



## LA CAQ ET LES TESTS STANDARDISÉS EN ÉDUCATION

8 décembre 2011

*« Ce qui a l'air intelligent dans une conversation, une réunion, et en particulier dans les médias, est tout particulièrement à prendre avec des pincettes. »*

NASSIM NICHOLAS TALEB, *Le Hasard sauvage*, 2009

### Introduction

Avant de se transformer en parti politique officiel le 14 novembre 2011, la Coalition Avenir Québec (CAQ) a présenté, au cours de cette même année, quatre documents de réflexion exposant sa vision, ses positions et ses recommandations sur certains enjeux québécois. L'un de ces documents portait sur le système d'éducation et contenait quelques propositions. Parmi celles-ci, on retrouve le nouveau pacte en enseignement et la réorganisation du réseau par l'abolition des commissions scolaires. Le présent document porte sur le nouveau pacte en enseignement (ci-après nommé le « Pacte ») et plus précisément sur l'utilisation de tests standardisés pour évaluer les écoles et les enseignants. Ce document est divisé en trois parties. La première décrit brièvement la nature de ce Pacte, la seconde présente les origines de cette proposition et la troisième en fait une analyse critique.

## **Partie 1 : Le nouveau pacte en enseignement de la CAQ**

Le Pacte propose une augmentation moyenne de 20 % des salaires des enseignants en contrepartie de la mise en place d'un système d'évaluation de leur performance et de celle des écoles. À cet égard, le texte de consultation de la CAQ mentionne que : « Chaque enseignant verrait son salaire augmenté, mais les augmentations pourraient varier afin de cibler les écoles jugées plus difficiles<sup>1</sup>. »

Toujours selon ce même document, le système d'évaluation des enseignants, mis en place pour « mesurer leurs compétences et leur motivation », serait fondé sur des évaluations bisannuelles effectuées par les directions d'école et porterait sur plusieurs aspects, notamment le taux de réussite des élèves aux tests standardisés et la satisfaction des parents. Ainsi, un enseignant qui n'atteindrait pas le niveau de performance souhaité recevrait, dans un premier temps, du soutien (perfectionnement, mentorat ou jumelage) et pourrait perdre son poste si la situation ne s'améliorait pas.

Face au tollé soulevé par cette proposition, François Legault a profité de la création de son parti politique en novembre 2011 pour modifier le tir quelque peu. Le plan d'action présenté au lancement de la CAQ mentionne que l'évaluation et la rémunération des enseignants ne seraient pas liées aux résultats des évaluations uniformes : « Le droit à une éducation de qualité exige d'évaluer le travail des enseignants. L'objectif de cette mesure n'est pas de lier le salaire à un quelconque indice de performance, de congédier des enseignants ou de menacer leur sécurité d'emploi<sup>2</sup>. » Il est donc encore difficile de connaître avec certitude la nature et les modalités exactes du système d'évaluation des enseignants que la CAQ souhaite mettre en place.

En ce qui a trait à la gestion des écoles, le Pacte prévoit le remplacement de l'actuel système de reddition de comptes par un « cadre de gestion de la performance axée sur les résultats<sup>3</sup>. » Ce cadre viserait « à mesurer les résultats des différents établissements selon leur milieu socioéconomique et leur clientèle<sup>4</sup> ». Les indicateurs d'un tel cadre seraient « basés sur de nouvelles évaluations annuelles standardisées, administrées dès la 1<sup>re</sup> année. Les tests auraient lieu chaque année dans les matières de base et seraient corrigés de façon indépendante<sup>5</sup> ». Le document de réflexion sur l'éducation et le plan d'action de la CAQ ne mentionnent pas clairement si le financement des écoles et la répartition des ressources entre elles seraient tributaires des résultats aux tests. Cependant, compte tenu de la feuille de route de François Legault en la matière, il y a tout lieu de croire que c'est un pas qu'il pourrait aisément franchir<sup>6</sup>.

---

<sup>1</sup> Coalition Avenir Québec, 2011a, p. 4.

<sup>2</sup> Coalition Avenir Québec, 2011b, p. 6.

<sup>3</sup> Coalition Avenir Québec, 2011a, p. 6.

<sup>4</sup> Ibidem, p. 6.

<sup>5</sup> Ibidem, p. 6.

<sup>6</sup> Lors de son passage au ministère de l'Éducation, François Legault a notamment mis en place les contrats de performance avec les universités québécoises en vertu desquels le financement de ces institutions était lié à l'atteinte de résultats.

## **Partie 2 : Les origines de cette proposition**

Pourtant présentée comme une nouveauté, l'idée d'utiliser des tests uniformes pour évaluer la performance des écoles et des enseignants n'est pas nouvelle. Chaptal et collab. (2008) fait un intéressant résumé des tentatives antérieures d'application de ce type de système d'évaluation, dont le premier répertorié est le *Revised Education Code* de Grande-Bretagne. De 1862 à 1897, le réseau scolaire britannique a utilisé des tests uniformes afin d'évaluer la performance des enseignants. Ces tests portaient sur la lecture, l'écriture et l'arithmétique. Leur utilisation n'a cependant pas eu les résultats escomptés. D'importants effets pervers furent constatés tels que la tricherie, l'appauvrissement pédagogique et le désintérêt envers les élèves ne permettant pas d'atteindre les objectifs de performance. Durant cette période, on a également constaté que le système d'éducation britannique a accumulé, de façon générale, un important retard sur ceux en usage dans l'Europe continentale (Chaptal et collab., 2008, p. 19). Sous la pression des enseignants et des inspecteurs scolaires, ce système d'évaluation a été abandonné une trentaine d'années après son implantation.

D'autres tentatives d'évaluation des enseignants ont également été mises en place au début du 20<sup>e</sup> siècle aux États-Unis et plus tard sous la gouverne du président Nixon, mais ont rapidement été abandonnées. Cependant, c'est vraiment la loi du président Bush *No Child Left Behind* (NCLB), votée par le Congrès en 2001, qui a remis l'utilisation des tests standardisés à l'avant-scène des politiques publiques en éducation. Ce programme, dont la création fut appuyée par les démocrates au Congrès américain, prévoit des investissements importants du gouvernement fédéral au bénéfice des États pour financer l'école publique en contrepartie de la mise en place d'un système d'évaluation par tests standardisés permettant d'évaluer l'atteinte d'objectifs de résultats dans certaines matières scolaires<sup>7</sup>.

En 2008, le président Obama continue sur la même lancée avec son programme *Race to the Top*. Ce programme octroie aux États d'importantes subventions fédérales destinées à l'éducation publique. En contrepartie, les États doivent toutefois instaurer un système de rémunération des administrateurs scolaires et des enseignants basé en tout ou en partie sur les résultats des élèves aux tests standardisés. Afin de bénéficier des sommes mises à leur disposition, plusieurs États ont récemment instauré de tels systèmes d'évaluation et de rémunération (Tienken, 2011).

Aujourd'hui, l'utilisation des tests standardisés n'est plus uniquement l'apanage des États-Unis. Les pays de l'Organisation de coopération et de développement économiques (OCDE) sont de plus en plus nombreux à les utiliser à diverses fins, notamment pour évaluer, récompenser ou punir les enseignants, les administrateurs scolaires ou les écoles (Morris, 2011).

---

<sup>7</sup> L'objectif du programme NCLB est que 100 % des enfants soient compétents (« *profficient* »), mais laisse aux États le soin de déterminer exactement ce que « *profficient* » veut dire.

### **Partie 3 : Reddition de comptes et rémunération basées sur les tests standardisés**

L'utilisation de tests standardisés afin d'évaluer les écoles ou de rémunérer les enseignants au mérite est une méthode découlant directement des principes de la Nouvelle gestion publique (NGP) et de la gestion par résultats. La NGP est une philosophie de gestion des organisations publiques copiée sur celle du secteur privé. Appliquée au secteur de l'éducation, cette philosophie de gestion s'est notamment traduite par des réformes visant la décentralisation vers les établissements scolaires, la concurrence entre les écoles, le libre choix des parents, la détermination d'objectifs de rendement, l'évaluation des résultats et l'attribution d'enveloppes budgétaires ou de primes liées aux performances des écoles ou des enseignants (Chaptal et collab., 2008, p. 39).

Selon cette philosophie de gestion, plus de latitude doit être accordée aux écoles dans leur gestion quotidienne (la décentralisation), mais en contrepartie ces dernières doivent atteindre des objectifs fixés par l'État et dont la réalisation est mesurée à l'aide de tests standardisés (gestion par résultat). Des incitatifs financiers (primes, récompenses, etc.) visant à favoriser l'effort des enseignants et la réussite des élèves sont parfois également mis en place. Le raisonnement qui sous-tend l'instauration de ces incitatifs est simple. Il tient pour acquis que si les enseignants sont évalués, rémunérés ou sanctionnés sur la base des résultats des élèves, ils seront encouragés à travailler plus fort et mieux afin d'améliorer la performance des enfants.

Séduisantes en apparence, ces idées ne tiennent pas la route quand on y regarde de plus près. L'utilisation de tests standards comme un mécanisme de reddition de comptes pour les écoles ou en tant qu'outil d'évaluation de la performance des enseignants comporte d'importants effets pervers qui affectent le travail en classe et le système scolaire dans son ensemble. Ces effets sont par ailleurs exacerbés lorsque le financement des écoles ou la rémunération des professeurs sont liés aux résultats des élèves à ces tests. En outre, la littérature scientifique récente est ambiguë quant à l'effet des systèmes de rémunération au mérite sur les résultats des élèves. Un nombre croissant d'études démontrent en effet que les incitatifs financiers proposés aux enseignants n'ont pas d'impact sur les résultats des élèves. Ainsi, la CAQ semble proposer une réforme qui ne produit pas les fruits escomptés et dont les effets pervers sont importants et de mieux en mieux documentés.

Ces effets pervers et l'inefficacité des systèmes de tests standardisés et de rémunération au mérite sont d'ailleurs au cœur du débat en cours aux États-Unis sur les programmes *No Child Left Behind* et *Race to the Top*. Plusieurs spécialistes dénoncent leur utilisation dans le réseau scolaire américain. Diane Ravitch, vice-ministre de l'Éducation des États-Unis de 1991 à 1993, est l'une de ces spécialistes qui remettent en question toute la philosophie de gestion basée sur les tests standardisés. Au début des années 2000, madame Ravitch était favorable au programme *No Child Left Behind* et appuyait le principe de la rémunération au mérite. Cependant, à la lumière des résultats de ces politiques, elle a radicalement changé d'avis. Elle

évoque les raisons de cette volte-face dans *The Death and Life of the Great American School System*, un essai publié en 2010. Elle y écrit<sup>8</sup>:

J'ai constaté avec désarroi que les tests étaient devenus une préoccupation centrale dans les écoles et n'étaient plus seulement un outil de mesure, mais une fin en soi (Ravitch, 2010, p. 12).

Un peu plus loin, elle poursuit :

J'ai vu mes espoirs pour une meilleure éducation se transformer en simple stratégie de mesure sans aucune vision éducationnelle. Finalement, j'ai réalisé que ces nouvelles réformes n'étaient que des réformes structurelles et comptables, n'ayant rien à voir avec les fondements de l'apprentissage (Ravitch, 2010, p. 16).

### **L'évaluation des enseignants : Mesurer le « non-mesurable »<sup>9</sup>?**

À la base des systèmes d'évaluation des enseignants se trouvent deux postulats :

1. Les enseignants ont un impact sur la performance des élèves ;
2. Il est possible d'évaluer cet impact à l'aide d'indicateurs.

Parmi les indicateurs fréquemment utilisés, la réussite ou la performance des élèves aux tests standardisés occupe une place de choix. L'évaluation des enseignants est d'ailleurs l'une des raisons qui motivent la CAQ à vouloir instaurer l'usage de tests standardisés. Pourtant, la réussite ou la performance d'un enfant dépend de plusieurs facteurs tels que le milieu socio-économique, la maîtrise de la langue (déterminant pour les enfants issus de l'immigration) ou encore l'implication des parents. Il est alors très difficile d'établir un lien de causalité entre l'effort ou la compétence d'un enseignant et la réussite des enfants.

Conscients que la réussite scolaire repose sur une multitude de facteurs, les partisans de l'évaluation basée sur les tests standardisés proposent d'utiliser les modèles d'évaluation dits « à valeur ajoutée » (en anglais « *value-added model* » ou « *growth model* »). Ces modèles tentent d'établir un lien de causalité entre les résultats des enfants et le travail des enseignants. Pour ce faire, différents moyens sont utilisés afin d'isoler le travail des enseignants des autres facteurs de réussite scolaire. Le moyen proposé par la CAQ consiste à moduler les cibles ou objectifs de performance en fonction du milieu socio-économique des élèves. Par exemple, les objectifs de réussite des élèves ne seraient pas les mêmes à Westmount que dans le quartier Hochelaga-Maisonneuve. Évidemment, la détermination d'objectifs équitables pour tenir compte des facteurs socio-économiques est un exercice éminemment difficile.

On apprenait récemment, dans un reportage consacré à l'évaluation des enseignants américains diffusé à l'émission *Une heure sur Terre* sur Radio-Canada<sup>10</sup>, qu'un tiers des

---

<sup>8</sup> Les deux passages suivants sont une traduction libre de l'auteur.

<sup>9</sup> Ce sous-titre est inspiré de l'article de Ellen R. Delisio accessible à l'adresse Internet suivante : [http://www.educationworld.com/a\\_issues/issues/issues374a.shtml](http://www.educationworld.com/a_issues/issues/issues374a.shtml) .

enseignants des quartiers les plus riches de la ville de Washington étaient considérés parmi les 25 % « meilleurs enseignants », alors que seulement 5 % des enseignants des quartiers les plus pauvres appartiennent à ce groupe sélect. Cet exemple démontre bien qu'il semble très difficile de moduler les cibles de résultats de façon à rendre le processus d'évaluation équitable envers tous les enseignants.

Un autre type de modèles d'évaluation à valeur ajoutée consiste à suivre des élèves de façon plus régulière et sur quelques années afin d'évaluer leur progrès. Les élèves étant suivis, il est présumé que l'amélioration des résultats découle du travail des enseignants. Ces modèles d'évaluation sont complexes et coûteux puisqu'ils nécessitent un suivi serré des élèves et plusieurs évaluations. Il demeure par ailleurs difficile d'établir un lien direct entre le progrès scolaire des élèves et le travail d'un enseignant en particulier, puisque les enseignants ne suivent généralement pas les élèves d'une année à l'autre.

En somme, pour la spécialiste Diane Ravitch, même « les modèles à valeur ajoutée font face à d'innombrables problèmes techniques<sup>11</sup> » (Ravitch, 2010, p. 187). Plus généralement, Ravitch soutient que les tests standardisés ne sont pas un instrument adéquat pour mesurer la performance des enseignants :

Le problème avec l'utilisation de tests standardisés, quand on les utilise afin de prendre des décisions qui ont une grande importance sur la vie des gens, est qu'ils ne sont pas un instrument de mesure précis. Malheureusement, les élus et la population en général ne le réalisent pas. Le public pense que les tests ont une validité scientifique, comme un thermomètre ou un baromètre, qu'ils sont objectifs et qu'ils ne sont aucunement affectés par le jugement humain faillible. Mais les résultats à ces tests ne sont pas comparables à ces instruments de mesure, ils n'ont pas la précision d'une balance chez le médecin ou d'un ruban à mesurer...

Les experts en matière de tests nous préviennent que les résultats à ces examens ne devraient pas être utilisés à d'autres fins que celles pour lesquelles ils ont été développés. Par exemple, un test de lecture d'un élève de cinquième année mesure les habiletés de lecture d'un élève de cinquième année et ne peut être utilisé de manière fiable comme mesure de la compétence de l'enseignant<sup>12</sup> » (Ravitch, 2010, p. 152).

L'utilisation de tests standardisés comme un outil d'évaluation des enseignants et l'établissement d'un lien de causalité entre le travail de l'enseignant et les résultats des enfants sont donc très contestés. L'efficacité de la rémunération au rendement basée sur la performance des étudiants l'est également.

---

<sup>10</sup> Le reportage fut diffusé le 25 novembre 2011 et peut être consulté à l'adresse Internet suivante : [http://www.radio-canada.ca/emissions/une\\_heure\\_sur\\_terre/2011-2012/](http://www.radio-canada.ca/emissions/une_heure_sur_terre/2011-2012/).

<sup>11</sup> Traduction libre de l'auteur.

<sup>12</sup> Traduction libre de l'auteur.

## Est-ce que ça marche? Des effets positifs non démontrés...

Dans une récente revue de littérature sur l'utilisation des tests standardisés, l'OCDE souligne l'ambiguïté des résultats des études empiriques portant sur l'efficacité de la rémunération au mérite (Morris, 2011, p. 35). Dans leur revue de littérature, Jones (2010) et Fryer (2011) arrivent eux aussi à la même conclusion. Jones écrit : « Les résultats empiriques de la rémunération au mérite sur la performance des étudiants sont peu concluants<sup>13</sup> » (Jones, 2010, p. 2).

Quelques études récentes apportent un éclairage fort intéressant quant à l'impact de la rémunération au mérite sur la performance des enfants. Pour l'une d'elles, Roland Fryer (2011), économiste de l'Université Harvard, a conduit la plus importante recherche empirique à ce jour aux États-Unis sur la rémunération au mérite. Le chercheur a divisé environ 20 000 enseignants provenant de 200 écoles de New York au sein de deux groupes. Le premier groupe se voyait appliquer un système de rémunération au rendement fondé sur les résultats de l'ensemble des élèves d'une école (« *school-based bonus*<sup>14</sup> ») et sur la satisfaction des parents et des élèves. Le second groupe, le groupe de contrôle, bénéficiait quant à lui, du système traditionnel de rémunération fondé sur l'expérience et le niveau de formation des enseignants.

Fryer résume ainsi les résultats de sa recherche :

Je ne trouve aucune preuve que les incitatifs financiers proposés aux enseignants augmentent la performance des élèves, leur présence en classe, l'obtention d'un diplôme, pas plus que je ne trouve de preuve que ces incitatifs influencent les comportements des élèves ou des enseignants<sup>15</sup> (Fryer, 2011, p. 1).

Springer et collab. (2010) rapporte les résultats d'une autre étude du même type (expérience aléatoire avec groupe bénéficiant de la rémunération au mérite et groupe de contrôle). Cette recherche a été conduite entre 2007 et 2009 auprès de 300 enseignants de mathématiques de la 5<sup>e</sup> à la 8<sup>e</sup> année œuvrant dans huit écoles publiques de Nashville au Tennessee. Les résultats de cette étude n'ont pas non plus été probants. Dans le communiqué de presse annonçant la publication de l'étude, l'auteur principal, Matthew Springer, affirme :

Nous avons testé la question fondamentale liée aux incitatifs à la performance – Est-ce que la rémunération au mérite suffit à elle seule à améliorer la performance ? – et nous avons trouvé que la réponse est non<sup>16</sup>.

---

<sup>13</sup> Traduction libre de l'auteur.

<sup>14</sup> Dans ce type de rémunération au mérite, tous les enseignants d'une école reçoivent une prime si les élèves de l'école atteignent un objectif de rendement basé sur leurs résultats aux tests.

<sup>15</sup> Traduction libre de l'auteur. Fait intéressant à noter, Fryer note à la page 6 que la nature collective de la prime ne peut expliquer l'inefficacité du système de rémunération au mérite.

<sup>16</sup> National Center on Performance Incentives, 2010. Traduction libre de l'auteur.

Glazerman et Seifullah (2010) rapportent quant à eux les résultats du programme TAP (*Teacher Advancement Program*) de Chicago. Ce programme offre aux enseignants des primes et des promotions dont l'attribution est basée sur leur contribution à la réussite des élèves à des tests standardisés et à une évaluation de leur performance en classe. Comme le mentionne l'étude :

L'idée sur laquelle repose le programme est que les incitatifs à la performance combinés aux outils mis à la disposition des enseignants pour se tenir au courant des performances et améliorer leur enseignement, devraient aider les écoles à attirer et à conserver les enseignants talentueux et aider les enseignants à hausser le taux de réussite des élèves<sup>17</sup> (Glazerman et Seifullah, 2010, p. xiii).

Une fois de plus, les résultats invalident la prémisse :

Deux années après l'introduction de TAP, nous ne trouvons aucune preuve que le programme a augmenté le résultat des élèves aux tests. Telle que mesurée par les résultats moyens en mathématiques et en lecture aux examens standards de l'Illinois<sup>18</sup>, la croissance du taux de réussite des élèves n'est pas significativement différente entre les écoles bénéficiant du programme TAP et les écoles comparables ne bénéficiant pas du programme<sup>19</sup> (Glazerman et Seifullah, 2010, p. xiii).

Finalement, Godmann et Turner (2010) arrivent à des conclusions similaires avec leur recherche portant sur le programme de rémunération au mérite<sup>20</sup> dans les écoles publiques de la ville de New York au cours des années 2007-2008 et 2008-2009. Les résultats montrent que ce programme n'a pas eu d'effets significatifs sur l'effort des enseignants, la performance des élèves en mathématiques et en lecture, les activités en classe, le taux de roulement des enseignants et la formation des nouveaux enseignants.

Nous arrêterons ici notre revue de la littérature sur l'efficacité de la rémunération au mérite pour hausser la réussite des élèves. Nous n'avons rapporté ici que quelques études montrant le peu d'effet de ce type de rémunération sur la performance des élèves. D'autres études concluent cependant à l'effet contraire. Les partisans des tests standardisés expliquent parfois cette ambiguïté de la littérature empirique par le fait que les incitatifs financiers mis en place sont parfois inadéquats et ne favorisent pas l'effort individuel des enseignants. On avance par exemple que les incitatifs de groupe au niveau de l'école (« *school-based bonus* ») n'auraient pas les effets escomptés parce qu'ils n'incitent pas à l'effort individuel.

Une interprétation diamétralement opposée des résultats peu concluants de ces systèmes de rémunération est que les enseignants, pourvu qu'ils se considèrent comme rémunérés correctement, sont peu motivés par les incitatifs financiers (Ramirez, 2010). Selon Ramirez, le choix de faire carrière dans l'enseignement se fait généralement en sachant très bien que la

---

<sup>17</sup> Traduction libre de l'auteur.

<sup>18</sup> *Illinois Standards Achievement Test*.

<sup>19</sup> Traduction libre de l'auteur.

<sup>20</sup> Comme pour Fryer (2011), il s'agissait ici de prime au niveau de l'école et non pas de prime individuelles.

profession ne procurera pas à ceux qui la choisissent des salaires mirobolants. En fait, ce que les enseignants rechercheraient avant tout, ce sont des conditions adéquates d'enseignement et un climat de coopération dans l'école. Aussi, « les mesures pour évaluer le mérite sont la plupart du temps considérées comme injustes et discrétionnaires » (Chaptal et collab., 2008, p. 43), ce qui mine la motivation des enseignants et l'efficacité des systèmes de rémunération au mérite.

En résumé, quelques que soient les raisons invoquées pour tenter d'expliquer la situation, il faut retenir que l'effet positif de la rémunération au mérite sur la performance des élèves est loin d'avoir été démontré. Aussi, il importe de mettre en relief ce constat avec les effets pervers de plus en plus documentés que génère l'utilisation des tests standardisés comme mécanisme d'évaluation des écoles et des enseignants.

### **Les effets pervers des tests standardisés**

L'introduction des tests standardisés pour évaluer les écoles ou les enseignants (qu'ils soient ou non liés à des avantages pécuniaires ou à des pénalités) change la dynamique dans les écoles. D'abord, les tests introduisent, sous diverses formes, une logique de compétition entre les écoles ou entre les enseignants, alors qu'il est reconnu que la coopération entre tous les acteurs du système scolaire est l'un des facteurs clés de la réussite.

Aussi, la littérature scientifique montre de plus en plus clairement que l'utilisation des tests standardisés comporte son lot d'effets pervers. Dans un rapport de l'OCDE, Allison Morris écrit :

Bien que les preuves sur l'efficacité des tests standardisés sur la performance des élèves soient mitigées, il y a plus de certitude à l'effet que ces tests augmentent les comportements stratégiques de la part des écoles et des enseignants. [...] En conséquence, quand les tests sont perçus comme pouvant avoir d'importantes conséquences pour les enseignants ou les écoles, il y a des risques que le système d'évaluation soit biaisé par les actions stratégiques suivantes : enseignement orienté vers la pratique des tests, réduction du curriculum, tricherie et exclusion des élèves<sup>21</sup> (Morris, 2011, p. 37).

Le biais causé par ces stratégies porte le nom anglais de « *score inflation* », c'est-à-dire le gonflement artificiel des résultats. Ce biais fait perdre toute précision aux tests en tant qu'outil d'évaluation de l'enseignement. En d'autres mots, ces stratégies gonflent artificiellement les résultats aux tests. En conséquence, ces tests ne fournissent pas d'information précise sur la qualité de l'enseignement, alors que c'est le postulat fondamental à la base de leur utilisation. Ce phénomène s'explique par la loi de Campbell (*Campbell's Law*) du nom du chercheur Donald T. Campbell, qui énonce que :

---

<sup>21</sup> Traduction libre de l'auteur.

Plus un indicateur social quantitatif est utilisé dans un processus décisionnel, plus il est susceptible de manipulation et plus il aura tendance à fausser et à altérer le processus social qu'il a pour objet d'évaluer [...] [Q]uand les résultats à un test deviennent le but du processus d'enseignement, ils perdent leurs valeurs en tant qu'indicateur du niveau d'apprentissage et altèrent le processus d'enseignement de manière indésirable<sup>22</sup>.

Voyons brièvement ces effets pervers.

### *La tricherie*

La tricherie est probablement le plus spectaculaire et le plus déplorable effet pervers de l'utilisation de tests standardisés en éducation. Dans son rapport, l'OCDE mentionne qu'aux États-Unis, plusieurs cas de tricherie ont été rapportés dans les médias. Ce qui a contribué au débat actuel sur la pertinence des tests standardisés utilisés comme un outil d'évaluation (Morris, 2011, p. 39). Parmi ces cas de tricherie, l'un des plus documentés s'est déroulé dans l'État de la Géorgie. Nathan Deal, le gouverneur républicain de cet État américain, a rendu publics, le 5 juillet 2011, les résultats d'une enquête sur des allégations de tricherie aux tests standardisés dans les écoles publiques d'Atlanta. Selon le communiqué de presse émis par le bureau du gouverneur, les résultats de l'enquête mènent à la conclusion que « la tricherie est généralisée dans les écoles publiques d'Atlanta<sup>23</sup> ». Certains résultats de l'enquête sont effectivement troublants<sup>24</sup> :

- Des cas de tricherie ont été découverts dans 44 des 56 écoles examinées ;
- Dans ces 56 écoles, 38 directeurs ont été reconnus responsables ou directement impliqués dans la tricherie ;
- En tout, l'enquête a trouvé que 178 enseignants et directeurs étaient responsables de tricherie. Et 82 d'entre eux ont reconnu les faits.

L'étude de l'OCDE démontre que la tricherie peut s'exécuter de différentes façons : en répondant aux questions non répondues par les élèves ; en changeant les mauvaises réponses pour des bonnes ; en accordant aux élèves plus de temps pour faire les tests ; en leur fournissant au préalable certaines réponses ou en leur permettant d'obtenir le test avant la date d'examen (Morris, 2011, p. 39). L'étude ajoute qu'aucun consensus n'existe sur les façons de limiter ce problème.

### *Enseignement orienté vers la pratique des tests et la réduction du curriculum*

L'enseignement orienté vers la pratique des tests (en anglais « *teach to the test* ») consiste à enseigner aux enfants des habiletés et des compétences spécifiques leur permettant de mieux performer aux tests. On peut, par exemple, pratiquer la mémorisation ou mettre l'accent sur les moyens efficaces et sur les trucs pour bien réussir les tests, qu'il s'agisse d'examens à

---

<sup>22</sup> Campbell (1976) est cité dans Forum on Educational Accountability (2011). Traduction libre de l'auteur.

<sup>23</sup> Gouverneur de Géorgie (2011). Traduction libre de l'auteur.

<sup>24</sup> Cette information est tirée du même communiqué de presse.

choix multiples ou d'essais. La réduction du curriculum consiste, elle, à carrément délaissier les matières qui ne sont pas sujettes à examen, comme l'histoire, la géographie ou les arts, afin de mettre l'accent sur les matières testées, souvent les mathématiques, la lecture et l'écriture. L'effet des tests sur la réduction du curriculum est bien documenté. Dans son analyse de 49 études sur la question, Wayne Au arrive à la conclusion suivante :

... les preuves présentées ici indiquent clairement que lorsque les enseignants font face à des tests standardisés aux fins de reddition de compte ou d'évaluation, ces tests entraînent une réduction du curriculum vers les matières sur lesquelles portent les tests<sup>25</sup> (Au, 2007, p. 264).

L'enseignement orienté vers la pratique soutenue des tests et la réduction du curriculum sont très liés et constituent le biais le plus important des résultats aux tests standardisés selon Adams et collab. (2009). À ce sujet, Shewbridge et collab. (2011, p. 120) mentionnent que :

La recherche aux États-Unis a montré que si les tests nationaux sont considérés par les enseignants et les écoles comme représentant un enjeu important, l'enseignement orienté vers la pratique des tests peut facilement provoquer une hausse des résultats et ainsi les rendre inefficaces comme mesure du progrès réel<sup>26</sup>.

### *Exclusion des élèves*

L'exclusion des élèves est une autre stratégie visant à améliorer les résultats aux tests standardisés. Souvent, les objectifs de résultats sont définis en termes de seuils à atteindre, par exemple, « que 75 % des élèves obtiennent la note de passage ». Confrontés à ces objectifs de rendement, les enseignants ou les écoles ont donc tendance à s'occuper plus attentivement des élèves dont les notes avoisinent celles du seuil de réussite, négligeant ainsi les plus faibles et les plus forts. Une autre technique permettant de hausser le seuil de réussite consiste à exclure certains élèves des examens. Le « Miracle du Texas » (le *Texas Miracle*) qui a fait les manchettes aux États-Unis est un exemple de ce type de manipulation. Alors que le district scolaire de Houston affichait un taux de décrochage fantastique de seulement 1,5 % (d'où le *Texas Miracle*), un reportage de l'émission *60 minutes* a démontré, à l'aide d'experts et d'éducateurs, que le taux de décrochage se situait en réalité entre 25 % et 50 % ! Booher-Jennings (2005) analyse en détail ce système de « triage » des élèves qui a été utilisé au Texas.

L'OCDE cite une synthèse de Figlio et Loeb (2011) sur la recherche en matière de reddition de comptes dans le système scolaire. Ces deux chercheurs concluent :

Les preuves sont très claires. Les écoles ont répondu à la pression des systèmes de reddition de compte en reclassant les élèves faibles en tant qu'élèves avec difficulté d'apprentissage<sup>27</sup> (Morris, 2011, p. 39).

---

<sup>25</sup> Traduction libre de l'auteur.

<sup>26</sup> Traduction libre de l'auteur. Passage cité dans Morris, 2011, p. 38.

<sup>27</sup> Traduction libre de l'auteur.

En résumé, toutes ces stratégies permettent de hausser artificiellement les résultats des élèves aux tests standardisés, ce qui les rend inutilisables comme indicateurs précis de l'apprentissage des enfants et encore moins de la performance des écoles ou des enseignants.

## **Conclusion**

L'idée de la CAQ d'utiliser des tests standardisés comme un mécanisme de reddition de comptes pour les écoles ou comme une méthode d'évaluation des enseignants n'est pas une nouveauté. Au 19<sup>e</sup> siècle, la Grande-Bretagne les a brièvement utilisés et a conclu à leur inefficacité.

Depuis le début des années 2000, sous l'impulsion des programmes fédéraux américains *No Child Left Behind* et *Race to the Top*, l'idée d'utiliser des tests standardisés a fait un retour en force. Cependant, l'efficacité de ces tests et de la rémunération au mérite pour améliorer les résultats scolaires des élèves est loin d'être démontrée, alors que leurs effets pervers, eux, sont de plus en plus documentés. Face à cette situation, les critiques et les remises en question se font de plus en plus soutenues et proviennent même de personnes, telle l'ancienne vice-secrétaire à l'éducation américaine Diane Ravitch, qui avaient plaidé en faveur de la mise en place de ces tests standardisés. Récemment, des centaines de directrices et directeurs d'école ont joint le mouvement de contestation en signant une pétition qui dénonce le système d'évaluation des directions d'école et des enseignants et l'utilisation des tests standardisés<sup>28</sup>.

En plus de tous ces éléments qui devraient alimenter la réflexion et inciter à la prudence, il ne faut pas oublier que la mise sur pied et l'utilisation de processus d'évaluation représentent des coûts considérables pour l'État. Ainsi, tout semble indiquer que l'utilisation de tests standardisés constitue une fausse bonne idée qui transformerait le système d'éducation québécois en un système dont la préoccupation première n'est plus l'apprentissage réel des enfants, mais bien la performance à ces fameux tests.

---

<sup>28</sup> La lettre est accessible à l'adresse Internet suivante : <http://www.newyorkprincipals.org/appr-paper> .  
Le *New York Times* a aussi publié un article au sujet de cette pétition (voir Winerip, 2011).

## Bibliographie

ADAMS, S. J., J. S. HEYWOOD et R. ROTHSTEIN (2009). *Teachers, Performance Pay, and Accountability: What Education Should Learn from Other Sectors*, Economic Policy Institute, Washington, 122 p.

[http://www.epi.org/publication/books-teachers\\_performance\\_pay\\_and\\_accountability/](http://www.epi.org/publication/books-teachers_performance_pay_and_accountability/)

AU, W. (2007). « High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis », *Educational Researcher*, vol. 36, n° 5, p. 258-267.

BOOHER-JENNINGS, J. (2005). « Below the Bubble: ‘Educational Triage’ and the Texas Accountability System », *American Educational Research Journal*, vol. 42, n° 2, p. 231-268.

CAMPBELL, Donald (1976). « Assessing the Impact of Planned Social Change, *The Public Affairs Center*, Dartmouth College.

<http://www.eric.ed.gov/PDFS/ED303512.pdf>.

CHAPTAL, et autres (2008). *Payer les profs au mérite ?* Institut de recherche de la FSU, Éditions Nouveaux regards, Paris, 102 p.

COALITION AVENIR QUEBEC (2011a). *Se doter de l'un des meilleurs systèmes d'éducation au monde d'ici 2020*, Texte de consultation en matière d'éducation, 12 avril 2011.

[http://archives.coalitionavenirquebec.org/files/Texte\\_consultation\\_education\\_final.pdf](http://archives.coalitionavenirquebec.org/files/Texte_consultation_education_final.pdf)

COALITION AVENIR QUÉBEC (2011b). *Agir pour l'avenir* (14 novembre).

[http://coalitionavenir.org/medias/Plan\\_daction\\_coalition.pdf](http://coalitionavenir.org/medias/Plan_daction_coalition.pdf)

DELISIO, Ellen R. (2011). « Pay for Performance: What are the issues? », *Education World*.

[http://www.educationworld.com/a\\_issues/issues/issues374a.shtml](http://www.educationworld.com/a_issues/issues/issues374a.shtml)

(consulté le 3 novembre).

FORUM ON EDUCATIONAL ACCOUNTABILITY (2011). *What Should Congress do about Teacher Evaluation?* (20 septembre)

[http://www.edaccountability.org/What\\_Should\\_Congress\\_Do\\_About\\_Teacher\\_Evaluation-\\_FEA\\_letter\\_9-20-11.pdf](http://www.edaccountability.org/What_Should_Congress_Do_About_Teacher_Evaluation-_FEA_letter_9-20-11.pdf)

FRYER, R. G. (2011). « Teacher Incentive and Student Achievement: Evidence from New York City Public Schools », *National Bureau of Economic Research*, Working Paper 16850 (mars).

GLAZERMAN, S. et A. SEIFULLAH (2010). *An Evaluation of the Teacher Advancement Program (TAP) in Chicago: Year Two Impact Report*, Mathematica Policy Research Inc. (17 mai).

GODMAN, S. et L. Turner (2010). « Teacher Incentive Pay and Educational Outcomes: Evidence from the NYC Bonus Program », *Program on Education Policy and Governance Working Papers Series*, PEPG 10-07, préparé pour la conférence *Merit Pay: Will it Work? Is it Politically Viable?*, Harvard Kennedy School (3-4 juin).

GOUVERNEUR DE GÉORGIE (2011). « Deal Releases Findings of Atlanta School Probe », Communiqué de presse ( 5 juillet).

[http://gov.georgia.gov/00/press/detail/0,2668,165937316\\_165937374\\_173112104,00.htm](http://gov.georgia.gov/00/press/detail/0,2668,165937316_165937374_173112104,00.htm)

JONES, M. (2011). *Teacher Behavior under Performance Pay Incentives*, Job Market Paper, Département d'économie, Université Notre Dame (octobre).

LEUNG, R. (2009). « The 'Texas miracle' », CBS News (11 février).

<http://www.cbsnews.com/stories/2004/01/06/60II/main591676.shtml>

MORRIS, A. (2011). « Student Standardised Testing : Current Practices in OECD Countries and a Literature Review », *OECD Education Working Papers*, n° 65, OECD Publishing.

[http://www.oecd-ilibrary.org/education/student-standardised-testing\\_5kg3rp9qbnr6-en](http://www.oecd-ilibrary.org/education/student-standardised-testing_5kg3rp9qbnr6-en)

NATIONAL CENTER ON PERFORMANCE INCENTIVES (2010). « NCPI Researchers Announce Results of POINT Experiment », Communiqué de presse (21 septembre).

<http://www.performanceincentives.org/news/detail.aspx?pageaction=ViewSinglePublic&LinkID=561&ModuleID=69&NEWSPID=1>

RAMIREZ, A. (2010). « Merit Pay Misfires », *The Effective Educator*, vol. 68, n° 4 (décembre 2010/janvier 2011) p. 55-58.

<http://www.ascd.org/publications/educational-leadership/dec10/vol68/num04/Merit-Pay-Misfires.aspx>

RAVITCH, D. (2010). *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*, Basic Books, New York, 283 p.

SPRINGER, M. G. et autres (2010). *Teacher Pay for Performance: Experimental Evidence from the Project on Incentives in Teaching*, National Center on Performance Incentives, Université Vanderbilt (21 septembre).

[http://www.performanceincentives.org/data/files/news/BooksNews/POINT\\_REPORT\\_9.21.10.pdf](http://www.performanceincentives.org/data/files/news/BooksNews/POINT_REPORT_9.21.10.pdf)

TIENKEN, C. H. (2011). « Pay for Performance Based on Standardized Test Scores: Twenty Questions », *AASA Journal of Scholarship and Practice*, vol. 8, n° 2 (summer), p. 3-12.  
[http://www.aasa.org/uploadedFiles/Publications/Journals/AASA\\_Journal\\_of\\_Scholarship\\_and\\_Practice/JSP\\_Summer\\_2011.FINAL.pdf](http://www.aasa.org/uploadedFiles/Publications/Journals/AASA_Journal_of_Scholarship_and_Practice/JSP_Summer_2011.FINAL.pdf)

WINERIP, M. (2011). « Principals Protest Role of Testing in Evaluations », *New York Times* (27 novembre).  
<http://www.nytimes.com/2011/11/28/education/principals-protest-increased-use-of-test-scores-to-evaluate-educators.html?pagewanted=all>