisation de la qualité qui s'étend au domaine de la formation professionnelle et de l'enseignement supérieur. L'assurance-qualité constitue aujourd'hui le référent de la Commission européenne qui tend à promouvoir les expériences et les projets pilote intégrant des modes d'évaluation de la qualité dans l'éducation et la formation. Ces nouveaux standards se substituent aux procédures d'agrément et aux contrôles réglementaires jusqu'à présent assurés par les États. Les procédures de contrôle-qualité confèrent une légitimité et un pouvoir de plus en plus important à des agences chargées de définir en leur sein des critères techniques, largement inspirés des conceptions modernes du

management. Au nom d'une meilleure "gouvernance" des systèmes d'éducation et de formation, cette logique de la qualité tend à déstabiliser les cadres nationaux tout en confiant des responsabilités accrues aux collectivités régionales et à des réseaux transnationaux d'experts et d'évaluateurs sans grande légitimité démocratique (Charlier 2003, Javeau 2003).

Ces défenseurs des évaluations par les tests et les standards avancent souvent l'idée que ceux-ci servent à promouvoir l'égalité des chances en améliorant la distribution des biens comme les salaires, les promotions, l'accès aux cours les plus prestigieux, les diplômes des grandes écoles. En termes de justice, on trouve là un mélange confus de libéralisme et de méritocratie, d'un côté la promotion du choix du consommateur ou de l'opinion éclairée, de l'autre un contrôle technocratique au nom de principes méritocratiques (Howe 1997). Toutefois, les tests sont incapables de prendre en compte les inégalités entre élèves à l'intérieur comme à l'extérieur de l'école. Ceux-ci ne bénéficient donc pas de la même égalité des chances sauf dans un sens très formel et très abstrait, parfois défendu par des conceptions utilitaristes de la justice, mais qui ne tiennent pas compte des épreuves de réalité affrontées par les élèves et les enseignants. En fait, cette idéologie tend à récuser l'idée d'un meilleur partage des bénéfices permettant la réduction de l'écart entre les plus et les moins favorisés, la technologie des tests servant davantage à éloigner les ressources éducatives des plus marginalisés et des plus pauvres en favorisant ceux qui ont la chance de devenir de bons ingénieurs et de bons scientifiques.

## Bibliographie

- AMREIN A.L., BERLINER D.C. 2002 "High-Stakes Testing, Uncertainty, and Student Learning", *Education Policy Analysis Archives*, 10(18)
- A NATION AT RISK: THE IMPERATIVE FOR EDUCATIONAL REFORM 1983, États-Unis, Washington DC, Government Printing Office
- ANGUS L. 1993 "The sociology of school effectiveness", *British Journal of Sociology of Education*, 14(3), 333-345
- BALL S.J. 2001 Performativities and Fabrications in the Education Economy, in *The Performing School. Managing, Teaching and Learning in a Performance Culture*, Routledge Falmer, London
- BLOOM J.S., ENGELDHART M.D. & FURST E.J. 1956 *Taxonomy of educational objectives: the Classification of Educational Goals*, Handbook I., New York, David Mc Kay Company
- BOTTANI N. & TUIJNMAN A. 1994 Les indicateurs internationaux de l'enseignement: cadre, élaboration, interprétation, in *Évaluer l'enseignement. De l'utilité des indicateurs internationaux*, Paris, CERI-OCDE
- BOWKER G. & STAR S.L. 1994 Knowledge and Infrastructure in International Information Management. Problems of Classification and Coding, in Bud-Friedman L. (ed.),

- Information Acumen. The Understanding of Knowledge in Modern Business*, London, Routledge, 187-213
- BOLTANSKI L. & CHIAPELLO E. 1999 *Le nouvel esprit du capitalisme*, Paris, Gallimard
- BROADFOOT P., OSBORN M., PANEL C. & SHARPE K. 2000 *Promoting Quality in Learning. Does England Have the Answer? Findings from the Quest Project*, London, Cassell
- BROWN M. 1998 The Tyranny of the International Horse Race, in Slee R., Weiner G. & Tomlinson S. (eds), *School effectiveness for Whom? Challenges to the School Effectiveness and School Improvement Movements*, London, Falmer Press
- CALLAHAN R.E. 1962 *Education and The Cult of Efficiency*, Chicago, The University of Chicago Press
- CHARLIER J.-E. 2003 Du Sud au Nord : que peut apporter l'analyse des systèmes éducatifs africains à la sociologie de l'éducation ?, in Van Haecht A. (dir.) *Sociologie, politique et critique en éducation, Revue de l'Institut de Sociologie*, Bruxelles, ULB
- COMMISSION EUROPÉENNE 2000a Direction Générale de l'Éducation et de la Culture, *Rapport européen sur la qualité de l'éducation. Seize indicateurs de qualité*
- COMMISSION EUROPÉENNE 2000b Document de travail des services de la commission, *Memorandum sur l'Éducation et la Formation tout au long de la vie*
- COMMISSION EUROPÉENNE 2001a Rapport de la commission. *Les objectifs futurs concrets systèmes éducatifs*
- COMMISSION EUROPÉENNE 2001b Direction Générale de l'Éducation et de la Culture. *Réaliser un espace européen de l'éducation et de la formation tout au long de la vie*. Document de travail des services de la commission. Les pratiques d'éducation et de formation tout au long de la vie et les indicateurs, novembre
- CREMIN L.-A. 1961 *The transformation of the School : Progressivism in American Education, 1876-1957*, New York, Knopf
- DEROUET J.-L. 1992 *École et justice. De l'égalité des chances aux compromis locaux?*, Paris, Métailié
- DEROUET J.-L. 2000 "Administration, sciences de l'administration et sociologie en éducation. Quelques exercices de traduction", *Revue de l'Institut de Sociologie*, 1997 1-4
- DEROUET J.-L. 2003 La sociologie des inégalités d'éducation dans les sociétés postmodernes. Petit guide à l'usage des sociologues pour explorer le pays de la pensée gnan-gnan, in Van Haecht A. (dir.) *Sociologie, politique et critique en éducation, Revue de l'Institut de Sociologie*, Bruxelles, ULB
- DEROUET-BESSON M.-C. 2003 Questions à la sociologie politique de l'éducation : peut-on maintenir l'unité du projet sociologique entre les sciences cognitives et le management ?, in Van Haecht A. (dir.) *Sociologie, politique et critique en éducation, Revue de l'Institut de Sociologie*, Bruxelles, ULB
- DESROSIÈRES A. 1993 *La politique des grands nombres*, Paris, La Découverte
- GLEESON D. & HUSBANDS C.T. 2001 *The performing school : managing, teaching and learning in a performance culture*, London, Routledge
- GOLDSTEIN H. 1995 *Interpreting International Comparisons of Student Achievement*, Paris, UNESCO
- GOODSON I. & FOOTE M. 2001 "Testing times : A School case study", *Education Policy Analysis Archives*, 9(2)
- GOULD S.J. 1983 *La Mal-mesure de l'homme : l'intelligence sous la toise des savants*, Paris, Ramsay

- HACKING I. 1990 *The Taming of Chance*, Cambridge U.K., Cambridge University Press
- HANEY W. 2000 "The myth of the Texas miracle in Éducation", *Education Analysis Policy Archives*, 8(41)
- HERMSTEIN R. & MURRAY C. 2002 *The bell curve : intelligence and class structure in American life*, New York, Simon & Schuster
- HEUBERT J.P. & HAUSER R.M. (eds.) 1999 *High stakes : Testing for tracking, promotion and graduation*, Washington DC, National Academy Press
- HUTMACHER W., COCHRANE D. & BOTTANI N. 2001 *In pursuit of equity in education. Using international indicators to compare equity policies*, London, Kluwer Academic Publishers
- JAVEAU C. 2000 Le savant et le politique revisités : les rapports d'aujourd'hui entre les champs politique et scientifique, in Van Haecht A. (dir.) *Sociologie, politique et critique en éducation*, *Revue de l'Institut de Sociologie*, Bruxelles, ULB
- KAMIN L. 1974 *The science and politics of I. Q. : complex human behaviour*, New York, Lawrence Erlbaum
- KLEIN S.P., HAMILTON L.S. MCCAFFREY D.F. & STECHER B.M. 2000 "What do test scores in Texas tell us?", *Educational Policy Analysis Archives*, 8(49)
- KOHN A. 2000 *The case against standardized testing : Raising the scores, ruining the schools*, Portsmouth, N.H., Heinemann
- KOSELLECK R. 2000 *Le futur passé. Contribution à la sémantique des temps historiques*, Paris, EHESS
- LEMAN N. 1999 *The Big Test. The Secret History of the American Meritocracy*, New York, Farrar, Strauss & Giroux
- MC LEAN M. 1992 *The Promise and Perils of Educational Comparison*, London, The Tufnell Press
- MC NEIL L.M. 2000 *Contradictions of school reform*, New York, Routledge
- NORMAND R. 2003 Le mouvement de la "School effectiveness" et sa critique dans le monde anglo-saxon, in Van Haecht A. (dir.) *Sociologie, politique et critique en éducation*, *Revue de l'Institut de Sociologie*, Bruxelles, ULB
- NUTTAL D. 1994 Choosing Indicators, in Riley A.K. & Nuttall D.L. (ed.) *Measuring Quality : Education Indicators-United Kingdom and International Perspectives*, London, The Falmer Press
- ORFIELD G. & KORNHABER M.L. 2001 *Raising standards or raising barriers? Inequality and high-stakes testing in public education*, New York, The Century Foundation Press
- PAPADOPOULOS G.S. 1994 *L'OCDE face à l'éducation 1960-1990*, Paris, OCDE, 1994
- PORTER T.M. 1995 *Trust in Numbers. The Pursuit of Objectivity in Science and Public Life*, Princeton N.J., Princeton University Press
- RAVITCH D. 2000 *Left Back. A Century of Battles Over School Reform*, New York, Touchstone
- SACKS P. 1999 *Standardized minds : The high price of America's testing culture and what we can do to change it*, Cambridge, MA, Perseus Books
- STEDMAN L.-C. 1997 "International achievement differences : an assessment of a new perspective", *Educational Researcher*, 26(3), 4-15
- STERNBERG R.-J. 1985 *Beyond IQ : a triarchic theory of human intelligence*, New York, Cambridge University Press

- THÉVENOT L. 1997 Un gouvernement par les normes. Pratiques et politiques des formats d'information, in Conein B. & Thévenot L. (dir.), *Cognition et information en société*, Paris, série "Raisons Pratiques" 8, Paris, EHESS
- THÉVENOT L. 2003 La mise en place d'un gouvernement par les normes à l'échelle européenne, in Derouet J.-L., Begyn F. & Normand R. (coord.), *L'Europe de l'éducation : entre management et politique*, Actes de l'université des personnels de l'encadrement de l'Éducation nationale, Clermont-Ferrand, (à paraître)
- TRAVERS R.M.W. 1983 *How research has changed American schools? : a history from 1840 to the present*, Kalamazoo, Michigan, Mythos Press
- VAN HAECHT A. 2003 Les politiques publiques d'éducation : un renouvellement nécessaire des outils théoriques, in Van Haecht A. (dir.) *Sociologie, politique et critique en éducation*, *Revue de l'Institut de Sociologie*, Bruxelles, ULB