

# Qualité de l'école un contrôle interne et externe pour éviter toute simplification

**Roger STANDAERT**

*Titulaire d'un doctorat en sciences de l'éducation de l'université de Louvain,  
Professeur d'éducation comparée à temps partiel à l'université de Gand,  
Directeur du développement éducatif au ministère de l'Éducation  
de la communauté flamande de Belgique*

*Traduit de l'anglais par Martine FAUCHER*

**Cet article présente les raisons pour lesquelles l'interaction entre l'auto-évaluation interne et le contrôle externe des établissements scolaires se révèle être l'approche la plus prometteuse pour parvenir à la qualité de l'École. Les arguments qui conduisent à cette conclusion s'appuient sur un modèle didactico-pédagogique qui s'oppose au modèle technico-économique. Les inconvénients de l'approche technico-économique dominante sont illustrés au travers de la description de la situation actuelle en Angleterre qui est entièrement fondée sur ce modèle. Dans le courant de cette analyse, les possibilités et la plupart des limites que présentent les examens nationaux et les palmarès comparatifs qui les accompagnent sont systématiquement mises en évidence.**

## **L'APPEL À LA QUALITÉ : DES OBJECTIFS, DES STRATÉGIES ET DE L'EFFICACITÉ**

Au cours des dernières décennies, l'éducation s'est de plus en plus retrouvée au centre de certaines réflexions sur la qualité. La notion de qualité est devenue omniprésente dans les discussions sur l'éducation et son renouveau.

Le contrôle de qualité suppose que l'on ait une idée de ce que qualité veut dire. La qualité de l'éducation dépend de la qualité de trois aspects importants : des objectifs utiles, des stratégies efficaces et un rendement suffisant en termes de satisfaction des objectifs (efficacité du rendement). La gestion de la qualité a effectivement à voir avec la gestion de ces trois aspects.

Premièrement, elle doit concerner des objectifs qui sont culturels, mais qui sont également jugés utiles par les consommateurs de l'école. Dans ce cas, le terme « consommateur » recouvre un vaste éventail de parties concernées : tout d'abord les élèves, mais aussi les parents, le commerce et l'industrie, la société, la communauté locale, différents segments de la société, les enseignants. C'est ce qui explique que la détermination d'objectifs utiles ne soit pas un processus simple et automatique. Les

diverses parties concernées ont toutes un ou des points de vue, plus ou moins tranchés, sur ce qu'il est utile d'apprendre à l'école. Il est normal, dans une société qui se veut démocratique, que chacun de ces groupes souhaite la mise en œuvre de son propre point de vue sur les objectifs utiles en matière d'éducation. Les divers points de vue vont, à un moment donné, converger dans une sorte d'équilibre dynamique, qui va être exprimé sous forme d'un programme scolaire. Le programme scolaire adopté est ensuite divisé en niveaux et en types d'enseignement, en prenant en compte l'âge des élèves et l'objectif final recherché. Dans quasiment tous les pays, il existe un programme scolaire établi par le gouvernement. Le gouvernement ou le parlement adopte ce programme (en général après de longues discussions et débats), parce que l'existence d'un système d'éducation transparent, avec un niveau comparable dans toutes les écoles et établissements, fait partie de la responsabilité centrale, la responsabilité d'assurer un enseignement obligatoire qui soit utile et juste.

Le deuxième aspect de la qualité concerne la manière dont le contenu qui a été adopté, les compétences et les attitudes sont enseignés. Cela touche aux procédés, aux stratégies, aux méthodes et aux structures d'organisation. Les thèmes peuvent être en rapport avec la didactique, la méthodologie, le regroupement des élèves, l'organisation scolaire ou encore les règlements scolaires. Il est évident

que les enseignants compétents, formés aux théories de l'éducation et aux méthodes d'enseignement, jouent un rôle important dans la qualité des procédés et des stratégies.

Le troisième aspect de la qualité concerne le résultat de l'éducation, son efficacité, sa productivité ou rendement. En fin de compte, le résultat de l'éducation doit être que le plus grand nombre possible d'élèves du groupe d'âge ou de la discipline concernés atteigne le comportement énoncé dans les objectifs. Ce résultat doit s'exprimer dans le cadre de tests (préliminaires), d'examens, d'observations ou d'activités pratiques.

Ces trois aspects de la qualité ne sont pas spécifiques à l'éducation. Dans le commerce et l'industrie, on accorde beaucoup d'attention à la « gestion par objectifs », c'est-à-dire le premier aspect. On prête attention aux procédés au travers d'approches comme le contrôle de qualité intégral, les normes ISO, le système EFQM de gestion de la qualité, largement utilisé en Europe. Tout cela sans oublier le classique contrôle des produits, qui repose sur des tests comparatifs souvent vivement encouragés par les organisations de consommateurs.

Les discussions sur la qualité présumée de l'éducation ne sont pas toujours rationnelles. Il arrive bien souvent que la notion de « qualité de l'éducation » soit utilisée avec un double sens. Par ailleurs, les discours sur la baisse du niveau, l'égalisation par le bas ou encore la baisse de la qualité ne sont souvent qu'une façade qui cache des motivations très diverses. Ces discours traduisent parfois la peur du changement, mais ils peuvent aussi être utilisés pour tenter de convaincre les personnes mal informées qu'il vaut mieux ne toucher à rien.

Tout acte en rapport avec le contenu ou le concept de chacun des trois aspects de la qualité décrits ci-dessus souligne le fait qu'il est impossible de décrire ce qu'est la qualité de l'éducation (le niveau, la norme) de

manière simple et univoque. L'acte en question est lié à des valeurs qui, à leur tour, tirent leurs racines des traditions, des exigences sociales, de la vision du développement de la société, des perceptions idéologiques et de plus en plus aussi de considérations en matière de psychologie éducative. Lorsque deux personnes discutent entre elles du niveau de l'éducation, il n'est pas toujours évident qu'elles soient sur la même longueur d'ondes. Au contraire, il y a bien souvent des questions de valeurs personnelles, qui peuvent ou non s'exprimer explicitement. Il en est de même pour les groupes qui, en société, expriment ce qu'ils entendent par qualité. La véhémence avec laquelle ces discussions sont bien souvent menées témoigne qu'il s'agit d'un thème intellectuellement et émotionnellement très chargé.

Les discussions sur le niveau de l'éducation s'expriment souvent sous forme d'oppositions de différentes sortes.

Il s'agit fréquemment d'oppositions sur la nature même de ce que l'on entend par niveau, par exemple :

- la quantité opposée à la qualité ;
- la spécialisation opposée à l'éducation de base ;
- l'éducation cognitive opposée à l'éducation générale ;
- l'éducation thématique opposée à l'éducation pluridisciplinaire ;
- la théorie opposée à la pratique ;
- la progression séquentielle opposée à la progression modèle ;
- les résultats immédiats opposés aux résultats permanents.

Quiconque a déjà participé au débat sur le renouveau de l'éducation sait très bien que ces idées refont surface régulièrement et que ceux qui en discutent adoptent toujours une position intermédiaire entre les deux extrêmes. Quoi qu'il en soit, il est évident que la qualité est un concept multiple, dans lequel toutes sortes d'aspects sociaux, idéologiques, psychologiques et pragmatiques se mêlent pour constituer un « axe » dynamique permettant de décrire le concept de la qualité.

## DES APPROCHES « CONTRASTÉES » DE LA QUALITÉ

Au risque de trop simplifier les choses, deux courants peuvent être distingués dans les arguments multiples et souvent contradictoires en faveur de la qualité dans l'éducation : une approche technico-économique de la qualité ; la démarche didactico-pédagogique.

### L'approche technico-économique

Comme son nom le suggère, l'approche économique découle de la réflexion sur la qualité engagée dans les cercles commerciaux des économies de marché. Quiconque entreprend quelque chose, et dans ce cas il s'agit de l'école, doit rendre compte des résultats obtenus. Les résultats déterminent la récompense ; vous obtenez ce que vous méritez. La personne qui fait plus obtient plus ; la personne qui fait moins obtient moins. La concurrence et la rivalité sont les moteurs de la récompense. Il s'agit d'une bataille pour conquérir l'avantage sur le marché.

Le marché est le principe moteur qui oriente et guide les réalisations. C'est le consommateur qui décide, en fin de compte, qui sera récompensé et dans quelle proportion.

Appliqué à l'éducation, ce principe se traduit par un contrôle rigoureux de l'enseignement par des programmes scolaires nationaux strictement encadrés, qui doivent être considérés comme la norme pour obtenir une « récompense en fonction des résultats ». Le rendement évalué au travers de tests centralisés détermine la qualité. L'extension des tests centralisés conduit à comparer les différentes écoles et établissements scolaires entre eux et à les classer en fonction de ce rendement. Le suivi du rendement exigé aboutit à un contrôle externe accru sur l'école au travers

d'exigences supplémentaires en marge du rendement (méthodes, organisation, regroupement des élèves, systèmes d'inspection rigoureux, etc.). L'une des conséquences immédiates d'une telle vision est également la dégradation de la position des professionnels, en particulier les enseignants et les chefs d'établissement. Ceux-ci ont en effet une marge de manœuvre de plus en plus réduite pour modeler eux-mêmes l'enseignement. Ils en sont de fait réduits, selon les termes de Michael APPLE, à n'être que des travailleurs semi-éduqués qui doivent se charger des tâches prosaïques du maintien de la discipline et de la préparation des tests.

### L'approche didactico-pédagogique

À l'exact opposé de la démarche purement économique, on trouve l'approche didactico-éducative. Là encore, l'expression parle d'elle-même. L'éducation prend sa source dans le droit des jeunes gens à une formation qui soit aussi bonne et aussi large que possible. La concurrence est ici moins importante puisqu'il s'agit de tous les futurs citoyens d'un pays qui doivent pouvoir bénéficier d'une éducation aussi riche que possible. La concurrence est remplacée par une dose considérable de solidarité. Les résultats sont bien sûr très différents, compte tenu de la grande diversité des élèves dans une école donnée. La réponse aux besoins particuliers impose la mise en œuvre d'une éducation individualisée, qui rend inutile la comparaison entre écoles<sup>1</sup>. L'adaptation aux individualités, qu'il s'agisse des élèves ou des écoles, exige des enseignants et des chefs d'établissement de faire preuve de créativité, en tenant compte des élèves présents, du contexte et des exigences locales. La qualité ne sera donc pas déterminée au niveau central, mais au niveau de l'école. Chaque école élaborera sa propre formule de qualité, le plus grand

apport à cet égard étant attendu d'une interaction créative entre le comportement des élèves d'un côté et celui des enseignants de l'autre. Cette faculté d'adaptation créative des enseignants aux situations impliquant les élèves et leurs comportements est caractéristique d'une relation éducative et didactique. Cette relation est en grande partie impossible à anticiper.

Une telle approche est plus facile à interpréter au niveau de l'école, puisqu'elle est élaborée à partir des caractéristiques réelles de la formation et de l'éducation : il s'agit d'une relation éducative dynamique, ascendante et en partie imprévisible. Dans cette approche, l'école est au centre du contrôle de qualité. C'est en grande partie aux professionnels de déterminer le climat éducatif et l'approche en matière éducative. L'évaluation externe est subordonnée à l'évaluation interne. L'école va renforcer son contrôle et son évaluation internes puisque chaque école aura sa propre interprétation de la qualité. La capacité de différenciation de l'école mettant en œuvre une éducation individualisée de manière optimale, par exemple la pédagogie de maîtrise, se voit accorder une place centrale.

Dans des circonstances où les gens pourraient penser que le second modèle est celui qui donne le ton dans l'éducation, il s'avère en réalité qu'à l'heure actuelle, c'est de moins en moins le cas. Le premier modèle a clairement gagné du terrain ces dernières années. Cela s'explique par une vision économique plus étroite de l'éducation ou par un malaise général vis-à-vis de la perte d'un certain nombre de repères sociaux, dans lequel l'éducation sert souvent de bouc émissaire (chômage, délinquance, performances dans le cadre de la concurrence internationale, etc.).

Cette évolution a donné naissance à un mouvement souvent présenté sous le nom de « retour à l'essentiel ». Ce mouvement est né dans

les années 80, principalement aux États-Unis. Une quantité de rapports, basés sur le célèbre rapport « *A nation at risk* », en est l'illustration. Un mouvement comparable remporte actuellement un succès sans précédent en Angleterre. Des programmes scolaires nationaux sont élaborés et des tests centralisés sont prévus pour tous les élèves âgés de sept, onze, quatorze et seize ans. Par ailleurs, les écoles doivent publier les résultats de ces tests et sont ensuite classées en fonction des résultats des tests. Ces mouvements sont actuellement sur la pente ascendante. Ils sont désormais également présents dans d'autres pays, et pas uniquement de manière latente.

Du fait de son manque d'esprit critique vis-à-vis des propositions concernant la qualité de l'éducation, la presse généraliste joue un rôle considérable dans la création d'un certain climat obsessionnel en ce qui concerne le niveau de l'éducation. Aux États-Unis, l'éminent pédagogue GOODLAD, parmi d'autres, a fortement critiqué ce que le *New York Times* du 30 janvier 1984 appelait le « Reaganisme éducatif », à savoir une démarche caractérisée par des pratiques de test rigoureuses, l'appel à la discipline et le paiement des enseignants en fonction des résultats des tests. GOODLAD qualifie cela de « préoccupation pathologique pour les effets sur les élèves » contre laquelle il revendique l'alternative de la qualité de la vie scolaire dans toute sa diversité.

#### NOTE

1. Dans cet article le mot « école », lorsqu'il est utilisé au sens de lieu d'enseignement, est à comprendre comme désignant à la fois les écoles primaires et les établissements d'enseignement secondaire.

## LE CAS DE L'ANGLETERRE

La situation en Angleterre illustre l'approche technico-économique actuelle de la qualité. Comme c'est souvent le cas, une vision de l'éducation et de l'évaluation de la qualité de l'éducation ne naît pas d'un coup. En Angleterre aussi, l'approche actuelle a été précédée d'une longue histoire, marquée par un certain nombre d'événements.

### Examens externes et rémunération au mérite

Le décret de 1862 concernant la « rémunération en fonction des résultats » est bien connu. Le vice-ministre de l'Éducation de l'époque, Robert LOWE (disciple du célèbre économiste libéral Adam SMITH), avait en effet promulgué un décret liant la rémunération des enseignants aux résultats des élèves en matière de lecture, d'écriture et de calcul. Avec ce système, les inspecteurs devaient noter les performances des élèves en six catégories ; les enseignants bénéficiaient alors d'une même rémunération supplémentaire par rapport au salaire minimum, en fonction de la performance des élèves.

Ce système a soulevé de nombreuses protestations, y compris de la part des inspecteurs. Il s'est en effet traduit par la pratique excessive d'exercices directement liés aux tests et par l'apparition d'un « groupe alimentaire » intermédiaire : les élèves ne procurant pas ce revenu « alimentaire » étaient négligés. On pouvait déjà repérer, dans ce système, les graines de l'éducation libérale, toujours prédominante en Angleterre. Le système honni de « rémunération en fonction des résultats » ne fut aboli qu'en 1897, après le succès du parti conservateur sur les libéraux.

Cette rémunération différenciée ou « rémunération au mérite » avant

la lettre s'est accompagnée de l'organisation d'examens externes.

L'origine des examens externes en Angleterre et au Pays de Galles peut être située juste après le milieu du dix-neuvième siècle, lorsque les universités d'Oxford, de Cambridge et de Londres introduisirent un examen d'entrée. La réputation de ces examens ne s'explique pas seulement par le prestige de ces universités, mais également par la politique mise en œuvre en 1868 par les Libéraux après leur succès aux élections. Le gouvernement de GLADSTONE introduisit ainsi une série de réformes qui, historiquement, lui valurent le nom de « ministère des grandes réformes ». Ces réformes s'inscrivaient dans le mouvement d'abolition des anciens systèmes de recrutement et de promotion des fonctionnaires, très favorables à l'aristocratie. Le recrutement et la sélection des fonctionnaires s'opérèrent désormais sur la base d'examens. De la même façon, il devint impossible pour l'aristocratie d'acheter une carrière dans l'armée. À l'avenir, seules les compétences et le mérite permettraient la promotion. Les citoyens des classes moyennes applaudirent à ces réformes. La concurrence se développa alors pour l'obtention de postes dans l'administration et les jurys des examens se retrouvèrent surchargés de travail.

La bonne disposition des classes moyennes vis-à-vis des examens s'étendit également à l'enseignement secondaire, de plus en plus fréquenté par les enfants de ces classes. Il ne s'écoula pas longtemps avant qu'un certain nombre d'écoles secondaires prestigieuses demandent aux universités d'Oxford et de Cambridge d'élaborer des examens pour elles. Le premier comité d'examens externes pour l'enseignement secondaire fut ainsi mis en place : le comité mixte d'Oxford et Cambridge. D'autres universités suivirent, avec toutes sortes de comités d'examens privés. Au début du vingtième siècle, on comptait déjà huit comités d'examens gérés par des universités. L'opinion

publique avait la plus haute estime pour ces comités. Le système suscita toutefois une certaine confusion, chaque université tentant d'imposer ses propres normes à l'enseignement secondaire.

Il fallu alors mettre un peu d'ordre dans ce système disparate. Mais les examens avaient entre temps remporté un tel succès qu'il ne pouvait absolument plus être question de rechercher des alternatives du côté d'examens internes. Il fut alors demandé au comité consultatif de la direction de l'éducation de proposer une solution pour garantir une plus grande uniformité des examens. En 1911, le comité publia le rapport intitulé : « les examens dans les écoles secondaires », qui concluait qu'une rationalisation était souhaitable et qu'il fallait pour cela uniquement charger les universités de faire passer les examens. Après bien des discussions, cela leur parut être la meilleure solution.

Un « certificat d'école secondaire » serait attribué après au moins trois ans d'enseignement secondaire, sur la base d'un âge d'entrée dans le secondaire de douze ans. L'âge moyen auquel ce certificat serait attribué serait donc de 16 ans.

Le comité proposa un examen portant sur une seule matière à la fois, à choisir dans des groupes de matières. Depuis cette époque, cette méthode de travail est appelée « examens par groupes », par opposition aux examens réformés qui, depuis 1851, étaient de simples « examens sur des matières optionnelles ». Le comité proposa également un second examen à l'âge de 18 ans, cet examen ne devant toutefois pas avoir de valeur de qualification, mais une valeur de sélection en vue des études universitaires.

Ces propositions furent l'objet de discussions entre la direction de l'éducation et les comités d'examens universitaires en 1914, 1915 et 1916. Les universités finirent par accepter d'adapter leur système en fonction de ces propositions.

En 1917, la direction de l'éducation introduisit les premiers examens de la « première école » (à l'âge de 16 ans). Les élèves devaient passer des examens dans cinq matières, avec toujours au moins une matière choisie dans chacun des trois groupes suivants : « humanités », langues étrangères et mathématiques. Il existait également un quatrième groupe, englobant l'art, la musique et l'économie domestique. Mais aucun sujet d'examen ne pouvait être choisi dans ce groupe. Selon le comité, ce n'est pas que ces sujets étaient de nature inférieure, mais ils ne pouvaient pas faire l'objet d'une mesure raisonnablement exacte. La vérité est probablement que les comités d'examens universitaires n'étaient que peu intéressés par ces disciplines. Il fallut attendre 1930 pour qu'il soit possible de choisir deux matières dans ce quatrième groupe pour le premier certificat.

En plus du premier certificat pour les jeunes de 16 ans, un certificat « supérieur » plus spécialisé fut introduit au même moment pour les jeunes de 18 ans qui souhaitaient entrer à l'université.

Ces examens standardisés eurent un grand impact sur l'évolution de l'enseignement en Angleterre et au Pays de Galles. Le prestige des universités était considérable, de sorte que le modèle fondamental de l'université traditionnelle se reflétait dans l'enseignement secondaire. Le prestige des universités avait également une influence auprès des employeurs. Eux aussi étaient de plus en plus demandeurs de ce certificat, avec là encore pour conséquence un accroissement de la pression des parents sur leurs enfants pour que ceux-ci obtiennent le certificat. À partir de la ratification officielle de ce système d'examens externes en 1917, le système anglais se trouva enserré dans un réseau où chaque fil allait dans la direction des examens externes. Depuis cette époque, il devint quasiment impensable, en Angleterre et au Pays de Galles,

d'envisager des alternatives à la certification des élèves, autres que les examens externes centralisés. Les structures créèrent une pensée qui s'exprimait en fonction de ces structures elles-mêmes. Les commentaires critiques sur le système n'allaient jamais remettre en question le système des examens externes, ils se limiteraient dès lors aux modalités de mise en œuvre de ces examens.

### **Le corps des inspecteurs**

En 1856 également, quelque 30 inspecteurs, membres du corps des inspecteurs de Sa Majesté (Her Majesty's Inspectorate [HMI]), furent nommés dans le service de l'éducation alors mis en place. Lorsque le service de l'éducation fut promu au rang de « Conseil de l'Éducation » ou « Direction de l'Éducation », ce conseil comptait 300 fonctionnaires et 350 inspecteurs. Le travail du HMI prit progressivement forme.

Aux environs de 1900, les inspecteurs faisaient des visites d'une journée dans les écoles primaires, seuls ou par deux. Chaque école primaire était ainsi inspectée tous les trois à cinq ans. En ce qui concerne l'enseignement secondaire, l'option « inspection complète » fut choisie dès le début : l'ensemble de l'école était inspectée par une équipe d'inspection. Toutes les écoles secondaires furent ainsi inspectées en 1912.

En raison du manque de temps, une variante abrégée fut par la suite imaginée. Cette variante se contentait du contrôle aléatoire de certains indicateurs. Malgré cela, avec l'ampleur du développement de l'enseignement secondaire, on en arriva à une situation où, en 1922, une école ne pouvait être inspectée que tous les dix ans. Le nombre des « inspections complètes » diminua alors en faveur de variantes limitées, plus thématiques.

Après l'adoption de la Loi BUTLER en 1944, le HMI se fit de plus en plus

discret, et ce, pendant une vingtaine d'années. Aux environs de 1950, les « inspections complètes » n'existaient même plus. Jusqu'aux années 70, le rôle du HMI devint de plus en plus consultatif. La raison d'être même du HMI fut mise en cause, en partie sous la pression des inspecteurs rivaux des directions régionales de l'éducation, les autorités éducatives locales (Local Education Authorities [LEA]). À partir de 1969, le paysage éducatif changea radicalement. Il était devenu nécessaire de faire des économies et les résultats de l'enseignement ne faisaient que susciter une défiance croissante. Dans les « Livres blancs » sur le système éducatif de 1969 et 1970, d'éminentes personnalités exprimèrent leur mécontentement vis-à-vis de la « dégradation » de l'enseignement. La tendance internationale était alors à « rendre des comptes » (accountability) : les écoles devaient rendre compte de leurs résultats en soumettant des informations claires sur la manière dont elles utilisaient les subventions qui leur étaient attribuées. Ce mouvement relança le HMI. Celui-ci élargit une nouvelle fois sa démarche. En plus des inspections traditionnelles, il organisa également des visites informelles et des inspections thématiques et publia des rapports d'études intéressants sur certains thèmes.

Tout cela aboutit, en 1974, à la mise en place d'une unité d'évaluation des performances (Assessment Performance Unit [APU]), sous-direction du HMI, chargée d'organiser des enquêtes ou des études systématiques sur certains thèmes, par exemple sur les groupes hétérogènes, l'enseignement d'une seconde langue, la lecture, le calcul, etc.

En 1983, la publication des rapports du HMI devint obligatoire. Le « Livre blanc » de 1985, intitulé « Better Schools » (de meilleures écoles), attirait de nouveau l'attention sur les avantages d'une approche centralisée des résultats de l'enseignement, contrairement à l'approche locale des LEA, qui

rendait possible l'existence de disparités considérables.

Le fameux « Education Reform Act » (loi de réforme de l'éducation) de 1988 (Loi BAKER), le joyau de la politique éducative des conservateurs, introduisit un « programme scolaire national » pour les élèves âgés de 5 à 16 ans et jeta les bases des « ordonnances » dans lesquelles les « programmes d'étude » allaient devoir être consignés. À l'avenir, des tests nationaux allaient devoir être passés à la fin de quatre « stades clés ». Un « Conseil du programme scolaire national » fut mis en place afin d'apporter des conseils sur les « objectifs de résultats » et de fournir du matériel à l'appui du nouveau programme scolaire.

### **L'orientation consumériste et les « tableaux comparatifs »**

Lorsque John MAJOR succéda à Margaret THATCHER au poste de premier ministre, en 1991, il se montra très intéressé par les questions d'éducation et promulgua la « Charte du citoyen », qui prévoyait que les résultats de chaque école seraient publiés dans des « tableaux comparatifs ». L'orientation consumériste était un concept central : cela signifiait qu'il devait y avoir une transparence totale quant aux performances de l'école, de manière à ce que les parents puissent choisir les écoles sur la base d'informations claires. L'inspection devait veiller à ce que l'école ait un bon « rapport qualité/prix ». Le HMI, qui n'avait toujours de comptes à rendre qu'à la Reine, était là encore très critiqué : les conclusions des inspections conduites par le HMI n'étaient pas assez concrètes et pas suffisamment documentées sur le plan statistique. Par ailleurs, il était reproché au HMI de ne pas être suffisamment tourné vers les parents en tant que consommateurs sur le marché de l'éducation. Comme les inspecteurs étaient recrutés parmi les enseignants

et les chefs d'établissement, ils risquaient de trop pencher en faveur de ces groupes dans leurs rapports. De plus, les gens étaient de plus en plus conscients que les inspecteurs étaient trop peu nombreux, ce qui avait pour résultat que les écoles n'avaient guère à craindre d'être visitées par eux.

L'indépendance du HMI provoqua également une controverse par rapport à la politique officielle. Le rapport très critique d'Éric BOLTON, dernier doyen de l'inspection générale de la vieille garde, concernant les effets négatifs de la politique du gouvernement sur la formation des enseignants, fut la goutte d'eau qui fit déborder le vase.

En 1991, le ministre CLARK proposa une organisation entièrement nouvelle de l'inspection dans la « Charte des parents ». L'inspection se retrouvait largement privatisée. Un certain nombre de membres du HMI restaient en place afin de former les inspecteurs privés et de donner des instructions. Le corps des inspecteurs allait à l'avenir être dirigé par l'Inspecteur en chef de Sa Majesté. Chaque école serait inspectée tous les quatre ans et le rapport d'inspection serait publié. Une synthèse serait publiée à l'intention des parents, avec mention de la position de l'école inspectée dans les « tableaux comparatifs ». Les directions d'écoles pourraient choisir elles-mêmes le contrat d'inspection qu'elles souhaitaient souscrire, sur la base du principe de concurrence entre les agences d'inspection. À l'avenir, chaque équipe d'inspection devrait comporter un « inspecteur profane » : un non-initié qui représenterait les intérêts de la communauté (« Education Schools Act » [loi sur les établissements d'enseignement], 1992).

En 1992, les Conservateurs remportèrent une nouvelle fois les élections, ce qui leur permit de pouvoir continuer à appliquer la politique qu'ils avaient choisie. Le « Livre blanc » intitulé « Choix et diversité » préparait le terrain pour l'adoption d'une nouvelle loi.

La « loi sur les établissements d'enseignement » de 1992 confirma finalement le point de vue du ministre Clark, mais en y ajoutant quelques éléments concrets. Ainsi le corps des inspecteurs fut-il transformé en un « Office for Standards in Education » (agence pour la qualité de l'enseignement [OFSTED]). Tous les inspecteurs devaient être agréés, suivre une formation spécifique, constituer une équipe et passer un contrat avec l'OFSTED. Dans la pratique, les anciens inspecteurs des LEA, les universités et les agences privées demandèrent l'agrément pour eux-mêmes ou pour leurs membres. En 1993, le système d'inspection fut précisé dans un « Guide de l'inspection des établissements d'enseignement » (Framework for the Inspection of Schools), qui fut envoyé à tous les établissements d'enseignement de manière à ce que ceux-ci sachent quoi attendre d'une inspection. Ce guide comprenait également tous les indicateurs que le corps des inspecteurs proposait d'adopter à titre de normes.

### **Les tendances actuelles**

Tony BLAIR a écrit, aussi bien en tant que leader de l'opposition que dans sa fonction actuelle de premier ministre, que la qualité de l'enseignement devait être au centre du programme du Parti Travailleiste. Son « Livre blanc » intitulé « L'excellence dans les établissements scolaires » (Excellence in Schools), repose sur six principes.

- L'éducation doit être au centre des préoccupations du gouvernement.
- Le système éducatif doit fonctionner au bénéfice du plus grand nombre, c'est-à-dire qu'il doit également prendre en compte les potentialités de tous les enfants, en évitant les erreurs du système précédant de sélection précoce (« eleven plus »).
- Les résultats doivent être plus importants que les structures. Sur cette base, il convient de mettre en place

une formation des enseignants, les enseignants incompetents doivent être rapidement congédiés, les bonnes pratiques doivent être rendues publiques et l'accent doit être mis sur la qualité des chefs d'établissement.

- Le contrôle externe par l'inspection doit avoir lieu au moins une fois tous les six ans, mais l'inspection doit être principalement centrée sur les établissements qui obtiennent des performances décevantes.

- La bienveillance vis-à-vis des établissements en situation d'échec sera très limitée : le document parle de « tolérance zéro à l'échec ». Chaque établissement doit avoir des objectifs clairs, établis sur la base des informations issues des objectifs nationaux, du matériel d'inspection récent et des évolutions dans des établissements comparables. Chaque LEA doit exposer sa contribution au progrès des résultats en matière d'enseignement dans un « Plan de développement de l'enseignement » (Education Development Plan). Enfin les LEA seront elles aussi soumises à des contrôles de l'OFSTED.

- Et enfin, le gouvernement souhaite constituer un partenariat avec tous ceux qui peuvent contribuer à améliorer les résultats de l'enseignement : 25 « Zones d'action éducative » (Education Action Zones) sont mises en place à cette fin. À l'avenir, les nouvelles subventions qui seront attribuées pour les bâtiments et les équipements scolaires le seront en fonction des performances éducatives ; c'est dans ce contexte que l'on utilise le terme « rémunération au mérite ».

La décision du Parti Travailleiste de nommer un secrétaire d'État aux « normes » ne peut pas prêter à confusion. Cette personne sera principalement responsable de la politique vis-à-vis des établissements en situation d'échec (la politique de « shame and blame ») et de l'amélioration des résultats aux tests centralisés par divers moyens.

Il s'agira notamment de l'amélioration des résultats en matière de lecture et de calcul, dans l'objectif d'amener 80 % des élèves de 11 ans au niveau 4 dans leur langue maternelle et 75 % au niveau 4 en calcul d'ici 2002. Cet objectif est basé sur une échelle de 10 pour tous les groupes d'âge scolaire. L'élément le plus nouveau de cette mesure est l'investissement dans des « enseignants experts ». S'ils sont sélectionnés comme experts, ces enseignants pourront gagner un salaire confortable.

Le cœur de cette politique de qualité est constitué par les résultats aux tests centralisés. Ils sont régulièrement publiés dans les « tableaux comparatifs », dans lesquels les établissements sont classés en fonction de leurs résultats à ces tests.

La communauté scientifique, en particulier, critique fortement la présentation simpliste de ces résultats scolaires sous forme de moyennes qui ne tiennent pas compte du recrutement des élèves et de l'environnement de l'établissement scolaire concerné.

Par ailleurs, il a été plus d'une fois constaté que des « départements » donnés dans un établissement réussissent mieux que d'autres, mais ils sont néanmoins exprimés dans le cadre d'une moyenne pour l'établissement. La critique la plus pertinente à cet égard se fonde généralement sur l'idée que ce type de système de classement ne permet pas de préciser la « valeur ajoutée » de l'enseignement.

Quoi qu'il en soit, dans leurs réflexions, les gens continuent à suivre la logique du classement central. Des établissements scientifiques se voient confier la tâche d'élaborer des tableaux comparatifs dont le contenu aurait plus de « valeur ajoutée ». La mesure des gains de l'enseignement doit devenir la nouvelle mission.

Des formes de tests de référence sont par exemple à l'étude. Dans les tests de référence, les performances de l'élève sont mesurées au début

d'une période d'enseignement et, après un certain temps, par exemple deux ou quatre ans, les performances sont évaluées et comparées aux performances de référence. Les travaux de l'Université de Durham (avec notamment Carole FITZ-GIBBON) font à cet égard autorité. Bien que ces chercheurs considèrent expressément leur système comme un instrument d'auto-évaluation des établissements scolaires et non comme un outil de classement externe, leur méthode de travail a déjà inspiré le gouvernement. Selon une mesure récente, les résultats à l'âge de 16 ans dans les tableaux comparatifs devaient être comparés aux résultats obtenus à l'âge de 14 ans. Ce système a pour l'heure été abandonné en raison de la forte protestation des enseignants.

En 1998, l'inspection anglaise a également introduit les « Performance and Assessment Reports » (rapports de performances et d'évaluation [Panda]). Vingt-quatre mille écoles et établissements scolaires anglais ont été répartis en catégories selon le milieu social des élèves. Les établissements ont ensuite été classés au sein de chaque catégorie.

Néanmoins, en Angleterre, la réflexion oscille toujours entre les examens centralisés et la comparaison des écoles. Les gens attendent beaucoup des statisticiens qui vont probablement trouver les bonnes formules pour corriger les résultats approximatifs des tableaux comparatifs sur la base de divers facteurs, comme le milieu social, l'environnement et l'intelligence.

La vérité n'existe qu'au sein d'une certaine grammaire, comme l'enseigna en son temps WITTGENSTEIN. La grammaire n'est toutefois elle-même jamais remise en question.

La situation anglaise décrite ici de manière assez détaillée traduit ce que l'on doit entendre par une approche technico-économique du contrôle de qualité. Mais de nombreux principes de base d'une telle approche de la qualité méritent discussion.

## LES RÉSERVES QUANT À UNE APPROCHE TECHNIQUE UNILATÉRALE

Les réserves concernant une approche exclusivement technico-économique peuvent être regroupées autour de quatre aspects, à savoir : les limites de la mesure ; une école n'est pas une entreprise ; la comparaison des écoles ; le redressement par l'environnement social.

### Les limites de la mesure

#### *Mettre les choses en perspective en s'appuyant sur le bon sens*

Il est peu de façons autant exploitées, pour convaincre les gens, que les chiffres et les statistiques. Pourtant, même si cela relève presque d'une sagesse populaire fondée sur le bon sens, la probabilité que les mesures et leurs résultats se transforment en un jugement équilibré ne semble toujours pas acquise.

Certains exemples de mise en perspective des choses enflamment l'imagination. Le mot d'esprit du célèbre écrivain néerlandais Godfried BOMANS décrivant la noyade d'une personne traversant une rivière d'un mètre de profondeur en moyenne est bien connu. Dans l'une de ses « pensées », l'illustre Pascal, qui était aussi un grand mathématicien, avait également clairement souligné l'importance relative des choses : *« Et ainsi il est rare que les géomètres soient fins et que les fins soient géomètres, à cause que les géomètres veulent traiter géométriquement ces choses fines, et se rendent ridicules, voulant commencer par les définitions et ensuite par les principes, ce qui n'est pas la manière d'agir en cette sorte de raisonnement. Ce n'est pas que l'esprit ne le fasse ; mais il le fait tacitement, naturellement et sans art, car l'expression de ces choses dépasse tous les hommes, et le*

*sentiment n'en appartient qu'à peu d'hommes ».*

Soyons clair d'emblée : traduire la réalité en chiffres est utile, mais cette démarche a également ses limites. Des différences importantes peuvent être observées dans les possibilités de mesurer. Il est par exemple possible d'utiliser de nombreuses mesures physiques et d'aller très loin, alors que le comportement humain devient de moins en moins exactement mesurable lorsque le niveau de complexité augmente. Lorsqu'il est question de qualité dans l'enseignement, la modification du comportement des élèves est un élément central. Il n'est pas surprenant qu'une approche éducative systématique tente d'avoir prise sur l'évolution du comportement et des performances humaines.

#### *La mesure du comportement humain*

Lorsque l'on mesure le comportement humain, il ne peut pas y avoir d'échelle de comparaison sur laquelle on mesurerait des quantités physiques (longueur, contenu, durée, vitesse, énergie, etc.). Une échelle de comparaison a un zéro reconnaissable et fixe. Par ailleurs, les distances entre les unités de l'échelle sont supposées égales.

Dans les sciences comportementales, les gens utilisent souvent des échelles d'intervalles. Les chiffres sont importants dans la mesure où les distances entre les points de mesure sont supposées égales. Il est donc également possible de calculer des moyennes. On ne peut pas dire que quelqu'un qui a 8 sur 10 est deux fois meilleur que quelqu'un qui a 4 sur 10. Après tout, il n'y a pas de zéro fixe. Dans de nombreux tests, le zéro est, de toutes façons, choisi arbitrairement. Si quelqu'un fait 30 fautes à la dictée de Bernard PIVOT, il est malgré tout considéré comme fort en orthographe. Une personne qui ne fait aucune faute obtient le niveau maximum absolu. Cet exemple

montre une nouvelle fois qu'un test permet de déterminer un niveau minimum et un niveau maximum.

La manière dont ces niveaux sont déterminés est toujours matière à discussion et dépend d'une norme fixée « quelque part ». Plus la norme est légitimée par un consensus, mieux elle est acceptée. Un tel consensus découle dans ce cas d'un certain nombre d'exigences déterminées de manière assez précise, par exemple, par un parlement, ou par une norme statistique, la valeur moyenne étant généralement prise comme point de départ. La subjectivité augmente lorsque le consensus est plus faible ou qu'il y a moins de matériel de comparaison.

La distance entre les points est encore un autre problème.

Si Marie obtient 8 sur 10, est-elle deux fois meilleure que Johnny, avec son 4 sur 10 ? On ne peut absolument pas dire cela. La seule chose évidente, c'est que les notes permettent un classement. Marie fait beaucoup mieux que Jean et légèrement mieux que Pierre, qui a obtenu 7 sur 10.

Et pourtant, là aussi, il y a de la subjectivité. Dans le cadre du test, on accorde un certain nombre de points par question. Dans un test avec des questions à choix multiples, par exemple, chaque question peut donner un point.

Mais les questions sont-elles toutes vraiment aussi importantes et pourquoi ? Et si une question donne plus ou moins de points, sur quoi cela se base-t-il ?

On peut par ailleurs dire qu'un score ne renseigne absolument pas sur le profil de connaissances de l'élève : à quel type de question Marie n'a-t-elle pas pu répondre et à quel type de question Pierre n'a-t-il pas pu répondre ? Là aussi, il manque des éléments de comparaison. La note ne permet qu'un classement et renseigne peu sur la qualité.

C'est encore plus difficile lorsque l'on note des questions ouvertes. Comment accorder des points sans que la marge de manœuvre soit trop

grande ? Les études sont nombreuses qui montrent que des enseignants différents ont noté les mêmes tests de façon très variée. L'alternative qui consiste à ne poser que des questions fermées se traduit quant à elle par un appauvrissement considérable de la matière à apprendre.

La subjectivité présente dans tous les tests ne va faire qu'augmenter si l'on ajoute les résultats de divers tests pour obtenir une note globale. Là encore, on se retrouve confronté à la tâche difficile d'attribuer, disons, un pourcentage pour la géographie et un pourcentage pour les mathématiques. Sur quoi baser la différence entre les deux pourcentages ? Qu'est-ce qui est important et qu'est-ce qui est moins important ? Et pour aller plus loin encore : qu'est-ce qui est important dans l'environnement A et qu'est-ce qui est important dans l'environnement B ?

Les choses sont encore plus difficiles si vous comparez des classes, des départements ou, *a fortiori*, des moyennes par école.

Quelle diversité d'environnements la moyenne unique cache-t-elle alors ? Quels types d'élèves a-t-on ? Quels enseignants ? Quels parents ? Quel type d'infrastructure ? Quel environnement ? Quelle culture scolaire ? Quels types d'idéologies sont inclus dans les résultats ? etc.

Il est évident que les psychométriciens professionnels connaissent ces difficultés. Bien des efforts ont également été faits pour réduire l'importance de la subjectivité. Un outil statistique sophistiqué et performant a été élaboré. Cela ne signifie toutefois pas que la notation directe du comportement soit toujours basée sur un élément de subjectivité. Les manipulations délicates réalisées à cette occasion n'ont pas réussi à éliminer « l'échec fondamental ». Les méthodes statistiques les plus performantes ne peuvent pas éradiquer totalement ce mal. Aucune acrobatie en matière de notes ne peut se passer totalement d'un filet de

sécurité pour tenir compte des facteurs subjectifs.

Le slogan « mesurer, c'est connaître » nécessite donc clairement une approche différenciée, où l'on ait toujours conscience de la subjectivité des décisions fondamentales et où l'on tienne compte du caractère réducteur, en termes de qualité, des notes, en particulier dans le cas du comportement humain.

### **L'école n'est pas une entreprise**

On souligne souvent que l'éducation a beaucoup à apprendre du contrôle de qualité tel qu'il est mis en œuvre dans le monde économique. C'est peut-être le cas, mais cela ne doit pas nous empêcher de regarder l'individualité de chacun des deux systèmes. Lors d'une réunion d'hommes d'affaires et de pédagogues, le journaliste flamand spécialiste de l'enseignement BODIFIEE a exprimé les choses de la manière suivante : « *Le monde de l'éducation et le monde des affaires doivent apprendre à se connaître, à s'apprécier et à se stimuler, mais ils ne doivent jamais devenir un seul et même monde* ». Les entreprises façonnent les technologies, mais pas les gens. Bien sûr, elles ne font pas que produire des machines frustes et aveugles, elles accordent également de l'attention aux questions humaines. Mais ce monde se heurte cependant à celui de l'éducation. Et c'est comme cela que ce doit être. Il n'y a pas besoin d'une masse informe qui proviendrait d'un vague amalgame entre l'éducation et les affaires. Nous devons accepter la tension, non pas comme une difficulté, mais comme une source de créativité. Chacun des deux mondes doit faire ce qu'il sait faire. Cette diversité joue un rôle moteur.

L'enseignement peut apprendre beaucoup des efforts des entreprises pour garantir la qualité. On parle souvent à cet égard de « qualité totale » ou de « qualité intégrale ».

L'entreprise dispose de plus de souplesse et de flexibilité et elle est capable de réagir plus rapidement en cas de modification de l'offre et de la demande.

Il est aussi par ailleurs également vrai que les efforts pour améliorer la qualité existent depuis longtemps dans l'enseignement, mais ces efforts ont été peut-être plus difficiles à mettre en pratique. Souvenons-nous par exemple de la pédagogie de maîtrise, des systèmes de suivi des élèves, des systèmes d'individualisation, etc.

Entre l'école et l'entreprise, il existe aussi des différences manifestes :

- l'enseignement n'est pas une initiative libre ; il est obligatoire pour les élèves jusqu'à l'âge de 16 ans ;
- l'enseignement est enserré dans des limites structurelles puissantes (liberté de l'enseignement, preuves des compétences, aménagement de l'année scolaire, etc.) ;
- l'objectif de l'école n'est pas de faire du profit, mais de former des élèves, quoi qu'on entende par là ; les objectifs de l'école sont plutôt vagues. Son efficacité et son rendement sont difficiles à évaluer ;
- il existe encore peu de paramètres clairs pour mesurer la qualité d'une école. Les paramètres utilisés sont souvent unilatéraux et limités aux résultats de l'école, sans tenir compte des efforts des élèves. Les différences de qualité sont souvent dues aux différences de recrutement ;
- la rémunération des employés ne dépend pas de la situation de certaines exigences de qualité.

Ayant dit cela, il semble évident que les processus de promotion de la qualité dérivés du monde des affaires doivent être interprétés pour pouvoir être adaptés à l'éducation. Il n'y a par ailleurs aucune raison de remplacer des formules de qualité qui ont fait leurs preuves dans l'éducation par d'autres systèmes moins spécifiquement adaptés à l'éducation.

## La comparaison des écoles

L'idée grandissante de comparer les écoles les unes aux autres et de les classer en fonction de leur qualité est parfaitement conforme à l'approche utilisée pour les entreprises. Les indications visant à informer le consommateur pour qu'il puisse faire la meilleure acquisition possible sont légion. Le consommateur doit être en mesure de choisir et cela n'est possible que s'il dispose d'indications claires en matière de qualité. La comparaison, lorsqu'il s'agit de marchandises, est là encore évidente. Les poudres à laver et les autocuiseurs peuvent être examinés relativement facilement quant à leurs propriétés physiques et chimiques et le résultat des comparaisons peut s'exprimer sous forme de chiffres. Ce n'est pas aussi évident pour les performances des élèves, et encore moins pour chaque école prise dans son ensemble.

Les notes que les élèves obtiennent à l'école lors des tests centralisés externes sont importantes. Le classement des écoles en fonction de leurs résultats à ces tests est très tentant, puisqu'il permet de réduire une réalité complexe à de simples notes. Les tableaux comparatifs sont très populaires en Angleterre ; en France, le classement de tous les établissements d'enseignement qui font passer le baccalauréat est publié. Chaque année, le périodique qui publie ce classement (*Le Monde de l'Éducation*) augmente sensiblement à cette occasion son chiffre d'affaires.

Les Pays-Bas aussi se sont vus contraints à publier les résultats des établissements scolaires et à les classer, après une demande du journal *Trouw*. Le *Volkskrant* est même allé plus loin. Contrairement aux souhaits du corps des inspecteurs, il a traduit les rapports des inspecteurs en notes. Cinq indicateurs dont les inspecteurs ont rendu compte dans quatre catégories différentes ont reçu des notes de 1 à 4. Ces notes ont

ensuite été additionnées et divisées par cinq. Le résultat obtenu représentait dès lors le score de l'école. La nécessité du principe de transparence fut invoquée.

Mais il s'agit là de simplifications très grandes de la réalité. Des notes aussi approximatives ne renseignent absolument pas sur la qualité d'un établissement si l'on ne tient pas compte du recrutement des élèves, plus déterminé par leur situation socio-économique et par leurs capacités. Il semble par exemple très clair que les établissements qui ont un fort pourcentage d'élèves issus des groupes défavorisés n'ont aucune chance de rivaliser avec des établissements qui pratiquent un recrutement sélectif. C'est une erreur logique considérable que de décider sur cette base qu'un établissement est meilleur qu'un autre. Il se pourrait bien, après tout, que les élèves de l'établissement accueillant des enfants issus des groupes cible apprennent beaucoup plus que les élèves d'un établissement où ne sont scolarisés que des enfants avec de fortes capacités. Cette comparaison simpliste est à juste titre rejetée dans les cercles scientifiques. La majorité de la population sera toutefois tentée par la simplicité du principe des notes. Le résultat étant, souvent à tort, que les écoles qui recrutent dans des milieux défavorisés sont considérées comme mauvaises.

Une étude intéressante de la London school of economics (université d'économie) de Londres apporte des éclaircissements nouveaux sur les différences entre écoles [ROBINSON, 1997]. Les facteurs réels qui déterminent essentiellement les différences dans les performances des écoles peuvent en grande partie être attribués à trois variables.

- Le niveau de pauvreté dans lequel un élève grandit. Les bulletins scolaires sont en rapport avec le niveau de pauvreté, qui à son tour génère toutes sortes de comportements non scolaires. La réduction de la pauvreté dans la société permet également

de réduire les bulletins scolaires défavorables.

- L'élargissement du recrutement influence de manière favorable les performances de tous les élèves de l'école, même celles des élèves issus des milieux plus pauvres. C'est un argument en faveur d'un recrutement équilibré des élèves. Les écoles « ghettos » sont à l'origine d'une spirale vers le bas. Socialement parlant, cela signifierait qu'il faille imposer certaines limites au libre choix. La plupart des gens se prononcent pour la justice sociale et la solidarité en tant que telles, mais pas lorsqu'il s'agit de leurs propres enfants. On peut également s'en rendre compte dans le succès limité que remportent les accords de non discrimination dans le recrutement des élèves en Flandre.

- L'intérêt et l'engagement des parents pour l'éducation de leurs enfants.

L'environnement social et un recrutement non équilibré semblent jouer un rôle bien plus important que des facteurs éducatifs typiques comme les méthodes, le regroupement des élèves, un bon matériel didactique, la quantité plus ou moins importante de devoirs, etc.

Les facteurs caractéristiques de la qualité de l'enseignement, comme, par exemple, les méthodes, les manuels, les enseignants, la classe, le regroupement des élèves, n'expliquent pas plus de 20 % de la variance des performances scolaires (mesurables) [CREEMERS, 1991, p. 9]. Le verdict est clair : 80 % de cette variance dépendent de l'environnement social et des capacités.

On arrive à la conclusion, assez choquante, qu'il n'est possible d'agir que sur 20 % des différences (mesurables) pour améliorer la situation par l'augmentation de la qualité propre de l'enseignement.

Tout cela ne doit pas occulter le fait qu'au cœur de ce mode de pensée, se trouve la recherche assidue de méthodes statistiques qui permettraient de corriger les facteurs

de l'environnement et de l'intelligence. Dans un paragraphe antérieur, nous avons évoqué les tentatives visant à constituer des catégories d'écoles sur la base de l'environnement socio-économique. Mais là encore, on note une tendance à la simplification. Comment définit-on un milieu socio-économique défavorisé ? Il est bien trop simple de le définir comme un milieu regroupant des élèves issus de familles avec des parents sans instruction (père ou mère) ou des enfants qui fréquentent les cantines scolaires subventionnées. Là encore, la réalité est bien plus complexe que les acrobaties sur les chiffres pourraient le laisser penser.

### **Des résultats corrigés pour tenir compte de l'environnement social**

La communauté pédagogique est unanime à ne pas vouloir comparer directement les résultats des établissements scolaires, mais à proposer d'introduire un facteur de correction pour tenir compte des caractéristiques personnelles, et plus particulièrement de l'intelligence et du milieu social. Comme nous l'avons déjà mentionné, les résultats scolaires sont souvent liés au milieu socio-économique dont les élèves sont issus : les facteurs de pauvreté et de stimulation au sein de la famille sont très importants pour déterminer les différences entre écoles.

La question logique qui se pose ensuite est donc de savoir comment définir un milieu social plus défavorisé et comment le quantifier. Les indicateurs classiques utilisés à cet égard sont la profession ou le niveau d'éducation des parents, ou, comme c'est aujourd'hui la norme en Angleterre, le fait de savoir si les élèves sont issus de familles qui sont en droit de bénéficier des cantines scolaires subventionnées. Il s'agit en réalité d'un indicateur d'exclusion sociale et de pauvreté.

Il semble de plus en plus certain que cette catégorisation est trop

imprécise. La pauvreté et l'exclusion sociale sont liées à un vaste éventail de facteurs. Une étude à grande échelle récente du *New Policy Institute* de Londres (Howarth, 1998) a identifié pas moins de 50 indicateurs-clés de la pauvreté et de l'exclusion sociale dans la population. Quatorze de ces indicateurs sont en rapport avec les enfants et les jeunes gens. Nous les résumons ci-après :

- chômage de l'un ou des deux parents ;
- revenu inférieur à la moitié du revenu moyen ;
- bébés à faible poids de naissance (25 % de plus chez les mères des classes sociales défavorisées) ;
- accidents mortels (deux fois plus nombreux dans les classes sociales défavorisées) ;
- enfants qui quittent l'enseignement sans qualification ;
- enfants renvoyés de l'école (trois fois plus nombreux dans la population noire) ;
- enfants de parents divorcés (deux fois plus nombreux chez les couples mariés sans instruction) ;
- filles-mères de moins de 16 ans (4,3 sur 1 000 naissances) ;
- enfants avec des troubles comportementaux en instituts ;
- chômage des jeunes ;
- difficultés de crédit bancaire des jeunes employés ;
- situation financière médiocre des 16-17 ans ;
- périodes de traitements médicamenteux ;
- suicide (60 % de plus dans les groupes plus défavorisés) ;
- jeunes de 19 ans sans aucune qualification (30 % en Angleterre) ;
- jeunes de 23 ans avec un casier judiciaire.

Tous ces indicateurs peuvent, bien sûr, être présents en même temps et il est évidemment aussi possible d'y ajouter les 36 indicateurs de pauvreté et d'exclusion sociale qui n'apparaissent pas ici.

La question qui nous préoccupe est de savoir comment il est possible d'exprimer en chiffres toutes les

combinaisons et différences possibles afin de redresser les résultats scolaires en fonction du milieu social.

Il faut dans ce cas prendre des décisions subjectives afin de calculer le rôle joué par chacun des facteurs contributifs. Si l'on procède ensuite au traitement informatique de ces chiffres, il devient réellement possible de parler d'acrobaties statistiques. Peut-être ces ambitions s'apparentent-elles même à de l'arrogance scientifique.

## **LES EFFETS CONNEXES DES TESTS CENTRALISÉS ET DES CLASSEMENTS**

Les tests centralisés dont les résultats doivent être publiés pour chaque établissement scolaire aboutissent inévitablement à des hit-parades des établissements, d'une part, et à la disgrâce de ceux qui ont de faibles scores, d'autre part. Les effets connexes de ces tests sont si lourds que les effets positifs possibles des tests en termes de miroir de la réalité disparaissent sans laisser de traces.

L'effet connexe le plus important est celui de réduire les objectifs de l'éducation à un niveau mesurable. Or, il n'est un secret pour personne qu'un nombre considérable d'objectifs de nature plus noble (la pensée créative, les facultés d'apprentissage), et également d'objectifs relevant des attitudes, sont extrêmement difficiles à mesurer. Le résultat est évident : l'enseignement en est réduit à préparer les élèves aux tests. L'éducation au sens large, la responsabilité civile et la création de valeurs deviennent automatiquement des objectifs de second rang.

Il est également très important de délimiter de manière très stricte le rôle des enseignants et des chefs d'établissement. Lorsque l'on se situe dans le cadre de la préparation à des examens qui ne cessent de se répéter,

on attend d'un enseignant une intervention strictement technique. Les signaux envoyés par la société donnent en revanche une image opposée. Les gens exigent de la créativité et de la souplesse de la part des enseignants dans leur travail avec les élèves et attendent d'eux qu'ils transmettent aux élèves bien plus que la simple connaissance d'une discipline et des expériences d'éducation cognitive. Si cette exigence est maintenue, il va falloir admettre une part d'imprévu dans l'éducation et accepter également l'auto-détermination des enseignants.

FITZ-GIBBON a dressé une liste supplémentaire d'effets connexes indésirables (1997) :

- Une *vision étroite* : la qualité se trouve réduite aux résultats d'examens d'un certain type. Les autres indicateurs sont annexes. Même l'accent mis sur les matières théoriques par opposition aux matières pratiques est l'une des conséquences ;

- La *sous-utilisation* : les établissements vont se concentrer sur les éléments de leur organisation qui vont leur permettre d'augmenter leur score moyen. Les élèves difficiles vont, par exemple, être renvoyés sans plus de détours ; un certain type de sélection à l'entrée va être organisé ; une attention étroite va être accordée aux élèves qui sont juste au-dessous de la moyenne, parce que ce sont eux qui vont générer le plus fort taux de succès. Il va peut-être être possible d'intervenir sur la difficulté des tests (en donnant des conseils aux élèves ; en jouant sur le temps accordé pour faire le test, etc.). Les enseignants vont être moins enclins au travail d'équipe, ils vont même plus ou moins l'abandonner, dans la mesure où ils seront évalués essentiellement sur leur contribution individuelle aux scores des élèves. En résumé, les enseignants vont avoir tendance à ne s'occuper que des élèves « alimentaires » pour eux, c'est-à-dire les élèves qui amènent les bons résultats ;

- La *myopie* : les gens ne s'intéressent qu'aux résultats des tests. Les résultats à long terme concernant, entre autres, les carrières professionnelles, la responsabilité civile, le comportement social, ne sont pas pris en compte. Alors que la société devrait s'intéresser aux effets à long terme de l'éducation ;

- La *fixation sur les notes* : la fixation sur des résultats mesurables met principalement l'accent sur les notes elles-mêmes et pas sur la réalisation des objectifs. Il est particulièrement fréquent, dans les sciences comportementales, que les gens ne puissent pas procéder à des mesures simples qui génèrent des réponses identiques. C'est également ce qui se passe avec des questions relativement simples. Par exemple, la question de savoir pourquoi un élève ne répond pas à certaines questions, ou même celle de savoir s'il réussit mieux que prévu ou moins bien, ne sont pas posées. En d'autres termes, un chiffre ne donne que trop peu d'informations sur une performance.

Il y a en plus un certain nombre de problèmes plus pratiques. Pour commencer, il y a l'organisation centralisée et qui prend beaucoup de temps. Les tests doivent rester confidentiels et les auteurs doivent s'engager à garder à 100 % le secret. Les sujets des tests doivent être imprimés sous le plus strict contrôle afin d'éviter les fuites. La distribution doit se faire simultanément et il ne doit y avoir aucun accroc. Et enfin, les tests doivent être passés en même temps. La pression organisationnelle sur les établissements est également très visible. Les établissements doivent absolument respecter les dates fixées pour les tests. Cette organisation qui prend du temps et la charge administrative qui lui est associée entraînent des coûts considérables. Et enfin, chaque test peut toujours être contesté quant au choix du sujet et à la construction du sujet. Des discussions sur la validité, la fiabilité et la représentativité des tests sont toujours possibles. Le test

absolument juste n'existe tout simplement pas.

Même lorsque les gens décident de passer à l'évaluation des gains de l'enseignement afin de corriger les indicateurs aléatoires, il existe malgré tout encore un certain nombre d'effets connexes indésirables. Il a été montré que des établissements s'arrangent pour obtenir des scores médiocres lors du test initial afin de pouvoir ainsi générer de meilleures performances éducatives. Et même dans les performances éducatives, on ne peut pas éviter que les élèves intelligents avancent plus vite que les élèves moins doués. Les élèves plus faibles sont souvent issus de milieux défavorisés, donc il existe une différence, déterminée par le milieu, dans la rapidité des performances éducatives. En ce qui concerne la mesure des progrès éducatifs, des exigences encore plus grandes sont formulées en matière de construction des tests. Les examens ou tests centralisés qui sont publiés, que ce soit ou non dans le cadre de la mesure des progrès éducatifs, sont tellement importants pour l'image d'une école qu'ils peuvent aboutir à une iniquité dans la formation et à des stratégies défensives afin d'obtenir les scores les plus élevés possibles.

Les avantages des tests centralisés sont l'objectivité des scores, la meilleure validité du contenu et du concept et la comparabilité des scores. Les inconvénients soulignés ici font pencher la balance dans le mauvais sens. La comparaison forcée des établissements n'incite pas à l'amélioration de la qualité des missions nobles de l'éducation. Seuls des objectifs limités, faciles à mesurer, et l'éducation à court terme vont tirer bénéfice de la comparaison des établissements. Il faut donc trouver une méthode de travail alternative.

## L'INTÉGRATION DE L'ÉVALUATION INTERNE ET DE L'ÉVALUATION EXTERNE

Il serait erroné de conclure de ce qui précède qu'il ne faut plus donner de notes. Cela aussi est une façon simpliste de voir les choses. Les processus de qualité progressent réellement lorsque les gens réussissent à quantifier certains aspects et indicateurs. Mais plus le comportement à mesurer devient complexe, moins il se prête à une quantification précise. Il faut tenir compte de cela dans l'interprétation des choses. Sous réserve de la prise en compte de ces limites, les analyses chiffrées sont très utiles pour mesurer sa propre pratique à l'aune des scores moyens ou des scores obtenus par des établissements comparables. Lorsque ces éléments sont utilisés par les établissements pour leur propre auto-évaluation, ils peuvent constituer une aide précieuse. Mais lorsqu'ils sont utilisés pour évaluer les établissements au travers d'un classement (ce qui est en grande partie le cas avec les tests centralisés), ils ne sont pas fonctionnels, au contraire, ils sont une entrave à l'amélioration de la qualité.

La mesure des performances n'est donc absolument pas inconvenante si les résultats de cette mesure sont utilisés à des fins d'auto-évaluation. Cette auto-évaluation repose toutefois sur un plus grand nombre d'éléments que simplement les performances de l'établissement.

### Les objectifs de la mesure des performances

Selon le Département du développement éducatif en Flandres (Dienst voor Onderwijsontwikkeling, DVO, 1998), la mesure des performances doit, pour être utile, satisfaire trois objectifs.

*Objectif 1 : Surveiller la réalisation du minimum.* Les autorités ont imposé des objectifs minimums afin de garantir une qualité minimale comparable pour tous les établissements scolaires du pays. Elles doivent donc avoir à leur disposition un instrument qui leur permette d'évaluer si les objectifs fixés ont été atteints. On espère que l'évaluation de la réalisation des objectifs aura un effet stimulant sur leur mise en œuvre.

*Objectif 2 : Contrôler les performances de chaque établissement.* Un établissement a de bonnes performances s'il atteint un niveau élevé de réalisation des objectifs – en tenant compte des caractéristiques du recrutement et du milieu socio-économique – et lorsque les élèves enregistrent des progrès sur le long terme plus importants que ceux que l'on attendait initialement d'eux (au regard des caractéristiques du recrutement et du milieu socio-économique). Il s'agit ensuite de contrôler le *profit éducatif* ou la valeur ajoutée de l'établissement pour un élève donné. Cela signifie que les résultats doivent être mesurés au moins à deux moments différents afin d'évaluer la progression. En utilisant ce principe, les écoles ne sont plus comparées les unes aux autres, mais chaque école mesure uniquement la progression de ses élèves.

*Objectif 3 : Conduire une politique et l'adapter.* Il est important pour les autorités de surveiller dans quelle mesure les objectifs éducatifs ou de développement sont réalisés au niveau du système. Cela peut se faire sans comparer les établissements.

Dans le cadre de la réalisation de ces objectifs, le préalable est une vision de la qualité qui soit en rapport avec l'approche didactico-pédagogique décrite précédemment. Le professionnalisme des enseignants et des chefs d'établissement est ensuite confirmé, voire même amélioré. De plus, l'école locale ou le lycée est au centre de la

préoccupation pour la qualité. Chaque école a sa propre nature, son propre recrutement et son propre environnement. Chaque école doit donc être jugée séparément. En d'autres termes : chaque école a sa propre qualité.

Une conséquence immédiate de cette approche est la priorité accordée à l'évaluation interne sur l'évaluation externe. Il doit y avoir une interaction manifeste entre les deux. L'évaluation externe présente à l'école un miroir qui peut aider celle-ci à réaliser sa propre évaluation de qualité. À l'inverse, l'évaluation interne envoie des signaux aux évaluateurs externes quant à la faisabilité et à l'utilité des objectifs à poursuivre et à réaliser.

### Les instruments de l'évaluation interne

L'école peut réaliser elle-même intégralement son évaluation interne, de différentes façons. Comme c'est déjà le cas, l'école utilise des tests et des examens pour mesurer les performances éducatives. Si l'on s'intéresse à l'expérience, aux attitudes ou à des processus de réflexion supérieurs, il faudra utiliser divers systèmes d'observation. Pour pouvoir être capable de suivre les gains éducatifs en plus des performances éducatives, des systèmes de suivi des élèves seront un atout supplémentaire à l'avenir.

En plus des performances éducatives, l'école utilisera naturellement tous les autres indicateurs qui déterminent la qualité des processus et de l'organisation, afin de contrôler leur présence et de les suivre.

### Les instruments de l'évaluation externe

Les autorités organisent l'évaluation externe essentiellement par le biais des inspections. Une équipe d'inspection professionnelle doit élaborer les procédures et les méthodes d'organisation afin d'obtenir un

enseignement satisfaisant. L'inspection scolaire va par ailleurs étudier le système d'évaluation interne de l'école : les tests, les examens, les systèmes de suivi des élèves et les moyens d'observation seront examinés en termes de qualité.

Pour ce qui concerne les performances éducatives, les autorités peuvent réaliser des inspections périodiques. L'efficacité des objectifs finaux peut être évaluée au niveau du système au moyen de contrôles aléatoires représentatifs mais anonymes.

Les inspecteurs peuvent utiliser une banque d'éléments pour déterminer le résultat à certaines échéances finales afin que l'inspection de l'école joue un rôle d'exemple. Ces éléments peuvent servir de miroir pour l'école. Ils peuvent également être rassemblés dans un rapport annuel sur les performances éducatives au niveau du système.

### **L'intégration de l'évaluation interne et de l'évaluation externe**

Si la qualité de l'évaluation interne augmente, la part de l'évaluation externe va pouvoir diminuer.

Si les gens souhaitent asseoir le principe de la priorité de l'évaluation interne sur l'évaluation externe, il faudra s'efforcer de parvenir à un plus grand professionnalisme des établissements en ce qui concerne l'auto-évaluation. La qualité des tests et des examens doit être augmentée. Les écoles devront également apprendre à recueillir des informations sur l'enseignement pratique et le développement des attitudes. Les écoles devront également apprendre à se remettre en cause de manière critique concernant un certain nombre d'indicateurs de la progression de la qualité (par exemple pour la formation continue, les rapports, les contacts avec les parents ou les personnes en charge des élèves, la qualité des discussions pédagogiques, etc.). Le développement de bons systèmes de suivi des élèves doit être encouragé, probablement avec le soutien des autorités. Les systèmes de suivi des élèves sont exclusivement réservés à l'usage interne. Afin de préserver leur efficacité, ils ne doivent en aucun cas être utilisés pour l'évaluation externe et, *a fortiori*, pour le classement des établissements scolaires. Il faudrait envisager de fournir une banque d'items

d'évaluation aux écoles pour qu'elles les utilisent, si elles le souhaitent, pour l'évaluation de la réalisation des objectifs des élèves.



Avec un modèle intégré d'évaluation interne et externe des performances des élèves et d'autres indicateurs de la qualité, un choix explicite a été fait en faveur d'un type de contrôle de qualité qui soit didactico-pédagogique. Cela signifie que la priorité du contrôle de qualité est entre les mains de l'école, le personnel enseignant et les chefs d'établissement intervenant comme professionnels. L'individualité de chaque école est ainsi fortement assurée.

Les comparaisons malhonnêtes et génératrices de résistances sont remplacées par un système modéré d'évaluation externe par le biais des inspections scolaires et d'enquêtes périodiques. Ces inspections et enquêtes suffisent à garantir les exigences légitimes de qualité, à les placer avant l'individualité de l'école et, en faisant cela, elles servent un intérêt social général. ■

CREEMERS, B.P.M. (1991). Une instruction efficace (publié en néerlandais : *Effectieve instructie*). La Haye : Instituut voor Onderzoek van het Onderwijs.

DE BOECK, P. (1992). Une évolution majeure dans la mesure des connaissances et des compétences intellectuelles (publié en néerlandais : *Een ommekeer in het meten van kennis en intellectuele vaardigheden*). Louvain : K.U., Département de Psychologie.

Dienst voor onderwijsontwikkeling (1998). Un modèle intégré pour la mesure des performances dans le contrôle de qualité et le contrôle de qualité gouvernemental (publié en néerlandais : *Een geïntegreerd model van prestatiemetingen voor kwaliteitsbewaking en kwaliteitsbewaking vanuit de overheid*). Bruxelles : DVO.

EISNER, E.W. (1977). *On the uses of educational connoisseurship and criticism for evaluating classroom life*. Teachers College Record. pp. 345-358 (78).

FITZ-GIBBON, C.T. (1997). *The value added national project*. Londres : Schools Curriculum and Assessment Authority.  
GOODLAD, J.I. (1975). *Schools can make a difference*. Educational Leadership, (33).

HOWARTH, C. et coll. (1998). *Key indicators of poverty and social exclusion*. Londres : New Policy Institute.

HOYWEGHEN, D. VAN (1997). Un corps d'inspecteurs d'un style nouveau en Angleterre et au Pays de Galles (publié en néerlandais : *Inspectie nieuwe stijl in Engeland en Wales*). Impuls, (28), n° 1, pp. 9-17.

National commission on excellence (1983). *A nation at risk. The imperative for education reform*. Washington : US Government Printing Office.

PUTTEN, C.M. VAN & KAMP, L.J.T. van der, (1998). Dronkers note les écoles pour l'enseignement secondaire : précision, validité et sens (publié en néerlandais : *Dronkers' rapportcijfer voor de middelbare scholen: nauwkeurigheid, geldigheid en betekenis*). Pedagogische Studiën, (75), n° 2, pp. 135-142.

ROBINSON, P. (1998). *Literacy, numeracy and economic performance*. Londres : London School of Economics.

SIKKES, R. (1998). L'offensive de la qualité. Le résultat est ce qui compte (publié en néerlandais : *Het kwaliteitsoffensief. Het resultaat telt*). (MESO focus). Houten : Educatieve Partners.

SILVERNAIL, O.L. (1996). *The impact of England's national curriculum and assessment system on classroom practice*. Educ. Policy, n° 1, pp. 46-62.

STANDAERT, R. (1989). Innovation dans l'éducation et politique éducative en Angleterre et au Pays de Galles (publié en néerlandais : *Onderwijsinnovatie en -beleid in Engeland en Wales*). Louvain : K.U. Leuven.

STANDAERT, R. & TROCH, F. (1998). Apprendre et enseigner (publié en néerlandais : *Leren en onderwijzen*). Louvain - Amersfoort : Acco. 4<sup>e</sup> édition révisée.

STANDAERT, R. (2001). *Inspectors of schools in Europe. A critical analysis*. Louvain - Amersfoort : Acco.

