

Les politiques d'évaluation en éducation. Et après?

Rédacteur invité:

Dany LAVEAULT

Liminaire

- 1** Les politiques d'évaluation en éducation. Et après?
Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

- 15** La formule actuelle d'évaluation ministérielle au Nouveau-Brunswick. Pour qui? Pourquoi?
Jimmy BOURQUE, Université de Moncton, Nouveau-Brunswick, Canada
Mathieu LANG, Université de Moncton, Nouveau-Brunswick, Canada
Jean LABELLE, Université de Moncton, Nouveau-Brunswick, Canada

- 31** La politique québécoise d'évaluation des apprentissages et les pratiques évaluatives
Michel LAURIER, Université d'Ottawa, Ontario, Canada

- 50** La politique d'évaluation du rendement en Ontario: un alignement qui se précise dans la persévérance et la durée
Dany LAVEAULT, Université d'Ottawa, Ontario, Canada
Louise BOURGEOIS, Université d'Ottawa, Ontario, Canada

- 68** Une recherche documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public (1^{re} à 12^e année) des provinces de l'Ouest canadien
Jules ROCQUE, Université de Saint-Boniface, Manitoba, Canada

- 85** L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds? Les enseignements de l'histoire récente de l'école primaire genevoise
Lucie MOTTIER LOPEZ, Université de Genève, Suisse

- 102** Comparer les systèmes éducatifs francophones à travers le monde grâce au PISA: pas si simple!
Ariane BAYE, Université de Liège, Belgique
Marc DEMEUSE, Université de Mons, Belgique
Nathanaël FRIANT, Université de Mons, Belgique

**VOLUME XLII : 3 – NUMÉRO SPÉCIAL,
AUTOMNE 2014**

Revue scientifique virtuelle publiée par
l'Association canadienne d'éducation de
langue française dont la mission est la
suivante : « Par la réflexion et l'action de
son réseau pancanadien, l'ACELF exerce
son leadership en éducation pour
renforcer la vitalité des communautés
francophones ».

Directrice de la publication

Chantal Lainey, ACELF

Présidente du comité de rédaction

Lucie DeBlois,
Université Laval

Comité de rédaction

Sylvie Blain,
Université de Moncton
Lucie DeBlois,
Université Laval
Nadia Rousseau,
Université du Québec à Trois-Rivières
Paul Ruest,
Université de Saint-Boniface
Mariette Théberge,
Université d'Ottawa

Directeur général de l'ACELF

Richard Lacombe

Conception graphique et montage

Claude Baillargeon

Responsable du site Internet

Étienne Ferron-Forget

Diffusion Érudit

www.erudit.org

Les textes signés n'engagent que
la responsabilité de leurs auteurs
et auteurs, lesquels en assument
également la révision linguistique.
De plus, afin d'attester leur recevabilité,
au regard des exigences du milieu
universitaire, tous les textes sont
arbitrés, c'est-à-dire soumis à des pairs,
selon une procédure déjà convenue.

La revue *Éducation et francophonie* est
publiée deux fois l'an grâce à
l'appui financier du ministère du
Patrimoine canadien et du Conseil
de recherches en sciences humaines
du Canada.

acelf

ASSOCIATION
CANADIENNE
D'ÉDUCATION DE
LANGUE FRANÇAISE

265, rue de la Couronne, bureau 303
Québec (Québec) G1K 6E1
Téléphone : 418 681-4661
Télécopieur : 418 681-3389
Courriel : info@acelf.ca

Dépôt légal
Bibliothèque et Archives nationales
du Québec
Bibliothèque et Archives du Canada
ISSN 1916-8659 (En ligne)
ISSN 0849-1089 (Imprimé)

Liminaire

Les politiques d'évaluation en éducation. Et après?

Dany LAVEAULT

Université d'Ottawa, Ontario, Canada

Canada



Conseil de recherches en
sciences humaines du Canada

Social Sciences and Humanities
Research Council of Canada

acelf

ASSOCIATION CANADIENNE
D'ÉDUCATION DE LANGUE FRANÇAISE

VOLUME XLII : 3 – NUMÉRO SPÉCIAL,
AUTOMNE 2014

Revue scientifique virtuelle publiée par l'Association canadienne d'éducation de langue française dont la mission est la suivante : « Par la réflexion et l'action de son réseau pancanadien, l'ACELF exerce son leadership en éducation pour renforcer la vitalité des communautés francophones ».

Directrice de la publication

Chantal Lainey, ACELF

Présidente du comité de rédaction

Lucie DeBlois,
Université Laval

Comité de rédaction

Sylvie Blain,
Université de Moncton
Lucie DeBlois,
Université Laval
Nadia Rousseau,
Université du Québec à Trois-Rivières
Paul Ruest,
Université de Saint-Boniface
Mariette Thérberge,
Université d'Ottawa

Directeur général de l'ACELF

Richard Lacombe

Conception graphique et montage

Claude Baillargeon

Responsable du site Internet

Étienne Ferron-Forget

Diffusion Érudit

www.erudit.org

Les textes signés n'engagent que la responsabilité de leurs auteurs et auteurs, lesquels en assument également la révision linguistique. De plus, afin d'attester leur recevabilité, au regard des exigences du milieu universitaire, tous les textes sont arbitrés, c'est-à-dire soumis à des pairs, selon une procédure déjà convenue.

La revue *Éducation et francophonie* est publiée deux fois l'an grâce à l'appui financier du ministère du Patrimoine canadien et du Conseil de recherches en sciences humaines du Canada.

acelf

ASSOCIATION
CANADIENNE
D'ÉDUCATION DE
LANGUE FRANÇAISE

265, rue de la Couronne, bureau 303
Québec (Québec) G1K 6E1
Téléphone : 418 681-4661
Télécopieur : 418 681-3389
Courriel : info@acelf.ca

Dépôt légal
Bibliothèque et Archives nationales
du Québec
Bibliothèque et Archives du Canada
ISSN 1916-8659 (En ligne)
ISSN 0849-1089 (Imprimé)

Les politiques d'évaluation dans le domaine de l'éducation

Rédacteur invité :

Dany LAVEAULT

Liminaire

1 Les politiques d'évaluation en éducation. Et après?

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

15 La formule actuelle d'évaluation ministérielle au Nouveau-Brunswick. Pour qui? Pourquoi?

Jimmy BOURQUE, Université de Moncton, Nouveau-Brunswick, Canada

Mathieu LANG, Université de Moncton, Nouveau-Brunswick, Canada

Jean LABELLE, Université de Moncton, Nouveau-Brunswick, Canada

31 La politique québécoise d'évaluation des apprentissages et les pratiques évaluatives

Michel LAURIER, Université d'Ottawa, Ontario, Canada

50 La politique d'évaluation du rendement en Ontario : un alignement qui se précise dans la persévérance et la durée

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

Louise BOURGEOIS, Université d'Ottawa, Ontario, Canada

68 Une recherche documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public (1^{re} à 12^e année) des provinces de l'Ouest canadien

Jules ROCQUE, Université de Saint-Boniface, Manitoba, Canada

85 L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds? Les enseignements de l'histoire récente de l'école primaire genevoise

Lucie MOTTIER LOPEZ, Université de Genève, Suisse

102 Comparer les systèmes éducatifs francophones à travers le monde grâce au PISA : pas si simple!

Ariane BAYE, Université de Liège, Belgique

Marc DEMEUSE, Université de Mons, Belgique

Nathanaël FRIANT, Université de Mons, Belgique

Les politiques d'évaluation en éducation. Et après?

Dany LAVEAULT

Université d'Ottawa, Ontario, Canada

Buts et conséquences des politiques : des résultats pas toujours au rendez-vous

L'éducation occupe une place importante dans nos sociétés et ses coûts figurent, avec ceux de la santé, parmi les dépenses les plus importantes des États modernes. Il n'est donc pas étonnant qu'une organisation internationale telle que l'OCDE se soucie d'évaluer l'efficacité des systèmes éducatifs de ses pays membres, et plus particulièrement l'efficacité de leur cadre d'évaluation (OCDE, 2009, 2013). Sans évaluation, l'éducation est un navire sans pilote. Sans évaluation valide, les systèmes d'éducation sont à risque de se méprendre quant au degré d'atteinte de leurs objectifs ainsi que sur l'efficacité des moyens mis en œuvre (Laveault, 2009a).

Au cours des vingt dernières années, deux principaux phénomènes ont accru l'impact de l'évaluation dans nos systèmes d'éducation. Le premier est la participation d'un nombre toujours plus grand de pays aux enquêtes internationales sur le rendement des élèves (PISA, TIMSS, PIRLS). Le second est la prise en compte grandissante du rôle de l'évaluation non seulement pour informer sur le rendement des élèves, mais aussi pour former ces derniers. L'évaluation formative s'est vue ainsi introduite comme composante nécessaire de toute politique d'évaluation visant, d'une part, à certifier les apprentissages des élèves et, d'autre part, à soutenir les enseignants, les élèves et les autorités scolaires dans leurs prises de décision.

Dans le cas des enquêtes internationales, les résultats sont largement diffusés auprès du public et font l'objet d'études approfondies tant par les chercheurs que par les autorités scolaires. Pour plusieurs pays, ces études comparatives du rendement

des élèves ont eu pour effet de remettre en question l'efficacité de leur système d'éducation. De plus, à l'intérieur de chaque territoire scolaire (districts, cantons, États ou provinces), il s'est avéré que le rendement des écoles ou des instances locales pouvait varier considérablement, toutes choses étant égales par ailleurs. Cela a fait en sorte que de nombreuses organisations scolaires ont mis en place leur propre système d'évaluation à grande échelle afin de faire une surveillance rapprochée du rendement de leur population scolaire.

Dans ce nouveau contexte d'évaluations extérieures à la salle de classe, le rôle des enseignants en matière d'évaluation s'est trouvé transformé, et ce, pour deux principales raisons. La première est la reconnaissance par la communauté scientifique de la contribution majeure de l'évaluation pour soutenir l'apprentissage des élèves (Black et Wiliam, 1998; Wiliam et Thompson, 2008; Wiliam, 2011; Scallon, 2000). Quoique l'ampleur de l'effet rapporté puisse varier selon les études, cet effet n'en demeure pas moins considérable, d'où l'importance que l'évaluation formative a prise peu à peu dans les politiques d'évaluation. De pratique facultative et artisanale qu'elle était à l'origine, l'évaluation formative est devenue une composante essentielle de la planification de l'enseignement et d'une véritable politique d'évaluation. La seconde raison est probablement plus subtile. Comparativement aux évaluations externes et uniformes, les évaluations sommatives réalisées par les enseignants – celles qui ont pour objectif de dresser le bilan des apprentissages et éventuellement de les certifier au moyen d'un document officiel tel qu'un bulletin – sont trop différentes et leurs niveaux d'exigence trop variables pour renseigner sur l'efficacité des systèmes d'éducation. C'est ainsi que les évaluations externes ont fait en sorte de diminuer le rôle des enseignants en matière d'évaluation sommative, alors que les attentes à l'égard de leur rôle lié à l'évaluation formative devenaient plus élevées.

Ce réaménagement des rôles n'est pas terminé. Il se traduit par des tensions qui peuvent être saines, surtout si elles donnent lieu à des synergies nouvelles qui contribuent ultimement à améliorer l'apprentissage des élèves et à de meilleures prises de décision aux différents niveaux des systèmes éducatifs (Laveault, 2009b; Laveault, Dionne, Lafontaine, Tessaro et Allal, 2014). Ces tensions peuvent aussi être contre-productives. De fait, les changements apportés par les évaluations externes et l'évaluation formative ne sont pas toujours à la hauteur des espérances. Il y a cependant lieu d'être optimiste. Comme annoncé par Fullan (2001, 2009), nous en sommes aux balbutiements d'une « professionnalisation » des réformes éducatives. Les systèmes d'éducation n'ont jamais fait l'objet d'autant d'études, de recherches et, bien entendu, d'évaluations. Il est à souhaiter qu'en assurant un meilleur suivi des systèmes éducatifs, le besoin de réformes importantes sera remplacé par des aménagements graduels et continus, tels que la révision périodique des programmes d'études (Legendre, 2002).

L'évaluation des apprentissages: des concepts en évolution

Pour l'ensemble des raisons précédentes, la mobilisation des connaissances dans le domaine de l'évaluation scolaire est de plus en plus cruciale. Les changements intervenus récemment dans les systèmes d'éducation ont non seulement suivi l'évolution des concepts en évaluation, mais dans certains cas les ont précédés et même provoqués. Ces changements ont fait apparaître des besoins nouveaux en matière de formation initiale et continue des enseignants en évaluation afin de mieux refléter et actualiser dans la pratique les objectifs des politiques d'évaluation et, aussi, de tenir compte des changements dans le rôle des enseignants. Pour que les attentes élevées à l'égard des systèmes éducatifs se matérialisent, force est de reconnaître que non seulement les politiques d'évaluation, leurs fondements scientifiques et les outils d'évaluation doivent évoluer, mais aussi que les enseignants doivent être informés et formés en plus d'être partie prenante de ces changements.

Parmi les concepts en évaluation appelés à évoluer, figurent l'évaluation formative et la validité de conséquence (Messick, 1988). Ces deux concepts sont par ailleurs liés indirectement: à quoi peut servir l'évaluation si celle-ci ne se traduit pas par des effets positifs ou si l'ensemble des conséquences positives ne l'emporte pas sur les inconvénients? Comme le disait déjà Cardinet, «l'évaluation scolaire est formative par nature et sommative par nécessité».

Le formatif et le sommatif: des objectifs qui se rejoignent

Historiquement, la responsabilité de soutenir l'apprentissage a été confiée à l'évaluation formative, l'évaluation sommative étant limitée au bilan des apprentissages et à la certification des apprentissages. Tourneur (1985, p.7) mettait déjà en garde contre cette fausse simplification à outrance:

L'évaluation sommative se trouve réduite au jugement de valeur sur la personne de l'élève ou à la notation des étudiants en termes de pourcentage ou de classement, quand elle n'est pas jugée responsable des abus liés à l'exercice du «pouvoir magistral».

Malgré le fait que les buts assignés à l'évaluation formative et à l'évaluation sommative soient différents, rien n'empêche a priori l'information recueillie initialement pour effectuer le bilan des apprentissages d'être utilisée dans un but formatif. De la même manière, l'information recueillie dans le but de soutenir l'apprentissage peut aussi renseigner sur la progression de l'élève et contribuer à son bilan, particulièrement dans les zones grises de prise de décision (Lafortune et Allal, 2008). Un groupe d'experts internationaux, principalement anglo-saxons, s'est entendu sur la définition suivante de l'expression évaluation-soutien d'apprentissage (EsA) («*assessment for learning*»), fréquemment utilisée en remplacement d'évaluation formative:

L'évaluation-soutien d'apprentissage fait partie des pratiques quotidiennes des élèves et des enseignants qui, individuellement et en interaction, réfléchissent sur et réagissent à l'information provenant d'échanges,

démonstrations et observations afin de favoriser les apprentissages en cours (Allal et Laveault, 2009).

Cette définition globale, tout comme les distinctions apportées par Stiggins (2005), cherche à établir un meilleur équilibre entre le rôle de l'enseignant et celui de l'élève. De plus, dans cette définition, l'éventail des situations d'évaluation pouvant servir à soutenir l'apprentissage est considérablement élargi pour y inclure autant des processus formels, comme des tests, qu'informels, comme les entretiens avec les élèves et les observations anecdotiques.

Ce qui importe selon Good (2011), ce n'est pas tant le qualificatif de l'évaluation que l'usage qui sera éventuellement fait de l'information recueillie. Par exemple, un bilan des apprentissages réalisé dans le cadre de l'évaluation sommative peut être réinvesti pour alimenter un entretien de l'enseignant avec l'élève et ses parents, identifier des cibles d'apprentissage prioritaires et orienter les choix scolaires de l'élève.

Peu importe le contexte dans lequel l'information sur l'apprentissage de l'élève est recueillie, un principe fondamental devrait être celui d'en optimiser l'utilisation pour l'apprentissage de l'élève. Plusieurs auteurs (De Ketele, 2006; Harlen, 2005; Looney, 2011; Taras, 2005) sont d'avis que le formatif et le sommatif ne doivent pas être utilisés en vases clos, l'un informant l'autre. De plus, les deux sont complémentaires: le sommatif porte principalement sur des informations à « gros grains », soit des apprentissages globaux se produisant sur des cycles de longue durée, alors que le formatif porte sur des apprentissages à « grains fins » dont les cycles sont de courte durée, jusqu'à l'interaction maître-élève. Peu importe la nature, fine ou grossière, de l'information recueillie sur l'apprentissage des élèves, l'important est de s'assurer que celle-ci contribue à améliorer les décisions qui sont prises pour favoriser l'apprentissage des élèves. Nous pourrions résumer ce principe par le slogan : formatif et sommatif, même combat!

La validité de conséquence : plus de bien que de mal

L'optimisation de l'usage de l'information d'évaluation répond à des impératifs d'efficacité et d'économie. Déjà, beaucoup de temps est utilisé pour tester les élèves et recueillir de l'information sur leurs acquis. Mais qu'en est-il du temps pris pour traiter ces renseignements, les analyser et prendre les décisions nécessaires pour soutenir leur progression? On ne peut indéfiniment accumuler des renseignements sur l'élève ou sur les systèmes d'éducation sans se soucier de l'impact réel de ces évaluations sur l'apprentissage.

À cet égard, l'évaluation n'en serait pas à ses premières errances. De mauvais usages de l'information recueillie peuvent avoir des conséquences néfastes, telles que le rétrécissement du curriculum, une évaluation incomplète du rendement des enseignants et des écoles qui ne tiendrait compte que des résultats des élèves et l'introduction de palmarès de rendement des établissements scolaires qui ne tiendrait pas compte du contexte propre à chacun. Même l'évaluation à intention formative – dont le rôle attendu consiste pourtant à soutenir l'élève dans ses apprentissages – a pu connaître des dérives et donner lieu à des pratiques d'évaluation qui ne vont pas

dans le sens de tels objectifs. Enfin, ajoutons que ces errances vont bien au-delà des simples aspects techniques qui se limiteraient à l'utilisation des résultats, au développement de tests et à des analyses statistiques des résultats selon les «règles de l'art». Elles interpellent au plus haut point nos valeurs éducatives et nécessitent une réflexion en profondeur sur l'éthique de l'évaluation scolaire (Hadji, 1997, 2012).

Parallèlement à la diversification des modalités d'évaluation et au décloisonnement entre le formatif et le sommatif, les exigences en matière de validité de conséquence ont progressivement été élevées. Si un usage plus diversifié de l'information recueillie sur l'apprentissage des élèves peut donner lieu à une meilleure utilisation du temps accordé à l'évaluation, cela ne signifie cependant pas que toutes les modalités de collecte d'information, ce qui est décrit en anglais par «*assessment*» (Laveault, 2010; Laveault, 2013), se valent les unes les autres. Une planification à rebours s'avérera nécessaire pour tailler la situation d'évaluation sur mesure, surtout lorsque les renseignements nécessaires ne sont pas disponibles, ou encore pour aider à prendre une décision bien précise. Bref, si une variété de renseignements recueillis sur l'élève, tant à «gros grains» qu'à «petits grains», peuvent contribuer à soutenir l'élève dans ses apprentissages, tout renseignement n'est pas formatif, et il est parfois nécessaire de concevoir des tâches qui permettront d'aller chercher une information utile à des fins bien précises.

Somme toute, la question n'est pas de savoir s'il faut recueillir plus ou moins de renseignements sur les élèves et sur le fonctionnement des systèmes éducatifs. La question est plutôt de savoir comment en optimiser l'usage au service de l'apprentissage de l'élève afin de réduire les effets négatifs de l'évaluation et d'en maximiser les effets positifs. Ultimement, l'évaluation scolaire a sa raison d'être lorsque l'utilisation des renseignements sur l'apprentissage de l'élève se traduit par un accroissement des probabilités, tant chez l'enseignant que chez l'élève, de prendre les bonnes décisions et de faire les bons choix afin de réguler au mieux la progression de ce dernier.

« Teach to the test »

La mise en place des récentes réformes éducatives a été accompagnée dans plusieurs pays industrialisés par l'introduction d'examens uniformes censés assurer un suivi et orienter les prises de décision à tous les niveaux du système éducatif. Un tel changement n'a pas eu que des effets positifs. L'une de ses conséquences négatives a été de provoquer un «rétrécissement du curriculum», les enseignants consacrant plus de temps aux matières testées, telles que la lecture, l'écriture, les mathématiques, et délaissant les autres matières. Poussée à l'extrême, cette pratique néfaste conduit les enseignants à entraîner les élèves à passer des tests similaires aux tests officiels et à ne faire porter les apprentissages que sur des bribes de connaissances qui font partie régulièrement de ces examens.

Une telle façon de faire, connue en anglais sous l'expression «*teach to the test*», relève plus de mauvaises pratiques d'enseignement que de mauvaises évaluations. S'opposer à l'utilisation d'examens uniformes et standardisés sur la base de telles façons de faire ne repose pas sur des motifs valables. Au contraire, ce serait priver les enseignants, les écoles et les systèmes scolaires dans leur ensemble de renseignements

précieux sur la qualité de l'alignement des apprentissages des élèves sur les priorités des programmes d'études dans les matières de base.

Cela étant dit, le meilleur moyen de mettre fin à ce genre de dérives consiste à améliorer non seulement l'enseignement, mais aussi les outils d'évaluation et la formation des enseignants. L'utilisation à outrance de questions à choix multiples peut contribuer à réduire la crédibilité des épreuves standardisées, notamment pour tout ce qui concerne l'évaluation des apprentissages de haut niveau, et à en réduire l'impact sur l'apprentissage des élèves. Même s'il ne faut pas sous-estimer la capacité de questionnaires à choix multiples à tester de tels apprentissages avec validité, une évaluation fine de compétences complexes nécessite beaucoup de temps et de supervision et une diversité de moyens d'évaluation, ressources rarement disponibles pour les évaluations à grande échelle (Beckers, Hirtt et Laveault, 2014). C'est pourquoi des modalités « hybrides » d'évaluation, combinant ce que les examens uniformes font de mieux – recueillir une masse de renseignements sur tous les élèves en peu de temps et à faible coût – avec ce que les enseignants font de mieux – observer de façon précise les réalisations des élèves dans des situations d'apprentissage authentiques et complexes –, constituent sans doute le meilleur compromis.

Dans les faits, plusieurs organisations scolaires optent déjà pour ce genre de modalités d'évaluation à grande échelle, dont le dosage entre partie « externe » de l'évaluation et partie « interne » sous la responsabilité de l'enseignant peut varier. Il en existe plusieurs témoignages dans ce numéro spécial d'*Éducation et francophonie* et dans plusieurs autres références récentes (OECD, 2011, 2013; Laveault, 2009b; Laveault, Dionne, Lafontaine, Tessaro et Allal, 2014). Peu importe, cependant, la modalité choisie, la clef de voûte de l'implantation réussie d'une politique d'évaluation passe dans tous les cas par une compétence accrue des enseignants à évaluer, que ce soit pour utiliser les résultats des évaluations externes à bon escient ou encore pour participer à la collecte de données probantes sur les aspects les plus complexes de l'apprentissage des élèves.

Les composantes d'une politique d'évaluation efficace

L'une des conséquences indirectes des études sur les évaluations internationales du progrès des élèves, telles que PISA, TIMMS, PIRLS, a été de mettre à l'avant-plan le rôle déterminant des enseignants dans la réussite scolaire des élèves. Dans la plupart des études sur la question, l'effet enseignant, quoique difficile à isoler d'autres effets de système (établissement, classe, etc.), rendrait compte d'une part non négligeable de la variance dans les résultats d'apprentissage des élèves : près de 30 % selon Hattie (2003) ou encore de 10 à 20 % selon Bressoux (1994). Selon l'OECD (2005), l'amélioration de la performance des enseignants est probablement la politique éducative la plus susceptible de se traduire en gains d'apprentissage pour les élèves. Darling-Hammond (2008) a démontré que les pays qui se classent le mieux aux enquêtes internationales sur les progrès des élèves se distinguent par le niveau d'exigence et la qualité de la formation du corps professionnel enseignant.

Cependant, la capacité des enseignants à exploiter l'information recueillie sur les apprentissages des élèves semble constituer un véritable défi. L'étude de Heritage, Kim, Vendlinski et Herman (2009) montre que, si les enseignants peuvent être du même avis quant à l'identification des faiblesses d'un élève, ils le sont moins souvent lorsqu'il s'agit de décider quelle devrait être l'étape suivante pour assurer sa progression. D'après Herman, Osmundson et Silver (2010), les enseignants ont besoin de développer les compétences qui leur permettront non seulement de définir les besoins d'apprentissage des élèves, mais aussi d'y répondre adéquatement. Ces deux études soulèvent des questions importantes quant à ce qui peut être mis en place pour améliorer la capacité des enseignants à intervenir adéquatement.

Dans les deux études précédentes, les enseignants devaient se prononcer sur la base de résultats obtenus au moyen d'outils d'évaluation qu'ils n'avaient pas eux-mêmes conçus. Les résultats de ces études demandent à être vérifiés dans des contextes où les enseignants jouent un rôle important dans la conception des instruments qui serviront à évaluer les élèves. Selon Looney (2011), les enseignants sont bien conscients de l'importance pour eux d'être plus systématiques dans leur approche de l'évaluation en salle de classe, étant donné que les interactions les plus efficaces avec les élèves requièrent une planification soignée. En conséquence, l'un des meilleurs moyens pour accroître l'usage formatif de l'information recueillie sur l'apprentissage des élèves consiste à accroître l'implication des enseignants dans la conception et la planification de l'évaluation (Laveault, 2014). Selon Webb (2009), l'amélioration des compétences professionnelles des enseignants en évaluation n'est pas un problème de ressources. Elle consiste plutôt à développer chez l'enseignant le regard d'un concepteur (*designers' eye*) pour choisir, adapter et développer les tâches nécessaires pour vérifier la compréhension des élèves.

Dans plusieurs États américains, de bons résultats d'évaluation peuvent être associés à des primes salariales, avec les dérives que de telles primes peuvent entraîner. Selon Baker, Barton, Darling-Hammond, Haertel, Ladd, Linn, Ravitch, Rothstein, Shavelson et Shepard (2010), la prétention selon laquelle les enseignants vont améliorer l'apprentissage des élèves s'ils sont récompensés financièrement pour les gains des élèves n'est simplement pas justifiée par la recherche. Les conséquences indésirables de telles politiques sont nombreuses : appauvrissement du curriculum, désintérêt des enseignants pour les élèves les plus en besoin, diminution de la collaboration entre les enseignants et démoralisation.

Parmi les pays industrialisés, les réussites de la Finlande aux examens internationaux ont beaucoup attiré l'attention. Selon Darling-Hammond et McCloskey (2008), ce pays attribue sa réussite à des investissements massifs en formation des maîtres, la plupart d'entre eux possédant une maîtrise en éducation et dans leur matière de prédilection. En matière d'évaluation, la formation des enseignants finlandais met l'accent sur l'évaluation formative au service de l'apprentissage. Dans un tel contexte, il n'est pas étonnant de constater que les décideurs politiques accordent beaucoup plus d'autonomie aux enseignants quant à ce qu'ils doivent enseigner qu'on ne le fait ailleurs et que le programme d'études finlandais est un document beaucoup plus «léger» que celui rencontré dans d'autres pays.

L'État australien du Queensland est souvent mentionné pour illustrer les avantages d'une implication accrue des enseignants dans les systèmes d'évaluation nationaux (Atkin, 2007; Black et William, 1998; Darling-Hammon et McCloskey, 2008). Ce système a été développé et raffiné durant une période de trente ans et a nécessité à ses débuts des investissements importants dans la formation des enseignants. Il ne tient pas pour acquis que le jugement professionnel des enseignants repose uniquement sur la communication descendante (*top-down*) de standards nationaux. Dans ce système, la notion de «modération sociale» joue un rôle capital (Maxwell, 2007, cité dans Wyatt-Smith, Klenowski et Gunn, 2010). La modération tient pour acquis que l'évaluation des enseignants possède une validité réelle, mais que la fidélité ne peut en être assurée qu'en permettant aux enseignants de se faire une représentation commune des standards et des indicateurs nécessaires à la reconnaissance des résultats qui démontrent l'atteinte de ces standards (Maxwell, 2001, p. 6). De telles pratiques font en sorte que les standards visés par les politiques prennent un sens qui se développe et se raffine avec le temps à l'intérieur d'une communauté professionnelle d'utilisateurs (Wyatt-Smith, Klenowski et Gunn, 2010).

L'ensemble des études internationales montre bien les limites d'une approche qui considère l'enseignant comme un simple technicien de l'évaluation. Au contraire, lorsque les politiques d'évaluation misent sur le développement des compétences professionnelles des enseignants, celles-ci s'accompagnent de meilleurs résultats. Les compétences en évaluation ne sont pas qu'un outil puissant de régulation de la pratique des enseignants en salle de classe. Elles sont aussi pour l'enseignant une source importante de réflexion sur sa pratique et de perfectionnement continu. Bref, pour qu'une politique d'évaluation fasse l'objet d'une mise en œuvre réussie, il faut prévoir améliorer non seulement les évaluations, mais aussi les évaluateurs. C'est à ce titre, également, que les enseignants pourront agir comme modèles et développer chez leurs élèves les habiletés réflexives et métacognitives qui les rendront plus aptes à s'autoévaluer et à apprendre par eux-mêmes. L'indicateur ultime de la réussite de la performance d'un système éducatif ne réside pas uniquement dans les résultats d'apprentissage eux-mêmes, que l'on peut gonfler momentanément de manière artificielle, mais aussi et surtout dans la capacité des enseignants et de leurs élèves à apprendre par eux-mêmes.

Les avancées dans le domaine de la recherche en évaluation scolaire nous ont permis de prendre conscience plus que jamais des impacts, tant positifs que négatifs, que peuvent avoir certaines politiques sur l'apprentissage des élèves. Mieux informés des conséquences, de ce qui se produit après la conception et la mise en œuvre de nouvelles politiques d'évaluation, nous pouvons maintenant mieux anticiper la validité de conséquence de plusieurs de nos choix futurs. À moins de pratiquer un aveuglement volontaire, à la prochaine vague d'innovations dans le domaine de l'évaluation on ne pourra certainement pas plaider l'ignorance, du moins pas entièrement. Il existe un corpus de savoirs à mobiliser dans ce genre d'entreprise et l'élaboration de nouvelles politiques ne pourra faire l'économie de ces connaissances accumulées grâce à de nombreuses recherches en évaluation des apprentissages et en évaluation de programmes.

Et maintenant?

Quel est l'état du corpus de connaissances en matière de politiques d'évaluation dans les pays francophones et en particulier dans les provinces et territoires du Canada? À la suite de la publication en 2011 du numéro sur la réussite scolaire et pour répondre à de nombreuses demandes de son lectorat, la revue *Éducation et francophonie* a voulu rendre compte de l'état des connaissances sur les politiques d'évaluation en milieu scolaire en faisant paraître un numéro spécial sur le sujet. Elle entend ainsi faire la lumière sur les politiques actuelles d'évaluation afin de voir comment elles influencent les systèmes d'éducation et contribuent à améliorer les apprentissages des élèves. Les auteurs du numéro ont donc pour objectif non seulement de présenter les politiques d'évaluation, mais aussi d'engager une réflexion sur leur valeur éducative. En considérant la genèse, l'historique et le contexte particulier des politiques, ils ont cherché à bien en expliquer les raisons d'être et à répondre aux questions suivantes :

- Qui a voulu ces politiques et à quelles fins?
- Les résultats obtenus ont-ils transformé les systèmes d'éducation?
- Les évaluations ont-elles contribué à améliorer les niveaux d'apprentissage?

Ce numéro spécial se présente donc comme une réflexion critique sur l'évaluation afin d'alimenter le débat qui a cours sur le sujet dans le domaine de l'éducation. Son objectif est de traiter des différents rôles, formatif et sommatif, que peut jouer l'évaluation à tous les niveaux des systèmes d'éducation, de la salle de classe aux plus hautes instances locales et nationales, en passant par les établissements scolaires. Une attention particulière a été accordée à l'évaluation des apprentissages dans le cadre de la scolarité obligatoire, c'est-à-dire au primaire et au secondaire. Pour ce qui concerne la composante canadienne, la problématique de l'évaluation scolaire en milieu minoritaire et la façon dont celle-ci est prise en compte dans les provinces ont fait l'objet d'une attention particulière. Enfin, ce numéro cherche à familiariser les lecteurs de la revue *Éducation et francophonie* avec les concepts importants de l'évaluation en prenant en considération le fait que ceux-ci ont évolué depuis les dernières réformes en éducation et que, dès lors, une mise à jour s'imposait.

Parce que les politiques d'évaluation relèvent d'instances gouvernementales, les articles de ce numéro spécial se distinguent d'abord par leur origine géographique. Selon le cas, ils décrivent la situation actuelle dans les provinces de l'Est du Canada (Bourque, Lang et Labelle), au Québec (Laurier), en Ontario (Laveault et Bourgeois) et enfin dans les provinces de l'Ouest et les territoires (Rocque). Pour faire écho à la situation prévalant dans la francophonie internationale, un article décrivant la situation vécue en Suisse fait également partie de ce numéro (Mottier-Lopez). Enfin, avec la publication toute récente des résultats PISA, nous avons voulu comparer les résultats des pays francophones. Les auteurs Baye, Demeuse et Friant vous expliqueront en quoi « ce n'est pas si simple! ».

Et la question est effectivement complexe : autant d'instances gouvernementales, autant de politiques d'évaluation, mais aussi autant de leçons à tirer et d'occasions

de parfaire nos connaissances et d'enrichir notre expérience. Toutes ces politiques tirent plus ou moins leur origine des dernières réformes en éducation. Dans les provinces de l'Ouest du Canada et dans les provinces maritimes, le management axé sur les résultats de performance (*New Public Management*) a grandement influencé les politiques d'évaluation. Du côté de l'Atlantique, les chercheurs ont cherché des moyens pour que «les évaluations standardisées puissent être utilisées d'une manière optimale sans qu'elles deviennent un obstacle aux finalités plus humanistes et émancipatrices proposées par *Les enfants au premier plan*». Pour ces auteurs, une approche cohérente du management axé sur les résultats est possible à condition «de ne pas confondre l'évaluation du système éducatif et l'évaluation des apprentissages».

Au Canada central, l'Ontario et le Québec semblent avoir connu des expériences fort différentes quant aux modalités d'implantation de leurs politiques. En ce qui concerne le Québec, Laurier conclut que la politique québécoise, fortement influencée par un modèle pédagogique axé sur le socioconstructivisme, «s'est déployée avec l'appui de documents complémentaires et de modifications aux règlements qui en précisent et parfois en réduisent la portée». En Ontario, Laveault et Bourgeois décrivent une implantation qui s'est faite dans la continuité, notamment par la mise en place de nouvelles institutions, dont un ordre professionnel des enseignants, mais surtout l'Office pour la qualité et la responsabilité en éducation (OQRE), chargé de faire le suivi des apprentissages des élèves et de vérifier l'alignement de la formation des élèves avec les attentes des programmes d'études. Ils en concluent que «c'est la refonte en profondeur des programmes-cadres en Ontario qui a le plus influencé les pratiques d'évaluation des enseignants et a soutenu une mise en œuvre progressive de la politique d'évaluation actuelle».

Une comparaison rapide entre les provinces canadiennes permet d'observer une différence importante dans l'implantation des politiques d'évaluation entre l'Est et l'Ouest du Canada et les deux provinces du centre. Dans le premier cas, la mise en place des politiques semble motivée et dirigée par des considérations d'ordre administratif, alors que dans le deuxième cas ce sont des considérations pédagogiques et curriculaires qui priment. Dans un cas comme dans l'autre, un équilibre optimal est sans doute souhaitable, car les deux aspects sont indissociables dans la pratique, même si les discours des théoriciens et des praticiens peuvent mettre l'accent sur l'un ou l'autre aspect.

L'article portant sur la Suisse apporte un éclairage différent. Devant l'objectif d'harmonisation de la scolarité actuelle obligatoire poursuivi par la Confédération helvétique, Mottier-Lopez se pose la question suivante : «Comment les fonctions formative et certificative de l'évaluation des apprentissages des élèves ont-elles été traitées dans les réformes en Suisse romande ?» La méthode employée pour répondre à cette question est ingénieuse. En examinant les offres de formation aux enseignants genevois entre 1976 et 2013, Mottier-Lopez parvient à décrire une sorte de «cours des valeurs» que prennent celles-ci dans le perfectionnement continu des enseignants. Elle observe «des déséquilibres successifs entre les fonctions formative et certificative de l'évaluation, ainsi que des éléments d'incompréhension et de débats sociaux qui en ont résulté».

Les difficultés de mise en œuvre des politiques d'évaluation ont ceci en commun qu'elles font toutes ressortir l'importance de bien former les enseignants en matière d'évaluation. De toute évidence, nous n'y parvenons pas très bien, puisque, plusieurs décennies plus tard, nous devons encore définir en quoi consistent évaluation formative et sommative et qu'«il apparaît qu'un enjeu crucial est de penser un modèle d'une évaluation pédagogique intégrative et systémique parvenant à trouver un équilibre entre ses différentes fonctions qui, toutes, devraient tendre vers une même finalité de justice et de professionnalité» (Mottier-Lopez).

Un autre défi est particulièrement saillant dans les pays francophones, et en particulier là où l'enseignement est donné dans la langue de la minorité. La mondialité des échanges et la mobilité accrue des personnes font en sorte que les systèmes d'éducation doivent accueillir et intégrer de plus en plus de nouveaux arrivants. Baye, Demeuse et Friant démontrent que «certains systèmes semblent mieux que d'autres permettre aux élèves ne parlant pas la langue d'enseignement à la maison de la maîtriser en réduisant les écarts de performances entre groupes d'élèves». À cet égard, il existe de grandes différences entre les pays quant au risque qu'un élève immigré fasse partie du quartile des élèves les plus faibles que les natifs. Ce genre de données intéresse particulièrement les éducateurs et les éducatrices qui enseignent le français en milieu minoritaire, tant aux personnes natives qu'aux nouveaux arrivants.

Naturellement, il aurait été intéressant d'inclure plus de pays francophones dans ce numéro spécial, mais des contraintes d'espace et de temps nous en ont empêchés. Cependant, la coopération internationale en matière d'évaluation entre francophones est particulièrement active, et il est important de mentionner la collaboration qui existe entre l'Association pour le développement de la mesure et des méthodologies de l'évaluation au Canada (ADMÉE-Canada) et son association sœur, l'ADMÉE-Europe. Les deux ont en partage une revue internationale de langue française, *Mesure et évaluation en éducation*, qui constitue une référence incontournable pour tout ce qui concerne la recherche en matière d'évaluation en français. Cette revue a d'ailleurs publié un numéro spécial intitulé «L'évaluation : des politiques aux pratiques» (Laveault et Paquay, 2009). Ajoutons que le symposium de l'ADMÉE-Europe tenu au Luxembourg en 2012 a donné lieu à un chapitre de livre effectuant une mise à jour sur la situation qui prévaut au Canada et dans les pays francophones d'Europe (Laveault, Dionne, Lafontaine, Tessaro et Allal, 2014). Enfin, mentionnons que l'*International Encyclopedia of Education* a ajouté une nouvelle entrée sur la question de l'évaluation des politiques dans les pays francophones. L'article, rédigé en anglais, s'intitule «Classroom evaluation in a policy context: French sources» (Laveault, 2009).

Somme toute, nos connaissances en matière de politiques d'évaluation connaissent un essor considérable depuis une décennie. Ce numéro spécial d'*Éducation et francophonie* en est un autre témoignage. Il ne reste qu'à espérer qu'à l'avenir les leçons que nous tirons de l'expérience actuelle inspireront les prochaines générations d'éducateurs et qu'ils contribueront non seulement à améliorer l'évaluation des élèves, mais aussi à favoriser la réussite du plus grand nombre.

Références bibliographiques

- ALLAL, L. et LAVEAULT, D. (2009). Évaluation-soutien d'apprentissage. Prise de position formulée par la Troisième Conférence internationale sur l'évaluation-soutien d'apprentissage. *Mesure et évaluation en éducation*, 32(2), 99-107.
- ATKIN, J. M. (2007). Swimming upstream. Relying on teachers' summative assessments. *Measurement*, 5(1), 54-67.
- BAKER, E.L., BARTON, P. E., DARLING-HAMMOND, L., HAERTEL, E., LADD, H. F., LINN, R. L., RAVITCH, D., ROTHSTEIN, R., SHAVELSON, R. J. et SHEPARD, L.A. (2010). *Problems with the Use of Student Test Scores to Evaluate Teachers*. Economic Policy Institute: Briefing Paper no. 278. Récupéré le 20 septembre 2014 de <http://www.epi.org/publication/bp278/>
- BECKERS, J., HIRTT, N. et LAVEAULT, D. (2014). Les compétences à l'école : controverses et défis actuels. Dans C. Dierendonck, E. Loarer et B. Rey (dir.), *L'évaluation des compétences en milieu scolaire et en milieu professionnel* (p. 7-70). Bruxelles, De Boeck.
- BLACK, P. et WILLIAM, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-75.
- BRESSOUX, P. (1994). Les recherches sur les effets-écoles et les effets-maîtres. *Revue française de pédagogie*, 108, 91-137.
- DARLING-HAMMOND, L. et MCCLOSKEY, L. (2008). Assessment for learning around the world: What would it mean to be internationally competitive? *Phi Delta Kappan*, 90(4), 263-272.
- DE KETELE, J.-M. (2006). La recherche en évaluation : propos synthétiques et prospectifs. *Mesure et évaluation en éducation*, 29(1), 99-118.
- FULLAN, M. (2001). *The New Meaning of Educational Change* (3^e éd.). New York: Teachers College Press.
- FULLAN, M. (2009). Large-scale reform comes of age. *Journal of Educational Change*, 10, 101-113.
- Good, R. (2011). Formative « use » of assessment information: It's a process, so let's say what we mean. *Practical Assessment Research & Evaluation*, 16(3), 1-6.
- HADJI, C. (1997). Pour une éthique de l'agir évaluationnel. *Mesure et évaluation en éducation*, 20(2), 7-26.
- HADJI, C. (2012). *Faut-il avoir peur de l'évaluation?* Bruxelles, De Boeck.
- HARLEN, W. (2005). Teachers' summative practices and assessment for learning – Tensions and synergies. *The Curriculum Journal*, 16(2), 207-223.
- HATTIE, J. (2003). *Teachers Make a Difference. What Is the Research Evidence?* Australian Council for Educational Research, octobre 2003. Récupéré de http://research.acer.edu.au/research_conference_2003/4/

- HERITAGE, M., KIM, J., VENDLINSKI, T. et HERMAN, J. (2009). From evidence to action: A seamless process in formative assessment. *Educational Measurements. Issues and Practice*, 28(3), 24-31.
- HERMAN, J. L., OSMUNDSON, E. et SILVER, E. (2010). *Capturing Quality in Formative Assessment Practice. Measurement Challenges*. CRESST Report 770. Los Angeles: CRESST/University of California. Récupéré de <http://files.eric.ed.gov/fulltext/ED512648.pdf>
- LAFORTUNE, L. et ALLAL, L. (dir.) (2008). *Jugement professionnel en évaluation: pratiques enseignantes au Québec et à Genève*. Québec: Presses de l'Université de Québec.
- LAVEAULT, D. (2009a). L'amélioration de l'efficacité du système éducatif: sur quels indicateurs s'appuyer? Dans X. Dumay et V. Dupriez (dir.), *L'efficacité dans l'enseignement: promesses et zones d'ombre* (p. 177-194). Bruxelles: De Boeck.
- LAVEAULT, D. (2009b). L'évaluation en classe: des politiques aux pratiques. *Mesure et évaluation en éducation*, 32(3), 1-22.
- LAVEAULT, D. (2010). Classroom assessment in a policy context (French sources). Dans P. Peterson, E. Baker et B. McGaw (dir.), *International Encyclopedia of Education*
- LAVEAULT, D. (2013). Plus de convergences que de divergences: rencontre des conceptions francophones et anglo-saxonnes de l'évaluation des apprentissages. *Recherches et applications – Le français dans le monde*, 53, 31-41.
- LAVEAULT, D. (2014). The power of learning-centered task design: An exercise in the application of the variation principle. Dans C. Wyatt-Smith, V. Klenowski et P. Colbert (dir.), *Designing Assessment for Quality Learning* (p. 109-121), Berlin: Springer.
- LAVEAULT, D., DIONNE, E., LAFONTAINE, D., TESSARO, W. et ALLAL, L. (2014). L'évaluation des compétences: régulations entre politiques et pratiques. Dans C. DIERENDONCK, E. LOARER et B. REY (dir.), *L'évaluation des compétences en milieu scolaire et en milieu professionnel* (p. 85-96). Bruxelles: De Boeck.
- LEGENDRE, R. (2002). *Stop aux réformes scolaires*. Montréal: Guérin.
- LOONEY, J. W. (2011). *Integrating Formative and Summative Assessment*. OECD Education Working Papers, No. 58. Paris: OECD Publishing.
- MAXWELL, G. 2001. *Moderation of assessments in vocational education and training*. Brisbane, Queensland: Department of Employment and Training.
- MESSICK, S. (1988). The once and the future issues of validity: Assessing the meaning and consequences of measurement. Dans H. Wainer et H. I. Braun, *Test Validity*. Hillsdale, NJ: Lawrence Erlbaum.

- OCDE (2009). *Examen par l'OCDE des cadres d'évaluation en vue d'améliorer les résultats scolaires*. Paris : OCDE, Direction de l'éducation. Récupéré le 20 septembre 2014 de <http://www.oecd.org/dataoecd/17/4/44568061.pdf>
- OCDE (2013). *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. Résumé en français de OECD Reviews of Evaluation and Assessment in Education. Paris : OECD Publishing. Récupéré de http://www.oecd.org/edu/school/Synergies_for_Better_Learning_Résumé.pdf
- OECD (2005). *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*. Paris : OECD. Récupéré le 20 septembre 2014 de http://www.oecd.org/document/52/0,3746,en_2649_39263231_34991988_1_1_1_1,00.html
- OECD (2011). *Evaluation and Assessment Frameworks for Improving School Outcomes. Common Policy Challenges*. Paris : OECD. Récupéré le 20 septembre 2014 de <http://www.oecd.org/education/school/46927511.pdf>
- REEVES, D. B. (2006). Of hubs, bridges, and networks. *Educational Leadership*, 63(8), 32-37.
- SCALLON, G. (2000). *L'évaluation formative*, Saint-Laurent, Québec : Éditions du renouveau pédagogique.
- STIGGINS, R. J. (2005). *Assessment For Learning Defined*. An international forum on promoting sound assessment practice in every classroom. Portland, OR, Assessment Training Institute. Récupéré le 20 septembre 2014 de <http://ati.pearson.com/downloads/afldefined.pdf>
- TARAS, M. (2005). Assessment – Summative and formative – Some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466-478.
- TOURNEUR, Y. (1985). La certification des compétences. *Mesure et évaluation en éducation*, 7(4), 5-20.
- WEBB, D. C. (2009). Designing professional development for assessment. *Educational Designer*, 1(2). Récupéré le 20 septembre 2014 de www.educationaldesigner.org/ed/volume1/issue2/article6
- WILIAM, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37, 3-14.
- WILIAM, D. et M. THOMPSON (2008). Integrating assessment with learning : What will it take to make it work? Dans C. A. Dwyer (dir.), *The Future of Assessment. Shaping Teaching and Learning*. New York : Erlbaum.
- WYATT-SMITH, C. M., KLENOWSKI, V. et GUNN, S. (2010). The centrality of teachers' judgment practice in assessment. A study of standards in moderation. *Assessment in Education : Principles, Policy & Practice*, 17(1), 59-75.

La formule actuelle d'évaluation ministérielle au Nouveau-Brunswick Pour qui? Pourquoi?

Jimmy BOURQUE

Université de Moncton, Nouveau-Brunswick, Canada

Mathieu LANG

Université de Moncton, Nouveau-Brunswick, Canada

Jean LABELLE

Université de Moncton, Nouveau-Brunswick, Canada

VOLUME XLII : 3 – NUMÉRO SPÉCIAL,
AUTOMNE 2014

Revue scientifique virtuelle publiée par
l'Association canadienne d'éducation de
langue française dont la mission est la
suivante : « Par la réflexion et l'action de
son réseau pancanadien, l'ACELF exerce
son leadership en éducation pour
renforcer la vitalité des communautés
francophones ».

Directrice de la publication

Chantal Lainey, ACELF

Présidente du comité de rédaction

Lucie DeBlois,
Université Laval

Comité de rédaction

Sylvie Blain,
Université de Moncton
Lucie DeBlois,
Université Laval
Nadia Rousseau,
Université du Québec à Trois-Rivières
Paul Ruest,
Université de Saint-Boniface
Mariette Thérberge,
Université d'Ottawa

Directeur général de l'ACELF

Richard Lacombe

Conception graphique et montage

Claude Baillargeon

Responsable du site Internet

Étienne Ferron-Forget

Diffusion Érudit

www.erudit.org

Les textes signés n'engagent que
la responsabilité de leurs auteurs
et auteurs, lesquels en assument
également la révision linguistique.
De plus, afin d'attester leur recevabilité,
au regard des exigences du milieu
universitaire, tous les textes sont
arbitrés, c'est-à-dire soumis à des pairs,
selon une procédure déjà convenue.

La revue *Éducation et francophonie* est
publiée deux fois l'an grâce à
l'appui financier du ministère du
Patrimoine canadien et du Conseil
de recherches en sciences humaines
du Canada.

acelf

ASSOCIATION
CANADIENNE
D'ÉDUCATION DE
LANGUE FRANÇAISE

265, rue de la Couronne, bureau 303
Québec (Québec) G1K 6E1
Téléphone : 418 681-4661
Télécopieur : 418 681-3389
Courriel : info@acelf.ca

Dépôt légal
Bibliothèque et Archives nationales
du Québec
Bibliothèque et Archives du Canada
ISSN 1916-8659 (En ligne)
ISSN 0849-1089 (Imprimé)

Les politiques d'évaluation dans le domaine de l'éducation

Rédacteur invité :

Dany LAVEAULT

Liminaire

1 Les politiques d'évaluation en éducation. Et après?

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

15 La formule actuelle d'évaluation ministérielle au Nouveau-Brunswick. Pour qui? Pourquoi?

Jimmy BOURQUE, Université de Moncton, Nouveau-Brunswick, Canada

Mathieu LANG, Université de Moncton, Nouveau-Brunswick, Canada

Jean LABELLE, Université de Moncton, Nouveau-Brunswick, Canada

31 La politique québécoise d'évaluation des apprentissages et les pratiques évaluatives

Michel LAURIER, Université d'Ottawa, Ontario, Canada

50 La politique d'évaluation du rendement en Ontario : un alignement qui se précise dans la persévérance et la durée

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

Louise BOURGEOIS, Université d'Ottawa, Ontario, Canada

68 Une recherche documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public (1^{re} à 12^e année) des provinces de l'Ouest canadien

Jules ROCQUE, Université de Saint-Boniface, Manitoba, Canada

85 L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds? Les enseignements de l'histoire récente de l'école primaire genevoise

Lucie MOTTIER LOPEZ, Université de Genève, Suisse

102 Comparer les systèmes éducatifs francophones à travers le monde grâce au PISA : pas si simple!

Ariane BAYE, Université de Liège, Belgique

Marc DEMEUSE, Université de Mons, Belgique

Nathanaël FRIANT, Université de Mons, Belgique

La formule actuelle d'évaluation ministérielle au Nouveau-Brunswick Pour qui? Pourquoi?

Jimmy BOURQUE

Université de Moncton, Nouveau-Brunswick, Canada

Mathieu LANG

Université de Moncton, Nouveau-Brunswick, Canada

Jean LABELLE

Université de Moncton, Nouveau-Brunswick, Canada

RÉSUMÉ

Cet article comporte une analyse critique de la composante évaluative du plan éducatif *Les enfants au premier plan* lancé au Nouveau-Brunswick. Il soutient la thèse selon laquelle le plan du ministère de l'Éducation présente une approche cohérente avec le management axé sur les résultats de performance et qu'il s'agit, dans ce contexte, de ne pas confondre l'évaluation du système éducatif et l'évaluation des apprentissages. Pour en arriver à cette conclusion, nous analysons les conséquences de l'évaluation standardisée en lien avec les théories managériales axées sur la performance. Par la suite, à l'aide du concept de l'outil convivial développé par Illich, nous montrons comment il serait possible que les évaluations standardisées soient utilisées d'une manière optimale sans qu'elles deviennent un obstacle aux finalités plus humanistes et émancipatrices proposées par *Les enfants au premier plan*.

ABSTRACT

The current ministerial assessment formula in New Brunswick For whom? Why?

Jimmy BOURQUE

University of Moncton, New Brunswick, Canada

Mathieu LANG

University of Moncton, New Brunswick, Canada

Jean LABELLE

University of Moncton, New Brunswick, Canada

This article provides a critical analysis of the evaluative component of the education plan *When Kids Come First*, launched in New Brunswick. It supports the argument that the Ministry of Education's approach is coherent with management focused on performance results, meaning in this context that it does not confuse assessment of the education system and the assessment of learning. To reach this conclusion, we analyze the consequences of standardized evaluation in relation to performance-based management theories. Then, using the concept of the user-friendly tool developed by Illich, we show how standardized assessments could be used optimally without becoming an obstacle to the more humanistic and emancipatory aims proposed by *When Kids Come First*.

RESUMEN

La actual fórmula de evaluación ministerial en Nuevo Brunswick ¿Para quién? ¿Por qué?

Jimmy BOURQUE

Universidad de Moncton, Nueva Brunswick, Canadá

Mathieu LANG

Universidad de Moncton, Nueva Brunswick, Canadá

Jean LABELLE

Universidad de Moncton, Nueva Brunswick, Canadá

Este artículo conlleva un análisis crítico del componente evaluativo del plan educativo *Les enfants au premier plan* lanzado en Nueva Brunswick. Avanza la tesis según la cual el plan del ministerio de Educación presenta una óptica coherente con la gestión centrada en cumplimiento de resultados, el cual no debe confundirse, en

este contexto, con la evaluación del sistema educativo y la evaluación de los aprendizajes. Para concluir, analizamos las consecuencias de la evaluación estandarizada en relación con las teorías de la gestión centradas en los resultados. Después, apoyándonos en el concepto de la herramienta convival desarrollado por Illich, mostramos cómo es posible utilizar las evaluaciones de manera óptima sin que se conviertan en un obstáculo a las finalidades más humanistas y emancipadoras propuestas por *Les enfants au premier plan*.

Introduction

L'acte de pensée qui conduit à poser un jugement en dehors d'une perception immédiate exige qu'on s'appuie sur des faits et qu'on accumule des preuves à l'aide d'une démarche rigoureuse. En ce sens, l'évaluation trouve toute sa légitimité, car ce processus d'appréciation et de mesure permet de porter un jugement sur la valeur d'un sujet ou d'un objet à partir de l'observation attentive et systématique, de même qu'en fonction de normes et de critères dûment établis. Cependant, force est de constater que l'évaluation et la mesure n'ont pas toujours bonne presse (Bernard, 2009). Certains reprochent à l'évaluation d'être souvent biaisée : c'est que l'évaluation s'inscrit dans un paradigme et est souvent tributaire d'un système de valeurs dominant qui oriente l'observation et préside à l'établissement de normes et critères (Hadji, 2012; Jeffrey, 2013). D'autres soulignent que l'évaluation est presque toujours axée sur les résultats et qu'elle tient rarement compte des conditions de l'obtention de ceux-ci, de même que des processus nécessaires pour y parvenir (De Ketele, 2013; Dodge, 2009).

Mais d'où proviennent ces dérives? Dans ce texte, nous montrons comment les sciences de l'administration, particulièrement la gestion axée sur la performance, ont exercé une certaine influence sur les sciences de l'éducation en matière d'évaluation au Nouveau-Brunswick. D'abord, nous situons le débat à partir de l'adoption par le ministère de l'Éducation du Nouveau-Brunswick (2006) de l'énoncé de politique *Les enfants au premier plan* et proposons un retour critique sur le texte de Cormier (2011), qui dénonce l'évaluation strictement axée sur les résultats. Puis nous examinons de plus près le concept de réussite scolaire comme construction sociale. Nous précisons ensuite quelques conséquences de la conception couramment adoptée qui sous-tend l'idée des évaluations standardisées de la réussite scolaire. Nous poursuivons en traçant le lien entre le management axé sur la performance et l'utilisation par le ministère de l'Éducation et du Développement de la petite enfance du Nouveau-Brunswick des évaluations standardisées et contrastons cette utilisation avec la posture humaniste aussi promue par le même ministère et par plusieurs acteurs du système éducatif. Nous concluons en situant plus précisément le rôle des évaluations standardisées dans le système éducatif néo-brunswickois et en spécifiant comment ce rôle peut être rendu cohérent avec les visées du plan stratégique du Ministère.

Les enfants au premier plan... et ses mécontents

Le débat commence avec la parution, en 2006, d'un plan intitulé *Les enfants au premier plan* (Ministère de l'Éducation du Nouveau-Brunswick, 2006). En y indiquant des cibles, en implantant des mécanismes de reddition de comptes et en multipliant le nombre d'évaluations ministérielles des élèves pour guider ses actions et mesurer le progrès du système éducatif, le Ministère affirme implicitement sa posture axée sur les résultats. Les 23 éléments de la fiche de rendement équilibré de son plan en font foi. Mais c'est le message du premier ministre de l'époque qui est le plus éloquent en ce sens : selon ses dires, son gouvernement répond à la demande d'établir un «système axé sur les résultats» en améliorant l'évaluation des apprentissages et en allouant les ressources pour tenir compte des résultats (Ministère de l'Éducation du Nouveau-Brunswick, 2006). Au paragraphe suivant, le premier ministre insiste sur l'idée que la province doit aller plus loin, puisqu'elle aura besoin de citoyens créatifs, ingénieux et faisant preuve d'esprit critique. Il revient alors sur des aspects liés au concept de performance et à des compétences liées au travail pour répondre à ces besoins.

Dès lors, les mesures de reddition de comptes imposées par le Ministère, notamment l'augmentation du nombre d'évaluations ministérielles standardisées, attirent particulièrement l'attention. Les réactions à la multiplication du nombre d'évaluations ministérielles annoncée dans *Les enfants au premier plan* (Ministère de l'Éducation du Nouveau-Brunswick, 2006) ne se sont pas fait attendre. Une revue de presse montre que sur les 37 articles publiés entre le lancement du plan éducatif et janvier 2014, 30 portent sur la présentation du plan et des différentes annonces s'y rapportant alors que sept y accordent un traitement plus critique. De ces sept articles, aucun ne remet en question la place accordée à l'évaluation. Plutôt, on y critique les actions subséquentes des gouvernements Graham puis Alward, jugées incohérentes avec les objectifs qu'ils se sont fixés, notamment en réduisant le financement en enseignement ressource. En plus de textes publiés dans la colonne d'opinion du lecteur des quotidiens locaux, Bourque (2009) consacre un éditorial de l'*Info-CRDE* à une remise en perspective de la tendance amorcée avec *Les enfants au premier plan*, alors que Cormier (2011) publie une critique de sa composante évaluative dans les pages de cette revue.

Retour sur l'analyse de Cormier (2011)

Il est intéressant de revenir sur l'article de Cormier, publié en 2011, parce que sa position semble avoir trouvé un écho chez plusieurs enseignants et parents, surtout francophones, de la province. Ce texte avance principalement trois arguments :

- a) Les évaluations ministérielles contribuent au «complexe minoritaire» et à l'insécurité linguistique, notamment en maintenant une relation oppressive envers la minorité francophone et en réitérant le constat de sa piètre performance.
- b) Les évaluations ministérielles affectent négativement l'enseignement (qui devient plus «transmissif et magistral») et l'apprentissage (qui s'oriente plus vers la performance que vers la maîtrise).

- c) Les évaluations ministérielles, à travers le portrait peu flatteur qu'elles renvoient, encouragent la contre-affirmation et la dévalorisation de l'éducation, alors qu'élèves, enseignants et parents s'efforcent de préserver leur estime de soi.

Si nous reconnaissons le bien-fondé d'une partie du raisonnement de Cormier, surtout à titre d'hypothèse explicative, nous en identifions également certaines limites. La première limite a trait à l'application des théories culturelles de l'éducation, surtout basées sur des situations où l'éducation d'une minorité est la responsabilité d'un groupe majoritaire. Ainsi, Cormier (2011) écrit :

Quand un groupe se retrouve dans une position subordonnée et endure une coercition de la part de la majorité, ce groupe peut internaliser le discours oppressif et critique de la majorité et peut alors dévaloriser ses propres écoles (p. 15).

Cette affirmation convoque les travaux théoriques d'Ogbu (voir par exemple Ogbu et Simons, 1998). L'idée qui les sous-tend part de la notion « d'analphabétisme de minorité », par oppression ou par résistance. La question est ici celle de l'adéquation du cadre invoqué à la situation critiquée, dans la mesure où les travaux théoriques d'Ogbu ont été développés en fonction de situations où un endogroupe minoritaire « hiérarchique » ou subordonné se voit imposer une politique oppressive par un exogroupe majoritaire dont le statut est socio-politiquement dominant par rapport à l'endogroupe (Ogbu et Simons, 1998). Des cas d'espèce explorés par Ogbu sont les Latino-Américains en Californie, les Autochtones du Canada et des États-Unis, ou alors les systèmes scolaires coloniaux africains (et leurs prolongements actuels). Comme la communauté francophone du Nouveau-Brunswick contrôle ses établissements scolaires et est scolarisée en français par des francophones venant majoritairement de l'Acadie du Nouveau-Brunswick, il serait étonnant d'y observer les mêmes phénomènes d'oppression et de résistance que chez les Maoris, par exemple. Autrement dit, sommes-nous ici dans la même situation que celle décrite par Ogbu et Simons (1998)? Ce n'est pas impossible, mais pas démontré non plus, aucune recherche empirique ne s'étant arrêtée à cette question, du moins à notre connaissance.

La deuxième limite touche le jugement implicite dans le texte de Cormier quant à la valeur d'approches pédagogiques plus directives lorsqu'elle déplore que « ... dans la majorité des cas, l'enseignement devient surtout transmissif et magistral » (p. 16). Notons d'emblée que la nature exacte de cet enseignement n'est pas définie : il devient donc difficile d'évaluer exactement en quoi consistent les méthodes pédagogiques ainsi désignées. Sans savoir ce qui se passe réellement en salle de classe, il nous semble hasardeux de souscrire à ce jugement. D'autant plus que les approches surtout centrées sur l'enseignant sont, en fait, loin d'être inefficaces. Certaines méthodes, que l'on pourrait qualifier de transmissives, l'enseignement explicite, par exemple, s'avèrent en fait plus efficaces que la plupart des approches que l'on qualifie de socioconstructivistes (Hattie, 2009). Il n'est donc pas clair qu'un enseignement plus transmissif soit négatif en soi, comme il n'est pas établi que cette

transition ait suivi inéluctablement la hausse du nombre d'évaluations ministérielles au Nouveau-Brunswick, encore une fois faute de recherches s'étant penchées sur la question.

La troisième limite touche la question de la dévalorisation de l'éducation sub-séquente aux constats d'échec qui se produit lorsque l'élève et ses parents se voient à travers le regard perçu comme méprisant de la majorité. Ce que nous désirons explorer davantage ici, c'est l'idée que l'augmentation du nombre d'évaluations ministérielles et la diffusion accrue des résultats entraîneraient une dévalorisation de l'éducation. Notons qu'il existe peu ou pas de recherches empiriques sur le sujet. Notre question est la suivante: considérant les résultats obtenus aux tests PISA et PIRS avant la publication des *Enfants au premier plan*, considérant le niveau d'alphabétisme au Nouveau-Brunswick francophone commenté par Corbeil en 2006, pourrait-on conclure que la contre-affirmation et la dévalorisation de l'éducation sont déjà bien installées? Ces situations prévalant avant 2006, il est certainement difficile de les attribuer à des changements apportés plus tard.

Une dernière limite à soulever est celle de l'estime de soi, menacée, selon Cormier, par les évaluations ministérielles et la diffusion des résultats. Il est d'abord important de noter que des recherches menées depuis les années 2000 suggèrent que le lien entre l'estime de soi et le rendement scolaire serait au mieux faible et inconsistent (Marsh et O'Mara, 2008) et que le renforcement inconditionnel de l'estime de soi pourrait même, à long terme, se traduire par des coûts (en matière d'autonomie, de compétence, d'apprentissage, etc.) qui dépasseraient les gains (Crocker et Park, 2004). Par ailleurs, l'effet de la performance scolaire sur l'estime de soi serait relativement faible en comparaison, par exemple, avec l'effet de la performance sur la perception de contrôle (Ross et Broh, 2000). La crainte de voir la multiplication des évaluations ministérielles anéantir l'estime de soi des élèves et, par conséquent, produire d'autres effets néfastes à long terme nous semble donc peu étayée.

Il demeure qu'une bonne part de l'analyse de Cormier semble pertinente (particulièrement le fait que si l'évaluation provinciale peut prétendre aider à diagnostiquer les difficultés des élèves, elle ne contribue en rien à le résoudre) et qu'une des finalités des mesures gouvernementales en matière d'évaluation consiste à mettre l'accent sur l'évaluation de la performance. De notre point de vue, l'augmentation de la fréquence des évaluations demeure inquiétante dans une certaine mesure en ce qu'elle semble avoir pour origine une confusion entre l'évaluation des apprentissages des élèves et l'évaluation du système éducatif. Par conséquent, il semble que cette initiative instrumentalise l'élève en se servant de ses résultats pour évaluer la performance d'un système éducatif. Si tel était le cas, les mauvais résultats des élèves aux évaluations signifieraient plutôt une mauvaise performance du système. Dans le texte qui suit, nous défendons la thèse suivante: la décision d'augmenter le nombre d'évaluations est cohérente avec une approche axée sur des résultats de performance et non avec une approche d'amélioration des conditions d'apprentissage.

La réussite scolaire comme construction sociale historiquement située

L'une des particularités du document *Les enfants au premier plan* est que si le mot «réussite» y est mentionné 23 fois en 32 pages, ce terme n'est défini nulle part. En fait, la Fiche de rendement équilibré (p. 29 et 30) donne probablement les meilleurs indices de la conception de la réussite scolaire véhiculée par ce document. La fiche présente 23 cibles, qui incluent des objectifs aussi divers que la préparation à l'entrée à l'école, l'engagement des parents, la performance aux épreuves nationales et internationales en lecture, mathématique et sciences, l'engagement du secteur privé, le recul de l'intimidation et l'intégration des technologies en classe. On y fait également référence au «rendement scolaire», toujours sans le définir. Le caractère polysémique donné à la réussite scolaire par le Ministère est ici évident et suggère que l'augmentation de la fréquence des évaluations ministérielles ne constitue pas une fin en soi. Elle est plutôt une mesure de l'atteinte de certaines des cibles énoncées et vise davantage la reddition de comptes et la gestion axée sur la performance du système que la stimulation des apprentissages.

Il demeure que cette absence de définition de la réussite scolaire s'avère problématique. En raison de l'équivoque créée, il devient difficile de saisir clairement ce que le Ministère entend par réussite scolaire. Par conséquent, il devient aussi difficile de s'entendre sur les moyens à mettre en place pour obtenir cette réussite scolaire. D'ailleurs, même si le Ministère avait proposé une définition claire de la réussite scolaire et levé l'équivoque sémantique, il n'en découle pas que tous auraient adhéré à cette définition. En ce sens, selon Hacking (1999), la réussite scolaire est une construction sociale. Comme le suggère implicitement l'absence de définition dans *Les enfants au premier plan*, la nature de la réussite scolaire est tenue pour acquise et inévitable. Or, la réussite scolaire n'est pas une évidence : dans certaines cultures, elle pourrait ne jamais exister comme concept ou alors prendre une forme totalement différente. En fait, la conception répandue de la réussite scolaire comme performance à un ensemble d'épreuves sommatives jusqu'à l'obtention d'un diplôme est critiquée, notamment par Freire (1974) et Illich (2003). Ces auteurs avancent même que, si cette conception de la réussite scolaire était radicalement transformée, voire éliminée, la société ne s'en porterait que mieux.

Pour notre part, nous posons que la définition de la réussite scolaire¹, comme performance à des épreuves sommatives, ne constitue en fait qu'une facette de la définition implicitement suggérée par *Les enfants au premier plan* et que de limiter la portée de ce document à cette facette et aux mesures qui en découlent ne lui rend pas justice. D'ailleurs, nonobstant les débats sur la définition de la réussite scolaire, il demeure que l'un des buts du parcours scolaire est la qualification dans une visée vocationnelle, au point où, au Québec par exemple, on parle d'approche orientante

1. Selon Baby (2013), si la réussite éducative est atteinte lorsque les trois éléments de la mission de l'école que sont l'instruction, la socialisation et la qualification le sont, la réussite scolaire découle de l'«atteinte des objectifs de la seule mission d'instruire» (p. 254).

(Ministère de l'Éducation du Québec, 2002). *Les enfants au premier plan* ne fait pas exception : le troisième but fondamental énoncé est que, « Avant la fin de leurs études secondaires, nos enfants auront eu l'occasion de découvrir leurs forces ainsi qu'un domaine qui les passionne » (p. 9). Les critiques qui reprochent au système scolaire l'adoption d'une orientation capitaliste néolibérale axée sur l'employabilité doivent cependant admettre 1) qu'il s'agit de la matrice socioéconomique prévalant pour le moment; 2) que la dimension « professionnalisante » suppose un contrôle de la qualité de la formation ainsi qu'une évaluation sommative et certificative des apprentissages (ce qui ne pose pas nécessairement de balises quant à la fréquence des évaluations); 3) que cette dimension n'est pas la seule qui sous-tend la formation scolaire au Nouveau-Brunswick : *Les enfants au premier plan* traite aussi d'inclusion scolaire, de citoyenneté et d'identité culturelle.

S'il y a un certain consensus autour du principe de l'évaluation sommative, la critique formulée à l'égard des *Enfants au premier plan* porte surtout sur l'augmentation considérable de la fréquence des évaluations. Nous ajoutons à cela la faible probabilité que cette mesure contribue à l'atteinte des cibles du Ministère. Hattie (2009) résume bien la principale lacune de cette mesure : « *Another form of feedback is repeated testing, but this is only effective if there is feedback from the tests to teachers such that they modify their instruction to attend to the strengths and gaps in student performance* » (p. 178). Plus précisément, l'effet de l'évaluation sur l'apprentissage double en présence d'une rétroaction comparativement à l'effet de l'évaluation seule (Hattie, 2009). Ainsi, le problème soulevé par la multiplication des épreuves standardisées au Nouveau-Brunswick est que le personnel enseignant n'en reçoit pas une rétroaction suffisamment détaillée pour pouvoir adapter son enseignement en conséquence et ainsi optimiser les apprentissages des élèves. Toujours selon Hattie, à mesure que la fréquence des évaluations augmente, cela produit des effets positifs décroissants. Par ailleurs, l'augmentation des résultats est plus marquée lorsque les élèves sont soumis à plusieurs tests courts plutôt qu'à des épreuves moins fréquentes, mais plus longues, toujours en supposant l'existence d'une rétroaction appropriée (Bangert-Drowns, Kulik, Kulik et Morgan, 1991).

Ainsi, s'il y a lieu de repenser la place accordée à l'évaluation standardisée dans le parcours scolaire, il semble difficile de l'exclure complètement. Il faut reconnaître que la place accordée aux évaluations et aux résultats ne l'est pas *de facto* aux dépens des enfants. Un équilibre est possible entre le contrôle de la performance du système et l'amélioration des apprentissages des élèves. En ce sens, si les engagements pris dans *Les enfants au premier plan* dépassent de loin la multiplication des évaluations standardisées, il n'en demeure pas moins que le lien entre les moyens (les évaluations standardisées) et les fins (de meilleurs résultats des élèves) est plutôt ténu.

Le curriculum caché : reproduction et sélection

Ayant jusqu'ici nuancé certains propos entourant l'évaluation standardisée, nous désirons maintenant avancer un autre argument. Il s'agit de la conceptualisation de l'examen comme une démonstration d'autorité pédagogique portant un potentiel de violence symbolique. C'est que, selon Bourdieu et Passeron (1970), « rien n'est mieux fait que l'examen pour inspirer à tous la reconnaissance de la légitimité des verdicts scolaires et des hiérarchies sociales qu'ils légitiment... » (p. 198). En amalgamant en une seule action l'évaluation ministérielle, l'analyse de l'efficacité du système et l'évaluation sommative des apprentissages des élèves, le Ministère se trouve en porte-à-faux entre deux idéologies simultanément présentes dans *Les enfants au premier plan*. En effet, le discours qui y est véhiculé est tiraillé entre un discours humaniste démocrate et un discours hérité de la révolution industrielle qui, finalement, instrumentalise l'élève comme rouage de la machine étatique visant la croissance de la productivité sans égard aux circonstances socioéconomiques particulières ou à l'unicité de chaque enfant.

C'est là que se manifeste le caractère oppressif de l'évaluation ministérielle : par l'élaboration, par une élite, de standards uniques de réussite qui désavantagent les élèves de certaines régions et de certaines classes socioéconomiques. La relégation de ces élèves à des filières moins prestigieuses se voit alors justifiée, comme l'écrivent Bourdieu et Passeron (1970), par l'analyse purement docimologique de l'examen : les analyses d'items effectuées avec des modèles statistiques complexes attestant ses qualités métrologiques, on ne peut qu'en accepter le verdict. Or, l'examen sociologique, lui, peut révéler une tout autre réalité : celle d'un examen qui favorise systématiquement certains types d'élèves plutôt que d'autres en ignorant le contexte dans lequel s'inscrit la performance. Ne serait-il donc pas possible de séparer les deux fonctions de l'évaluation ministérielle : conserver sa fonction d'évaluation du système (qui ne requiert pas de communiquer les résultats aux élèves, parents et enseignants) et confier la responsabilité de l'évaluation des apprentissages aux écoles?

L'éducation et la gestion axée sur la performance

Précisons maintenant en quoi consiste ce mode de gestion de la performance et comment il inféode le domaine de l'éducation par l'entremise des évaluations standardisées. D'abord, selon Normand (2005, 2011), les théories psychologiques, inspirées des tests d'intelligence de Stanford-Binet, ainsi que le taylorisme ont modifié la structure et la dynamique des écoles. Si les tests standardisés ont servi à évaluer et à classer les élèves en fonction de leur quotient intellectuel pour accéder à tel ou tel programme d'études, le taylorisme a permis de diviser et de coordonner le travail en fonction des principes de son organisation scientifique. Or, ces principes, c'est-à-dire une stricte division entre les tâches de conception et d'exécution, de même que la décomposition du processus de production en une suite de tâches simples confiées chacune à un ouvrier spécialisé, poursuivaient un objectif, soit celui

d'obtenir une plus grande efficacité par une plus grande efficience. Ce faisant, l'acquisition des connaissances fut fragmentée en différentes matières et réduite à sa dimension mesurable. Il en fut de même avec l'organisation de l'école, axée sur sa dimension managériale et productive, de même qu'avec la gouvernance du système éducatif, qui fait de la réussite scolaire la base de ses politiques. Du reste, Petrella (2000) met en évidence cinq pièges tendus à l'éducation, dont le premier consiste en l'influence marquée des écoles de management axées sur les impératifs de la productivité et de la performance afin que l'éducation pour la ressource humaine prenne le pas sur l'éducation pour et par la personne humaine. Il en irait de même avec les nouvelles tendances d'évaluation institutionnelle qui s'appuient sur des concepts tels que l'obligation de rendre compte (*accountability*), l'étalement (*benchmarking*), le palmarès des écoles et le classement au *Program for International Student Assessment* (PISA) (Nusche et Shewbridge, 2013; OCDE, 2009, 2013, 2014). Ainsi, nous avançons que c'est dans ce cadre que *Les enfants au premier plan* inscrit l'évaluation ministérielle: un cadre de gestion de l'atteinte des cibles du système entier, censé être le meilleur au pays en 2013 selon une liste d'indicateurs qui inclut la performance globale des élèves en littératie, en numératie et en sciences (Ministère de l'Éducation du Nouveau-Brunswick, 2006).

Des dérives de la focalisation exclusive sur une gestion axée sur la performance

L'implantation de cette stratégie afin de classer la province en tête de peloton parmi les provinces canadiennes conduit à identifier, développer, puis mesurer des résultats d'apprentissages spécifiques en les considérant comme des capacités clés. Or, selon Wilson, Croxson et Atkinson (2006), cette situation fait en sorte que les indicateurs de performance changent les comportements des éducateurs, notamment des directions d'école, qui exercent une pression afin de mettre l'accent sur l'atteinte des objectifs, par exemple les matières de base au détriment d'autres contenus d'apprentissage tels que la citoyenneté et l'actualisation de la personne. Ainsi, des habiletés attribuables habituellement aux arts et aux humanités n'ont plus la même importance que des objectifs d'apprentissage prioritaires et plus facilement mesurables. Les dysfonctions engendrées par la gestion axée sur la performance sont aussi documentées par Wiggins et Tymms (2000), qui soulignent leurs effets pernicieux à long terme sur la gestion scolaire lorsque l'on privilégie certains indicateurs de performance considérés isolément sans tenir compte d'une vision systémique.

L'individualité et la créativité propres à l'éducation humaniste peinent à se développer dans ces conditions, car, lorsque ce qui est testé risque d'influencer l'avenir d'un système, les échanges entre les enseignants et les élèves deviennent de plus en plus orientés directement vers ce qui sera testé (rétrécissement du curriculum). Une routine s'installe autour des tests et les élèves perdent le pouvoir d'influencer l'orientation que pourraient prendre les apprentissages et les activités qui y sont associées. La pression (parfois auto-imposée) exercée sur les enseignantes et les

enseignants est forte pour adopter une pédagogie de l'apprentissage plus structurée dans un cadre plus facile à contrôler (Demazière, Lessard et Morissette, 2013; Dutercq et Cuculou, 2013; Maroy et Vaillancourt, 2013).

Dans ce contexte, à la fois la pédagogie et les contenus sont touchés, puisque l'un et l'autre sont subordonnés à l'objectif d'accroître continuellement les taux de réussite. Comme l'impératif est celui de la réussite de l'évaluation, le questionnement et la responsabilisation individuelle de l'élève dans ses apprentissages deviennent superflus et, à la limite, contre-productifs. Freire (1974) le souligne. Dès le commencement de la lutte pour l'humanisation, pour le dépassement de la contradiction oppresseur/opprimé, il est nécessaire que les apprenants sachent que cette lutte exige d'eux, à partir du moment où ils s'y engagent, une responsabilité totale. Car la justification de cette lutte n'est pas la conquête de la liberté de manger à sa faim, mais d'une «liberté pour créer et construire, pour découvrir et aller à l'aventure» (p. 47).

Ainsi, le danger est bien réel de mal interpréter ce qui pourrait vraiment compter dans le développement de l'enfant en recourant, dans le but de définir les objectifs et les cibles, à des critères qui découlent de l'évaluation qui sont eux-mêmes influencés par des impératifs de performance. Même si l'on admet que l'évaluation vise la mesure de la performance du système, en dernière analyse, quand l'objectif explicitement déterminé est de devenir le meilleur système d'éducation du pays, il est clair que la comparaison se fera à partir des résultats des élèves. Dans ces conditions, le danger est grand de confondre les moyens avec les fins. Il serait judicieux d'être conscient qu'il existe une différence fondamentale entre l'évaluation d'un système et l'évaluation des personnes. Il faudrait éviter que l'évaluation du système entraîne une sorte de condamnation des élèves à être confinés dans un rôle consistant à améliorer la performance du système. La composante humaniste de la mission de l'école publique du Nouveau-Brunswick incite à penser que tout le système d'éducation sera au service du développement des personnes et que toute évaluation consistera à déterminer si les moyens mis en place permettent d'arriver aux fins souhaitées. Comme le souligne Baby (2013), les objectifs institutionnels de réussite et les objectifs individuels sont distincts l'un de l'autre.

C'est dans ce sens qu'Illich (2003, p. 13) parle de «grand enfermement». Lorsque l'élève, pour satisfaire des impératifs de performance du système, perd les moyens nécessaires pour se créer une identité propre, il se voit imposer un modèle d'existence qui lui vient d'une organisation extérieure. Foucault (1994) soulignait à quel point la norme vient articuler la société aux dépens de la loi qui, malgré tout, tend davantage à imposer une surveillance et un contrôle des personnes et de leurs actions. L'élève perd ainsi de sa liberté à devenir lui-même, puisque l'institution prétend savoir ce qui est bien pour lui et l'oriente en ce sens. Castoriadis (1999, p. 164) mentionne d'ailleurs que «[...] ce qui pourrait être vu "au départ" comme un ensemble d'institutions au service de la société devient une société au service des institutions». L'éducation et la culture ne remplissent plus leur mission de contribuer à la construction du sens de l'expérience des individus en inscrivant leur existence singulière et collective dans l'histoire de l'humanité. Bien au contraire, comme le

soulignent Abelhauser, Gori et Sauret (2011), la culture propage de nouvelles valeurs suggérées par les évaluations.

Illich (2003, p. 43) utilise aussi le concept de l'outil pour illustrer l'activité autant fabricatrice qu'organisatrice de l'humain. Ainsi, « pour autant que je maîtrise l'outil, je charge le monde de mon sens; pour autant que l'outil me domine, sa structure me façonne et informe la représentation que j'ai de moi-même » (Illich, 2003, p. 44). Dans le cas de l'évaluation, il semble plutôt s'agir de la deuxième option. Nous pourrions ainsi voir que, sur le plan scolaire, les personnes ne sont plus en mesure d'être les maîtres de leurs apprentissages. En fait, ajoute-t-il, « [l']outil peut croître de deux façons, soit qu'il augmente le pouvoir de l'homme, soit qu'il le remplace » (*ibid.*, p. 124). L'outil contre-productif serait donc celui qui, se voulant au départ un simple moyen, modifie l'expérience et devient une fin en soi.

Afin de renverser cette tendance, Illich (2003) plaide pour un rétablissement d'un équilibre entre la vie sociale et les outils. Il suggère de modifier le rapport qu'entretient le sujet avec l'outil en rendant celui-ci convivial :

L'outil convivial est celui qui me laisse la plus grande latitude et le plus grand pouvoir de modifier le monde au gré de mon intention. L'outil industriel dénie ce pouvoir; bien plus, à travers lui, un autre que moi détermine ma demande, rétrécit ma marge de contrôle et régit mon sens. [...] L'outil est convivial dans la mesure où chacun peut l'utiliser, sans difficulté, aussi souvent ou aussi rarement qu'il le désire, à des fins qu'il détermine lui-même (Illich, 2003, p. 44-45).

Ce faisant, les individus ont la possibilité d'agir d'une manière plus autonome puisque les outils exercent moins de contrôle sur eux. Certes, l'évaluation demeure inévitable, mais la valeur fondamentale de l'éducation, apprendre à être une personne libre, autonome et responsable, demeure préservée. Dans un plaidoyer en faveur d'un système d'éducation qui laisse le soin aux disciplines des sciences sociales d'éduquer à la citoyenneté, Nussbaum (2010) note que les lycées britanniques ont un système d'évaluation où les élèves doivent rédiger des essais que lisent plusieurs enseignants correcteurs. Évidemment, cette modalité d'évaluation est plus coûteuse, mais elle a l'avantage d'évaluer des dimensions, comme la capacité à analyser et à penser de manière critique, qui s'évaluent difficilement autrement. Le questionnement soulevé par Cormier (2011) prend ici une autre forme : est-ce que le Ministère promeut l'usage des bons outils pour atteindre les bonnes cibles? En d'autres mots, à quelles cibles devrait-on associer l'outil qu'est l'évaluation standardisée et, une fois cette cible correctement identifiée, quel est le mode d'emploi optimal de l'outil pour assurer qu'il ne devienne pas un obstacle aux autres cibles proposées par *Les enfants au premier plan*?

Le rôle de l'évaluation standardisée dans un dispositif public d'éducation

Les deux sections précédentes articulent deux positions, la montée du management axé sur la performance en éducation et les fondements humanistes de l'entreprise éducative dans les états démocratiques, qui peuvent sembler difficilement conciliables. En fait, nous soutenons qu'elles peuvent s'avérer raisonnablement compatibles, mais pas à travers les mêmes outils. L'évaluation des apprentissages rompt déjà avec l'éducation libertaire issue d'un humanisme radical, elle-même difficile à concevoir dans le contexte du système d'éducation public d'un État capitaliste post-révolution industrielle. Cela dit, l'évaluation peut prendre des formes très différentes de celles proposées par une approche du management axé sur la performance, notamment en prévoyant une rétroaction fréquente et détaillée. Les écoles du Nouveau-Brunswick regorgent d'exemples de pratiques enseignantes et évaluatives qui sont axées davantage sur le développement de l'élève et la différenciation pédagogique. Cependant, nous devons admettre, avec Cormier (2011), que les examens ministériels témoignent d'une autre philosophie, plus cohérente avec le management axé sur la performance. Il suffit, pour s'en convaincre, de retourner aux principes docimologiques élémentaires, par exemple la transparence de l'évaluation et la rétroaction subséquente à l'élève sous forme non seulement d'une note, mais aussi d'une appréciation qualitative de ses forces et faiblesses susceptible de favoriser ses apprentissages ultérieurs (évaluation au service de l'apprentissage) (Richard, 2004). Les examens ministériels dérogent à ces principes élémentaires et, du coup, ne peuvent se prétendre au service direct de l'apprentissage, quel que soit le contenu.

Force nous est donc de constater que, pour que le titre *Les enfants au premier plan* s'avère une description juste du plan stratégique mis en œuvre, une importance égale, voire supérieure, devra être accordée aux cibles plus humanistes privilégiant le développement harmonieux de l'enfant plutôt qu'aux cibles (et moyens) relevant davantage du management axé sur la performance. Quant aux outils du management axé sur la performance, ils devront être déployés de façon cohérente avec leur finalité, c'est-à-dire celle d'informer les décideurs sur l'évaluation de l'efficacité du système éducatif et non pas pour clouer publiquement élèves et enseignants au pilori si les performances attendues ne sont pas réalisées. Par exemple, l'évaluation formative et sommative pourrait être confiée aux enseignantes et enseignants. Les résultats à ces évaluations, accompagnés d'une rétroaction constructive détaillée, seraient transmis promptement et régulièrement aux parents et aux élèves eux-mêmes. Quant aux évaluations standardisées, elles pourraient se limiter à quelques moments charnières du parcours scolaire. Leurs résultats ne serviraient qu'à informer le Ministère du rendement du système éducatif.

Références bibliographiques

- ABELHAUSER, A. GORI, R. et SAURET, M.-J. (2011). *La folie évaluation : les nouvelles fabriques de la servitude*. Paris : Mille et une nuits.
- BABY, A. (2013). *Qui a eu cette idée folle? Essais sur l'éducation scolaire*. Québec, QC : Presses de l'Université du Québec.
- BANGERT-DROWNS, R. L., KULIK, C. L. C., KULIK, J. A. et MORGAN, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213-238.
- BERNARD, M. (2009). L'évaluation des apprentissages des élèves : un cafouillis. *Le Nouvelliste*. Récupéré de <http://www.lapresse.ca/le-nouvelliste/opinions/200906/22/01-877863-levaluation-des-apprentissages-des-eleves-un-cafouillis.php>
- BOURDIEU, P. et PASSERON, J.-C. (1970). *La reproduction : éléments pour une théorie du système d'enseignement*. Paris : Éditions de Minuit.
- BOURQUE, J. (2009). La construction sociale de la réussite scolaire et le rôle des utopies. *Info-CRDE*, 13, 2-3.
- CASTORIADIS, C. (1999). *L'institution imaginaire de la société*. Paris : Les Éditions Points.
- CORBEIL, J.-P. (2006). *Le volet canadien de l'Enquête internationale sur l'alphabétisation et les compétences des adultes de 2003 (EIACA) : état de la situation chez les minorités de langue officielle*. Ottawa, ON : Statistique Canada.
- CORMIER, M. (2011). Au premier plan : les enfants ou les résultats? *Éducation et francophonie*, 39(1), 7-25.
- CROCKER, J. et PARK, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin*, 130(3), 392-414.
- De KETELE, J.-M. (2013). Les épreuves externes servent-elles l'action? *Revue française de linguistique appliquée*, 18(1), 9-27.
- DEMAZIÈRE, D., LESSARD, C. et MORISSETTE, J. (2013). Les effets de la Nouvelle Gestion publique sur le travail des professionnels : transpositions, variations, ambivalences. *Éducation et sociétés*, 2(32), 5-20.
- DODGE, A. (2009). Heuristics and NCLB standardized tests: A convenient lie. *International Journal of Progressive Education*, 5(2), 1-19.
- DUTERCQ, Y. et CUCULOU, S. (2013). La performance comme outil de gouvernance : quelles conséquences sur le travail des enseignants de l'école primaire? Dans C. Maroy (dir.), *L'école à l'épreuve de la performance : les politiques de régulation par les résultats*. Bruxelles : De Boeck.

- FOUCAULT, M. (1994). La psychologie de 1850 à 1950. Dans *Dits et écrits*, I. Paris : Gallimard.
- FREIRE, P. (1974). *Pédagogie des opprimés; suivi de Conscientisation et révolution*. Paris : Éditions F. Maspero.
- HADJI, C. (2012). *Faut-il avoir peur de l'évaluation?* Bruxelles : De Boeck.
- HATTIE, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. New York, NY : Routledge.
- ILLICH, I. (2003). *La convivialité*. Paris : Les Éditions Points.
- JEFFREY, D. (2013). *L'éthique dans l'évaluation scolaire*. Québec, QC : Presses de l'Université Laval.
- MAROY, C. et VAILLANCOURT, S. (2013). Le discours syndical face à la nouvelle gestion publique dans le système éducatif québécois. *Éducation et sociétés*, 2(32), 93-108.
- MARSH, H. W. et O'MARA, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality and Social Psychology Bulletin*, 34(4), 542-552.
- MINISTÈRE DE L'ÉDUCATION DU NOUVEAU-BRUNSWICK (2006). *Les enfants au premier plan. Faire de notre système d'éducation publique le meilleur au Canada : un défi à relever ensemble!* Fredericton, NB : Ministère de l'Éducation du Nouveau-Brunswick.
- MINISTÈRE DE L'ÉDUCATION DU QUÉBEC (2002). *À chacun son rêve. Pour favoriser la réussite : l'approche orientante*. Québec, QC : Gouvernement du Québec.
- NORMAND, R. (2005). La mesure de l'école : politiques des standards et management par la qualité. *Cahiers de la recherche sur l'éducation et les savoirs*, 1, 67-82.
- NORMAND, R. (2011). *Gouverner la réussite scolaire : une arithmétique politique des inégalités*. Berne : Peter Lang.
- NUSCHE, D. et SHEWBRIDGE, C. (avril 2013). Evaluation and assessment is for everyone. *Education Today*. Récupéré de <http://oecdeducationtoday.blogspot.fr/2013/04/evaluation-and-assessment-is-for.html>
- NUSSBAUM, M. (2010). *Les émotions démocratiques : comment former le citoyen du XXI^e siècle?* Princeton, NJ : Princeton University Press.
- OCDE (février 2014). *Résultats du PISA 2012 : les clés de la réussite des établissements d'enseignement* (vol. IV). Paris : Éditions OCDE.
- OECD (juillet 2009). *Review on Evaluation and Assessment Frameworks for Improving School Outcomes*. Paris : OECD Reviews of Evaluation and Assessment in Education.

- OECD (avril 2013). *Synergy for Better Learning: An international Perspective on Evaluation and Assessment*. Paris: OECD Reviews of Evaluation and Assessment in Education.
- OGBU, J. U. et SIMONS, H. D. (1998). Voluntary and involuntary minorities: A cultural-ecological theory of school performance with some implications for education. *Anthropology & Education Quarterly*, 29(2), 155-188.
- PETRELLA, R. (2000). *L'éducation, victime de cinq pièges*. Montréal: FIDES.
- RICHARD, J.-F. (2004). *L'intégration de l'évaluation dans le processus enseignement-apprentissage*. Fredericton, NB: Direction de la mesure et de l'évaluation, Gouvernement du Nouveau-Brunswick.
- ROSS, C. E. et BROH, B. A. (2000). The roles of self-esteem and the sense of personal control in the academic achievement process. *Sociology of Education*, 73(4), 270-284.
- UNESCO (2011). *Faire des choix fondamentaux pour l'assurance qualité externe: module 1*. Récupéré de http://www.iiep.unesco.org/fileadmin/user_upload/Cap_Dev_Training/Training_Materials/HigherEd/AQE_ES_1.pdf
- WILSON, D., CROXSON, B. et ATKINSON, A. (2006). What gets measured gets done: Headteachers' responses to the English Secondary School Performance Management System. *Policy Studies*, 27(2), 153-171.

La politique québécoise d'évaluation des apprentissages et les pratiques évaluatives

Michel LAURIER

Université d'Ottawa, Ontario, Canada

VOLUME XLII : 3 – NUMÉRO SPÉCIAL,
AUTOMNE 2014

Revue scientifique virtuelle publiée par l'Association canadienne d'éducation de langue française dont la mission est la suivante : « Par la réflexion et l'action de son réseau pancanadien, l'ACELF exerce son leadership en éducation pour renforcer la vitalité des communautés francophones ».

Directrice de la publication

Chantal Lainey, ACELF

Présidente du comité de rédaction

Lucie DeBlois,
Université Laval

Comité de rédaction

Sylvie Blain,
Université de Moncton
Lucie DeBlois,
Université Laval
Nadia Rousseau,
Université du Québec à Trois-Rivières
Paul Ruest,
Université de Saint-Boniface
Mariette Thérberge,
Université d'Ottawa

Directeur général de l'ACELF

Richard Lacombe

Conception graphique et montage

Claude Baillargeon

Responsable du site Internet

Étienne Ferron-Forget

Diffusion Érudit

www.erudit.org

Les textes signés n'engagent que la responsabilité de leurs auteurs et auteurs, lesquels en assument également la révision linguistique. De plus, afin d'attester leur recevabilité, au regard des exigences du milieu universitaire, tous les textes sont arbitrés, c'est-à-dire soumis à des pairs, selon une procédure déjà convenue.

La revue *Éducation et francophonie* est publiée deux fois l'an grâce à l'appui financier du ministère du Patrimoine canadien et du Conseil de recherches en sciences humaines du Canada.

acelf

ASSOCIATION
CANADIENNE
D'ÉDUCATION DE
LANGUE FRANÇAISE

265, rue de la Couronne, bureau 303
Québec (Québec) G1K 6E1
Téléphone : 418 681-4661
Télécopieur : 418 681-3389
Courriel : info@acelf.ca

Dépôt légal
Bibliothèque et Archives nationales
du Québec
Bibliothèque et Archives du Canada
ISSN 1916-8659 (En ligne)
ISSN 0849-1089 (Imprimé)

Les politiques d'évaluation dans le domaine de l'éducation

Rédacteur invité :

Dany LAVEAULT

Liminaire

1 Les politiques d'évaluation en éducation. Et après?

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

15 La formule actuelle d'évaluation ministérielle au Nouveau-Brunswick. Pour qui? Pourquoi?

Jimmy BOURQUE, Université de Moncton, Nouveau-Brunswick, Canada

Mathieu LANG, Université de Moncton, Nouveau-Brunswick, Canada

Jean LABELLE, Université de Moncton, Nouveau-Brunswick, Canada

31 La politique québécoise d'évaluation des apprentissages et les pratiques évaluatives

Michel LAURIER, Université d'Ottawa, Ontario, Canada

50 La politique d'évaluation du rendement en Ontario : un alignement qui se précise dans la persévérance et la durée

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

Louise BOURGEOIS, Université d'Ottawa, Ontario, Canada

68 Une recherche documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public (1^{re} à 12^e année) des provinces de l'Ouest canadien

Jules ROCQUE, Université de Saint-Boniface, Manitoba, Canada

85 L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds? Les enseignements de l'histoire récente de l'école primaire genevoise

Lucie MOTTIER LOPEZ, Université de Genève, Suisse

102 Comparer les systèmes éducatifs francophones à travers le monde grâce au PISA : pas si simple!

Ariane BAYE, Université de Liège, Belgique

Marc DEMEUSE, Université de Mons, Belgique

Nathanaël FRIANT, Université de Mons, Belgique

La politique québécoise d'évaluation des apprentissages et les pratiques évaluatives

Michel LAURIER

Université d'Ottawa, Ontario, Canada

RÉSUMÉ

En 2003, le ministère de l'Éducation du Québec faisait paraître une politique sur l'évaluation des apprentissages en milieu scolaire. Il fallait alors tenir compte des approches pédagogiques qui sous-tendaient le nouveau curriculum (notamment le socioconstructivisme), du mode d'organisation en matière de compétences, de la réorganisation du cursus amenée par la notion de cycle et de la redéfinition des rôles des divers acteurs. Après plus de dix années marquées par de nombreux questionnements, certaines incohérences apparaissent entre les pratiques mises en place et les orientations de la politique. On observe que celle-ci s'est déployée avec l'appui de documents complémentaires et de modifications aux règlements qui en précisent et parfois en réduisent la portée.

ABSTRACT

The Québec learning assessment policy and evaluation practices

Michel LAURIER
University of Ottawa, Ontario, Canada

In 2003, the Québec Ministry of Education issued a policy on the assessment of learning in schools. It then became necessary to consider the pedagogical approaches underpinning the new curriculum (including social constructivism), the skill organization mode, the reorganization of the curriculum based on the cycle concept, and the redefinition of the roles of various actors. Ten years and many questions later, some inconsistencies have appeared between the policy's practices and guidelines. We observe that the policy was deployed with the support of complementary documents and amendments to the regulations that clarify it and sometimes reduce its scope.

RESUMEN

La política quebequense de evaluación del aprendizaje y las prácticas evaluativas

Michel LAURIER
Universidad de Ottawa, Ontario, Canadá

En 2003, el ministerio de Educación de Quebec publicó la política de la evaluación del aprendizaje escolar. Fue necesario tomar en consideración los enfoques pedagógicos que servían de base al nuevo currículo (en especial el socio-constructivismo), la organización de las materias por competencia, la reorganización de los cursos provocada por la noción de ciclo y la redefinición de los roles de los diversos actores. Después de más de diez años marcados por innumerables interrogaciones, aparecieron ciertas incoherencias entre las aplicaciones y las orientaciones de la política. Se observa que ésta última se ha desplegado gracias al apoyo de documentos complementarios y modificaciones a los reglamentos que han precisado y a veces reducido su alcance.

Introduction

La question de l'évaluation des apprentissages a toujours suscité des débats passionnés dans les milieux éducatifs au Québec. Autour des enjeux de l'évaluation, se noue souvent l'issue des réformes éducatives. On y voit se cristalliser les résistances au changement qui poussent certains acteurs à reproduire les situations antérieures, de même que les failles d'un curriculum qui ne permet pas toujours de réunir les conditions nécessaires à la mise en œuvre d'une évaluation rigoureuse et pertinente. Dans cet article, nous retracerons l'évolution des politiques et des pratiques évaluatives dans le contexte de l'implantation de la réforme amorcée au Québec au début des années 2000. Une décennie s'étant écoulée depuis la publication de la dernière politique québécoise d'évaluation des apprentissages, cette analyse se construit à partir du regard rétrospectif et critique qu'y jette un membre de l'équipe chargée de son élaboration et qui apportait le point de vue d'un chercheur dans le domaine de l'évaluation des apprentissages.

Nous décrirons brièvement le contexte dans lequel a été introduite la politique québécoise d'évaluation des apprentissages, puis nous en présenterons les composantes, c'est-à-dire les valeurs sur lesquelles elle repose, la démarche d'évaluation qui est proposée et les orientations qui sont mises en avant. Dans un second temps, nous adopterons une position plus critique pour expliquer les difficultés d'implantation et les enjeux que soulèvent les directives en ce qui a trait à la notation.

Le contexte

L'éducation étant, au Canada, de compétence provinciale, les programmes québécois ont évolué selon la logique imposée par le ministère de l'Éducation du Québec (MEQ) et les tendances inspirées d'une recherche en évaluation qui se situe au confluent de la tradition nord-américaine et de la tradition européenne.

Au début des années 1980, le ministère de l'Éducation du Québec (MEQ) publiait une première politique sur l'évaluation des apprentissages (MEQ, 1981) dans la foulée d'une remise en question des orientations de l'école dont les résultats étaient résumés dans un document (MEQ, 1979) paru deux ans plus tôt. Il s'agissait de la première réflexion collective sur l'état du système éducatif québécois depuis le rapport Parent, un document déterminant paru dans les années 1960 dans la foulée de la Révolution tranquille. Dans une perspective néo-béhavioriste, on prônait alors l'élaboration de programmes d'études par objectifs (généraux et spécifiques), ce qui a amené un découpage fin de la matière auquel devait répondre l'évaluation des apprentissages. La politique de 1981 adoptait donc cette perspective et mettait l'accent sur l'évaluation systématique des objectifs des programmes d'études. On y trouvait une distinction, héritée de Scriven (1967) et de Bloom (1969), entre l'évaluation formative et l'évaluation sommative. Il faut noter qu'à l'origine cette distinction avait été proposée pour l'évaluation de programmes; c'est par la suite qu'elle a été appliquée à l'évaluation des apprentissages. L'introduction du concept d'évaluation

formation dans la politique québécoise n'est d'ailleurs pas étrangère à la transposition théorique du concept dans une perspective de régulation des apprentissages, dans la recherche francophone (Allal, 1991; Bonniol et Vial, 1997; Perrenoud, 1998), et à travers la notion d'évaluation centrée sur l'élève, dans la recherche anglo-saxonne (Stiggins, 2005; Shepard, 2007; Black et Wiliam, 2009; Andrade, 2009).

La politique de 1981 posait comme valeurs fondamentales la justice et l'égalité. Elle reflétait l'émergence de préoccupations d'inclusion en prônant l'idée d'égalité des chances de réussite. Elle témoignait aussi d'un souci, parfois opposé à la notion d'inclusion, d'évaluer de façon uniforme et donc indifférenciée. Enfin, la politique de 1981 définissait les rôles et les responsabilités des différents acteurs (enseignants, directions d'école, parents et élèves), annonçant par le fait même une vision selon laquelle l'évaluation était une responsabilité partagée et non pas une responsabilité découlant du pouvoir du maître. La politique faisait de l'enseignant le maître d'œuvre de l'évaluation formative. Toutefois, peu de temps après la publication du document, un avis du Conseil supérieur de l'éducation (1982) reconnaissait la difficulté pour les enseignants de concevoir l'évaluation formative autrement que comme un exercice d'évaluation sans notes.

Dans la seconde moitié des années 1990 s'est imposée la nécessité de repenser le système éducatif québécois. Un groupe de travail a été mis sur pied par le MEQ pour mener une vaste consultation sur les enjeux de l'éducation de base au Québec à l'aube du 21^e siècle. Cette consultation a mené au dépôt d'un rapport, *Réaffirmer l'école* (MEQ, 1997), connu sous le nom de «rapport Inchauspé», du nom du président du groupe de travail. Ce rapport proposait des moyens pour réaliser les trois missions de l'éducation (instruire, socialiser et qualifier), de même que des correctifs quant à l'évaluation et à la sanction des apprentissages. Pour faire suite à ce rapport, le MEQ s'est engagé à lier les changements dans l'organisation scolaire à un renouvellement du curriculum, du préscolaire (maternelle) à la fin du secondaire (11^e année). Le rapport visait à établir les bases d'une compréhension commune du «renouveau pédagogique» en précisant la mission éducative, les domaines d'apprentissage, le profil de formation, les disciplines communes ainsi que le contenu de formation «transversale». Sans imposer ni organisation curriculaire ni approche pédagogique particulières, le document ouvrait la voie à l'introduction du concept de compétence et à l'application de principes issus des théories constructivistes.

La nouvelle politique d'évaluation des apprentissages annoncée dans *Réaffirmer l'école* a mis du temps à paraître. De fait, lorsque la Politique d'évaluation des apprentissages (MEQ, 2003, ci-après la Politique) a été publiée, le nouveau curriculum était déjà implanté aux deux premiers cycles de l'enseignement primaire. Cette parution tardive reflète les longues discussions, voire les tergiversations, qui entouraient la question de l'évaluation des apprentissages dans le cadre du renouveau pédagogique. Elle reflète aussi l'ampleur de la consultation qui s'est effectuée afin que le document traduise les consensus en ce qui a trait à l'évaluation des apprentissages dans le contexte d'une réforme que les milieux scolaires et l'opinion publique jugeaient parfois sévèrement.

Les valeurs

La politique de 2003 reprend les valeurs fondamentales de justice et d'égalité. La justice revêt une importance particulière dans les situations d'évaluation liées à la sanction des études. Cette fonction implique, par exemple, la mise en place d'un dispositif permettant l'exercice d'un droit de reprise ou d'un droit d'appel. Quant à l'égalité, elle implique que tous les élèves ont des chances égales de démontrer les apprentissages qu'ils ont réalisés. Par exemple, des exigences uniformes doivent être définies en fonction des résultats attendus et des critères d'évaluation spécifiés dans les programmes de formation. Inspirés par les théories sur la justice (Rawls, 1971), les auteurs de la politique ont constaté qu'une intervention conçue comme juste et égale ne garantissait pas la mise en œuvre de moyens assurant la réussite du plus grand nombre que préconisait le rapport *Réaffirmer l'école*. C'est pourquoi la valeur d'équité s'est ajoutée à celles de justice et d'égalité pour former un ensemble de trois valeurs fondamentales en interaction entre elles. C'est en vertu de l'équité que les pratiques d'évaluation doivent respecter les caractéristiques individuelles ou communes à certains groupes. Il faut non seulement s'assurer que l'évaluation est exempte de biais, mais aussi mettre en application au moment de l'évaluation les principes de différenciation pédagogique que prône le renouveau pédagogique.

Une politique étant un document qui s'adresse à un public relativement large, on doit éviter l'utilisation de termes trop techniques. C'est ainsi que les termes « validité » et « fidélité » sont quasi absents. Toutefois, même si le document n'y fait pas explicitement référence, ces exigences de mesure s'expriment à travers trois valeurs instrumentales qui se superposent aux valeurs fondamentales.

- **Cohérence** – C'est en vertu de la cohérence qu'il faut s'assurer que l'évaluation est en lien direct avec l'apprentissage et le programme de formation qui l'encadre. La cohérence dépasse la représentativité des éléments de contenu ainsi que la congruence entre les objets de l'évaluation et la performance qui est observée. Elle suppose aussi que les tâches soumises à l'élève pour amener cette performance sont alignées vers les finalités du programme et, plus largement, vers les missions de l'école.
- **Rigueur** – La valeur de rigueur impose que l'évaluation soit menée avec un souci de précision et d'exactitude. Le respect de cette valeur est donc garant de la fidélité. L'évaluation doit être planifiée et il importe de mettre en œuvre des moyens permettant de recueillir une information pertinente et suffisante en vue des inférences. Il faut aussi interpréter les données de façon à poser un jugement qui découle de ces données plutôt que d'une lecture impressionniste. Enfin, une démarche rigoureuse commande un suivi : une évaluation est inutile si elle ne se traduit pas par des décisions et des actions appropriées.
- **Transparence** – La Politique met l'accent sur l'importance de communiquer les enjeux et les résultats de l'évaluation. La valeur de transparence renvoie à la clarté et à l'intelligibilité de l'information qui est transmise dans le cadre de l'évaluation. Le respect de cette valeur se manifeste sur plusieurs plans. D'abord, le contexte réglementaire qui entoure l'évaluation doit être communiqué

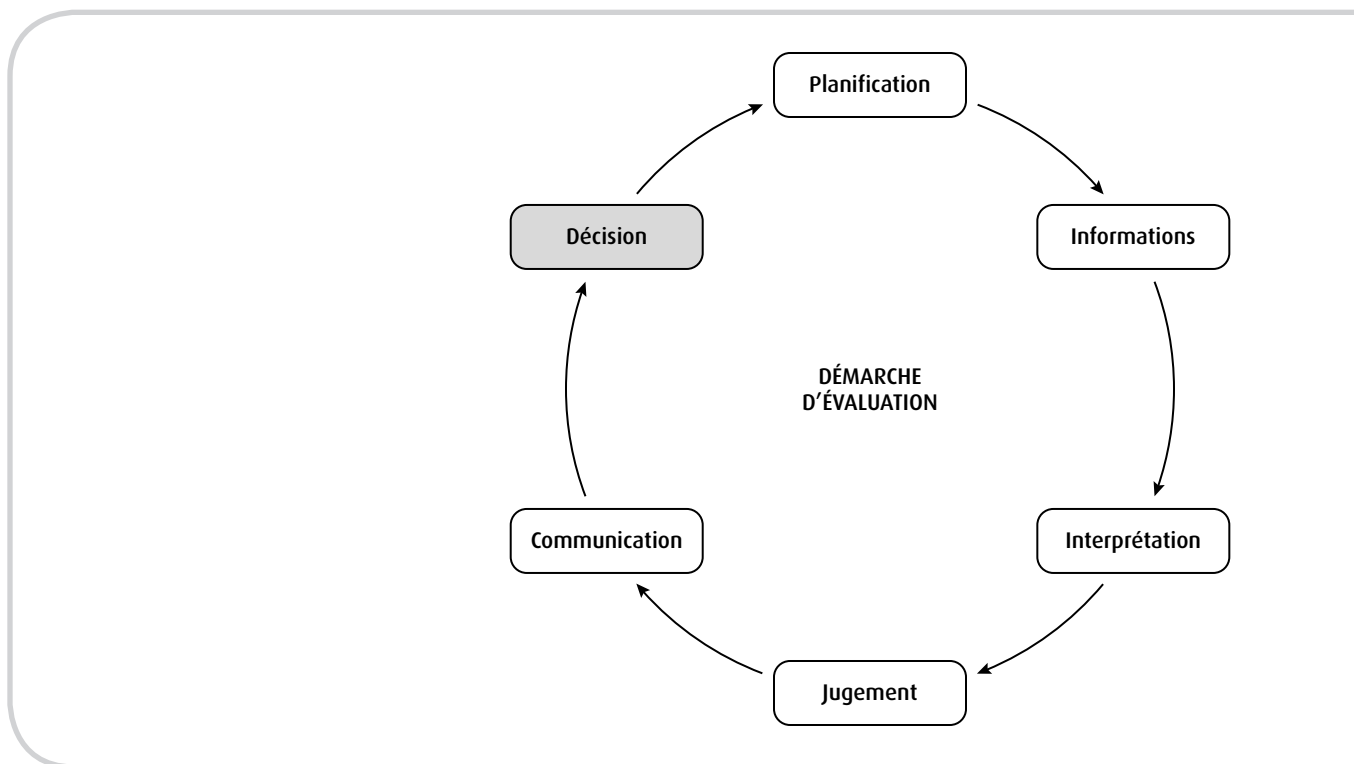
adéquatement : les normes et les modalités d'évaluation doivent être connues et comprises par tous les acteurs. De plus, l'élève doit savoir sur quoi il sera évalué, ce qu'on attend de lui; il doit aussi comprendre les jugements et les décisions qui le concernent. Une évaluation de type formatif s'accompagne d'une rétroaction pertinente et claire sur les apprentissages. Enfin, l'information concernant les apprentissages de l'élève doit être accessible aux destinataires, qui doivent facilement la comprendre.

La démarche d'évaluation

La Politique (p. 3) donne la définition suivante de l'évaluation des apprentissages :

L'évaluation est le processus qui consiste à porter un jugement sur les apprentissages, soit des connaissances et des compétences disciplinaires, à partir de données recueillies, analysées et interprétées, en vue de décisions pédagogiques et, le cas échéant, administratives (MEQ, 2003).

Figure 1. Les étapes de la démarche d'évaluation



Le processus dont il est question peut se présenter sous la forme illustrée par la figure qui précède. La case ombragée indique que la décision n'est pas à proprement parler une étape de l'évaluation, bien qu'elle en soit l'aboutissement et qu'elle détermine donc la forme des étapes précédentes du processus (Newton, 2007). Dans le but de sortir des ornières creusées au fil des ans par une interprétation un peu simpliste de la distinction entre évaluation formative et évaluation sommative où la première devient un exercice en vue de réussir la seconde, la Politique propose une distinction fondée sur le type de décision. La décision a une portée pédagogique lorsque l'intention est de fournir de l'aide à l'apprentissage, c'est-à-dire en vue d'une action de régulation de l'apprentissage. Dans l'exercice de la régulation, l'évaluation devient un moyen de planification, de diagnostic et de réflexion métacognitive (Allal, 2007; Morrissette, 2010; Mottier Lopez, 2012; Black et Wiliam, 2012). La comparaison des résultats entre les élèves occupe alors peu de place, puisqu'il s'agit d'informer ceux-ci sur leurs difficultés, leurs réalisations et leurs progrès personnels. La décision a une portée plus administrative lorsque l'intention est la reconnaissance des compétences, ce qui conduit à des actions qui revêtent un caractère plus formel.

Il faut aussi noter le caractère cyclique du processus, puisque la décision entraîne une action qui aura un effet sur l'apprentissage. La figure ne rend cependant pas compte du fait que la séquence des étapes n'est pas nécessairement linéaire : la direction des flèches n'exclut pas des allers-retours. Par exemple, la planification peut être revue en cours de processus.

La figure indique qu'une évaluation suppose d'abord une collecte d'information. La Politique s'écarte des approches antérieures fortement influencées par la psychométrie où cette information devait être recueillie à l'aide d'un instrument de mesure. On établit ici que l'information peut être de nature quantitative (une mesure) ou de nature qualitative (une observation). Plutôt qu'une interprétation normative en fonction d'un groupe de référence, on privilégie une interprétation de cette information en fonction des attentes du programme (interprétation critériée) ou une interprétation en fonction des progrès de l'élève.

Ainsi que le précise la définition ci-dessus, l'interprétation doit conduire à un jugement. Celui-ci est porté par l'enseignant, sauf dans des situations d'auto-évaluation ou d'évaluation par les pairs où l'objectif est formatif et où le jugement appartient aux élèves ou, à l'opposé, dans des situations d'évaluation certificative, particulièrement lorsqu'un organisme externe est responsable de porter le jugement. De ce point de vue, la Politique fait jouer un rôle central à l'enseignant dont le jugement professionnel est déterminant dans le processus d'évaluation. Enfin, le processus d'évaluation comporte une étape où les résultats doivent être communiqués. Cette communication peut prendre différentes formes selon le but fixé. En contexte formatif, où la régulation est intégrée à l'enseignement et doit fournir une rétroaction à l'élève, la communication devient l'étape finale du processus. Par contre, en contexte certificatif, la communication prend un aspect beaucoup plus formel et s'accompagne de suivis d'ordre plus administratif.

Les orientations

Sur la base des valeurs qui sous-tendent l'évaluation et de la définition qu'elle en donne, la Politique met l'accent sur une dizaine d'orientations.

1. *L'évaluation en cours de formation doit être intégrée à la dynamique des apprentissages de l'élève.* Qu'il s'agisse d'un bilan périodique consigné dans un bulletin ou, a fortiori, d'une intervention à des fins de régulation, l'évaluation n'est pas une fin en soi. Elle est plutôt au service de l'apprentissage. Non seulement l'évaluation doit fournir une information diagnostique pour guider l'enseignement et l'apprentissage, mais une tâche conçue pour l'évaluation peut elle-même devenir une occasion d'apprentissage. De ce fait, pour l'enseignant l'évaluation est une partie intégrante de sa tâche et ne peut pas en être dissociée.
2. *L'évaluation des apprentissages doit reposer sur le jugement professionnel de l'enseignant.* Évaluer est un acte professionnel auquel l'enseignant doit être préparé. Ce jugement, au cœur de la démarche d'évaluation, ne peut s'exercer de façon arbitraire, mais doit se fonder sur une information qui est recueillie soigneusement et que l'enseignant est apte à interpréter.
3. *L'évaluation des apprentissages doit s'effectuer dans le respect des différences.* Cette exigence suppose que l'enseignant pratique la différenciation pédagogique non seulement dans son enseignement, mais aussi dans son évaluation. Il lui faut tenir compte de facteurs tels que la maîtrise de la langue d'enseignement par l'élève, ses références culturelles, le niveau de développement de ses connaissances antérieures, ses difficultés à apprendre, ses limitations physiques, etc.
4. *L'évaluation des apprentissages doit être en conformité avec les programmes de formation et d'études.* Le curriculum du Québec est structuré autour du développement de compétences disciplinaires et transversales qu'il faut évaluer dans toute leur complexité et leur globalité. Cela n'exclut pas l'évaluation régulière des connaissances, lesquelles sont des composantes essentielles des compétences, particulièrement à des fins de régulation. Cela suppose cependant qu'on ait recours à des tâches intégratives (complexes et contextualisées) pour évaluer une ou des compétences.
5. *L'évaluation des apprentissages doit favoriser le rôle actif de l'élève dans les activités d'évaluation en cours d'apprentissage, augmentant ainsi sa responsabilisation.* La seconde orientation ne signifie pas que l'enseignant est le seul responsable de l'évaluation. Sauf quand il s'agit de sanction des apprentissages, l'évaluation est en effet une responsabilité partagée qui engage l'élève. Celui-ci doit apprendre à porter des jugements sur son propre apprentissage. L'auto-évaluation et l'évaluation par les pairs doivent être vues comme des moyens de rendre l'élève actif dans son apprentissage et de développer des habiletés métacognitives.
6. *L'évaluation des apprentissages doit s'effectuer dans un contexte de collaboration entre différents partenaires tout en tenant compte de leurs responsabilités*

propres. L'élève et l'enseignant ne sont pas les seuls concernés par l'évaluation. La Politique se pose comme un complément aux encadrements réglementaires qui confie des responsabilités à divers acteurs. Ainsi, les directions d'école doivent approuver les modalités d'évaluation proposées par leurs enseignants. Les commissions scolaires doivent collaborer aux épreuves imposées par le ministère de l'Éducation, du Loisir et du Sport (MELS, autrefois le MEQ) qui s'est lui-même engagé à mettre en place des épreuves à la fin de certains cycles. Le MELS est aussi responsable de la sanction des études et de la délivrance des attestations qui en découlent. Par ailleurs, la priorité du MELS de favoriser la création d'équipes au niveau du cycle et de l'école à travers l'implantation de la réforme touche également les questions d'évaluation. Enfin, d'autres intervenants sont amenés à collaborer, qu'il s'agisse des professionnels qui apportent du soutien aux élèves, des entreprises ou organismes qui reçoivent des élèves dans le cadre de programmes à vocation professionnelle ou des parents eux-mêmes, qui doivent être considérés comme des partenaires incontournables.

7. *L'évaluation des apprentissages doit refléter un agir éthique partagé par les différents intervenants.* L'exercice du jugement professionnel de l'enseignant doit être balisé par des règles d'éthique qui vise à assurer le respect des valeurs prônées par la Politique et de règles déontologiques souvent codifiées (Rogers *et al.*, 1994). Ces règles éthiques font l'objet d'un consensus plus ou moins explicite parmi les différents partenaires pour régir les pratiques évaluatives. Ces règles incluent le respect de la confidentialité des résultats de l'évaluation, l'absence de biais pouvant causer de la discrimination, l'interprétation juste des données, etc.
8. *L'évaluation des apprentissages doit contribuer à l'amélioration de la qualité de la langue parlée et écrite de l'élève.* Cette orientation fait écho à un principe directeur du nouveau programme qui fait une place plus large aux compétences langagières tant sur le plan disciplinaire que sur le plan transversal. Dans cet esprit, il s'agit de reconnaître que la responsabilité de la qualité de la langue incombe à l'ensemble des intervenants et de fournir du soutien constant à l'élève. On doit éviter de contaminer les résultats de l'évaluation par des méthodes punitives telle que la soustraction de points.
9. *L'évaluation en vue de la sanction des études doit rendre compte de l'acquisition des compétences et ainsi garantir la valeur sociale des titres officiels.* Cette orientation, tout comme la suivante, s'inscrit dans la fonction de reconnaissance des compétences. La sanction des études fait l'objet d'un encadrement réglementaire résumé dans un guide (MELS, 2013). La maîtrise d'œuvre de la sanction des études appartient au MELS qui fixe les seuils de réussite et est responsable de l'émission des titres officiels – attestations, certificats ou diplômes.
10. *La reconnaissance des acquis doit permettre de reconnaître les compétences d'une personne, indépendamment des conditions de leur acquisition.* Par cette orientation, on admet que l'environnement de l'école n'est pas le seul

milieu où les compétences attendues peuvent être développées. Cela est d'autant plus important que la Politique vise non seulement les apprentissages réalisés dans le cadre des programmes réguliers du primaire et du secondaire, mais couvre aussi les programmes de formation technique du secondaire et les programmes de formation à l'intention des adultes.

Les suites...

Les orientations proposées par la Politique ne sont pas uniques au Québec. Un peu partout en Occident, on observe la promotion de tels principes. Laveault (2009) signale que l'évaluation des apprentissages à travers les réformes des dernières années est caractérisée par l'importance accordée à la régulation des apprentissages à l'aide de l'évaluation formative et par la définition de l'objet à évaluer sous la forme d'une compétence.

Dans le prolongement de la Politique, le MELS a publié différents documents d'accompagnement afin d'outiller les enseignants dans la mise en application des orientations préconisées. Il était prévu que la Politique allait être suivie de cadres de référence. De fait, le cadre de référence pour l'éducation préscolaire et l'enseignement primaire (MEQ, 2002) a paru avant même la Politique. On y définit ce qu'est une compétence et comment l'évaluer. On reprend les principes de la Politique en les illustrant à partir d'exemples pour le préscolaire et le primaire et l'on y définit les principaux termes relatifs à l'évaluation dans un glossaire. Quant au cadre de référence pour le secondaire, la version définitive n'a pas été complétée.

Il faut s'interroger sur ce qui a pu contribuer à retarder autant la parution de la Politique et, par la suite, celle des documents d'accompagnement. Quatre facteurs peuvent expliquer les difficultés qui s'observent dans l'implantation de la politique québécoise d'évaluation des apprentissages.

1) *La viabilité du concept de compétence en milieu scolaire*

Crahay (2006) note que la dimension intégrative de la compétence permettrait de faire le lien entre les connaissances construites à l'école et les savoirs mobilisés dans l'action, mais que cette attente est de fait difficilement comblée. Au Québec, le débat sur la question s'est polarisé autour de la distinction entre l'évaluation des connaissances et l'évaluation des compétences. Il s'agit en quelque sorte d'un faux débat, puisque le développement d'une compétence implique l'acquisition de connaissances. Pourtant, les programmes de formation proposés à l'origine par le MEQ faisaient peu de place aux connaissances, alimentant ainsi la perception qu'elles avaient été évacuées. C'est dans ce contexte qu'ont été entrepris les travaux autour de la progression des apprentissages. À l'origine prévue pour le programme de français langue d'enseignement, la progression des apprentissages s'est

généralisée à l'ensemble des disciplines¹. Trop peu, trop tard, diront certains : la parution de ces documents complémentaires a vite été interprétée comme un retour massif des connaissances, d'autant mieux accueilli par certains que le concept de compétence était déjà perçu comme l'héritage d'une vision utilitariste de l'éducation. Les documents sur la progression des apprentissages ont été reçus comme une invitation à accorder une place prépondérante à l'évaluation des connaissances. De fait, il faut se demander si le concept de compétence, développé, d'une part, dans le domaine de l'apprentissage des langues et, d'autre part, dans le domaine de l'enseignement technique, satisfait les attentes sociales en ce qui a trait au contenu d'un programme qui rassemble toutes les matières du primaire et du secondaire. En ce sens, la question de savoir si le programme mis en œuvre peut effectivement rendre compte des résultats attendus à la fin de la scolarité obligatoire devient secondaire. C'est en effet l'acceptabilité sociale d'un programme par compétence, quel qu'il soit, qui est en cause.

2) *Le passage difficile d'une politique à de nouvelles pratiques*

Lessard, Kamanzi et Larochelle (2008) expliquent que les représentations défavorables d'une politique éducative peuvent se transformer en frein à sa mise en œuvre. De ce point de vue, on peut penser que la mauvaise presse qui a entouré les programmes par compétence a pu contribuer aux difficultés d'implantation de la politique québécoise d'évaluation des apprentissages. L'implantation d'une politique est un processus complexe (Fullan, 2007; Carpentier, 2012) et il est toujours possible que les acteurs la réinterprètent en fonction de leurs représentations et leurs objectifs (Howlett, Ramesh et Perl, 2009). Dans cette perspective, il est important que les enseignants comprennent et s'approprient les éléments d'une politique éducative si l'on veut éviter une prolifération de pratiques divergentes. Or, non seulement la Politique fait la promotion d'un concept de compétence dont on conteste la pertinence en milieu scolaire, mais en plus elle fournit peu d'exemples clairs pour les praticiens qui veulent l'appliquer. Qui plus est, le contexte difficile dans lequel elle a été lancée a empêché la mise en œuvre d'une stratégie d'implantation efficace pour susciter l'adhésion et clarifier les enjeux. Le décalage entre la publication des cadres de référence par rapport au texte de la politique de même que le retard à produire du matériel d'appoint (par exemple, les fiches qui ont été mises en ligne pour aider les enseignants) sont liés à la réception mitigée qu'a connue la politique dans les milieux éducatifs.

1. Les documents sur la progression des apprentissages sont disponibles à l'adresse <http://www1.mels.gouv.qc.ca/progressionPrimaire/> pour le primaire et à l'adresse <http://www1.mels.gouv.qc.ca/progressionSecondaire/> pour le secondaire.

3) *La lourdeur de la tâche demandée aux enseignants*

Pour beaucoup d'enseignants, les attentes associées aux deux fonctions principales de l'évaluation des apprentissages (régulation des apprentissages et reconnaissance des compétences) étaient trop élevées. On leur demandait de concilier divers rôles qui demandent du temps et des efforts supplémentaires et les amènent à adopter des postures parfois inconciliables allant de celle du conseiller et ami à celle du juge et contrôleur (Jorro, 2006). Les enseignants disent manquer de temps pour faire tout ce qu'on attend dans des classes jugées encore trop nombreuses, composées d'élèves qui ont des besoins très diversifiés. Exacerbées par un contexte de relation de travail tendu, il est clair que les exigences nouvelles, à la fois pour pratiquer la différenciation pédagogique et évaluer comme le prescrit la Politique, ont amené une contestation relativement au réalisme des attentes. Ces problèmes ont sans doute été amplifiés par des exigences administratives accrues, une division du travail (surtout au secondaire) qui ne favorise guère la concertation sur le plan de l'évaluation et le manque de matériel.

C'est d'ailleurs dans le but de faciliter la tâche des enseignants que des échelles descriptives ont été proposées (MELS 2007; MELS 2009a, 2009b). Le projet autour des échelles descriptives visait à fournir aux enseignants des descriptions succinctes et compréhensibles de ce que globalement l'on peut observer chez un apprenant typique à divers degrés de l'évolution de chaque compétence des programmes d'études. Ces descriptions ont été mises au point à l'aide d'une méthodologie qui permettait de faire ressortir les consensus au sein de la communauté des enseignants par rapport à des indicateurs caractéristiques de performance en regard des cinq degrés suggérés par les programmes, à partir d'une compétence insuffisante jusqu'à une compétence marquée. Le but poursuivi était de permettre aux enseignants de porter un jugement global sur la compétence afin de positionner l'élève sur le continuum de son développement. Bien que plusieurs enseignants continuent de s'y référer, ces échelles ont été peu à peu mises en veilleuse, avec la diffusion des documents sur la progression des connaissances.

4) *L'alignement entre curriculum et évaluation*

Le retard à publier la Politique tient en partie au fait que le programme de formation avait été établi sans tenir compte des problèmes d'évaluation qui allaient se poser. C'est avec un certain retard qu'on a constaté que des attentes du programme étaient irréalistes ou que les manifestations des compétences n'étaient pas toujours observables. L'idée de concevoir des tâches intégratives a alors été avancée comme solution au problème d'évaluation. Ces tâches doivent d'abord être contextualisées afin de répondre à l'exigence d'authenticité qui assure une certaine validité et motive la personne évaluée dans son exécution. La contextualisation pose un premier problème en ce

qui a trait à la possibilité d'inférer que l'élève pourra exécuter un ensemble de tâches appartenant à la même famille, puisque les éléments contextuels peuvent influencer sur la difficulté de la tâche (Crooks, Kane et Cohen, 1996; Linn, Baker et Dunbar, 1991). Celle-ci doit aussi être complexe. Pour Dierendonck et Fagnant (2010), une tâche est considérée comme complexe dans la mesure où elle « nécessite l'identification, la mobilisation et l'intégration de plus d'une ressource et qu'elle nécessite dès lors une interprétation (ou un cadrage) de la situation et une organisation de la démarche de résolution » (p. 13). La complexité pose un deuxième problème, puisqu'il s'avère parfois impossible de distinguer dans l'exécution l'effet d'une compétence particulière par rapport à l'interaction entre plusieurs compétences. Cela est encore plus difficile si l'on retient l'idée de Carette (2007) d'exiger de la tâche qu'elle permette un certain diagnostic. Certes, une attention particulière à l'élaboration des tâches et à la mise au point d'instruments d'observation pertinents et pratiques peut permettre de faire des inférences justes sur le développement des compétences ou même d'éléments de compétence. Il n'en reste pas moins que certaines compétences peuvent être difficiles à évaluer avec une validité et une fidélité acceptables. Il peut en résulter une réduction de la portée de l'évaluation au point où des pans du programme d'études risquent de passer sous le radar de l'évaluateur (Rey, 2014).

Le bulletin

Aucune étude complète n'est disponible pour déterminer avec précision comment, dans ce contexte, les orientations de la Politique ont été mises en vigueur par les enseignants. Un élément est toutefois venu influencer de façon déterminante leur appropriation par les membres de la communauté éducative. En effet, une modification au Régime pédagogique a imposé, à compter de l'année scolaire 2011-2012, l'usage d'un bulletin unique. De nouvelles modalités d'évaluation ont été établies afin de baliser la confection des bulletins². Ces modalités ont pour but déclaré :

- de réduire le nombre de résultats à consigner et à communiquer dans les bulletins;
- de fournir des critères d'évaluation plus simples et plus uniformes;
- de réduire le temps consacré à l'évaluation et ainsi d'augmenter celui qui est consacré à l'apprentissage.

Les modifications au Régime pédagogique définissent deux types de communications aux parents. D'une part, on trouve des communications dont la forme est déterminée par l'école, mais qui doivent comprendre au moins un résumé des modalités en début d'année et une évaluation préliminaire de la part de l'enseignant avant la mi-octobre. D'autre part, les bulletins doivent suivre le modèle imposé par

2. <http://www.mels.gouv.qc.ca/sections/publications/index.asp?page=fiche&id=1607>

le MELS et leur nombre est réduit, passant de cinq à trois : un premier avant le 20 novembre, un deuxième avant le 15 mars et un troisième avant le 15 juillet.

La réduction du nombre de résultats à consigner tient non seulement à la réduction du nombre de bulletins, mais aussi au fait qu'il n'est plus requis de porter un jugement sur chacune des compétences. Pour ce qui est des matières «secondaires», on donne un résultat global pour la discipline. En ce qui concerne la langue d'enseignement (français ou anglais), la langue seconde (anglais ou français), les mathématiques et les sciences, les compétences ont été reformulées et, dans certains cas, fusionnées. Outre le fait que cela réduit le nombre de jugements à porter, on justifie cette initiative par un souci d'assurer l'intelligibilité des énoncés du bulletin. En contrepartie, l'alignement entre les programmes et l'évaluation n'est plus aussi évident.

Alors que le programme ne précise pas l'importance relative des compétences, le bulletin unique pondère chaque compétence. Le tableau qui suit reproduit les cases à remplir dans la section «Résultats» du nouveau bulletin en ce qui a trait à la langue d'enseignement. Dans ce cas, on retient trois compétences qui ont été formulées le plus simplement possible et qui reçoivent un poids différent. Ainsi, la compétence *Communiquer oralement* contribue au résultat disciplinaire deux fois moins que ne le font *Lire* ou *Écrire*. Le résultat disciplinaire représente le total sur 100 des appréciations pour chacune des compétences. La référence antérieure au seuil de réussite de 60 % est rétablie. Plus encore, on voit réapparaître la moyenne de groupe, héritage d'une tradition qui privilégie une interprétation normative plutôt que de déterminer si les objectifs d'apprentissage ont été atteints. En effet, la moyenne de groupe permet de situer un élève par rapport à l'ensemble du groupe. Il va sans dire que le retour des notes en pourcentage en vue de calculer des moyennes de groupe s'éloigne des orientations de la Politique.

Tableau 1. **La consignation des résultats de l'évaluation au bulletin unique**

Français, langue d'enseignement	Étape 1 20 %	Étape 2 20 %	Étape 3 60 %	Final
Lire 40 %				
Écrire 40 %				
Communiquer oralement 20 %				
Résultat disciplinaire				
Moyenne de groupe				

Le tableau ci-dessus montre aussi que les résultats des trois étapes sont additionnés pour entrer dans le calcul de la note finale de l'école. Même si la dernière étape compte pour 60%, alors que les deux précédentes ne comptent que pour 20%, l'interprétation d'une telle note est problématique dans le contexte d'une évaluation de la compétence (Scallon, 2007). En effet, l'évaluation qui devrait être retenue est celle qui porte sur le moment où la compétence est le plus achevée. La prise en

compte des stades antérieurs ne permet pas de considérer la note comme l'indication du niveau de compétence atteint par l'élève. L'objectif de rendre compte d'une performance scolaire tout au long de l'année et celui d'informer sur le degré de développement d'une compétence paraissent ici incompatibles.

Il faut aussi signaler que le bulletin du secondaire maintient la pratique antérieure de donner un poids égal à la note finale de l'école et à la note obtenue aux épreuves ministérielles à la fin de la 4^e année du secondaire dans certaines matières et à la fin de la 5^e année du secondaire dans d'autres. Par ailleurs, pour le primaire, le MELS a mis au point au cours des dernières années des épreuves en français langue d'enseignement et en mathématiques pour la fin du 2^e cycle (après quatre années de scolarité) et la fin du 3^e cycle (après six années de scolarité). La construction de ces épreuves s'inspire de celle des épreuves du secondaire en ce qu'elle tente d'intégrer les principes d'une évaluation authentique visant à évaluer une compétence dans toute sa complexité et sa globalité. Ces épreuves ministérielles du primaire comptent pour 20 % de la note finale, la proportion restante (80 %) provenant de la note finale de l'école.

Conclusion

Les débats qui se sont tenus autour du bulletin unique ne sont pas sans rappeler ceux qui avaient eu cours vingt-cinq ans plus tôt autour du bulletin descriptif (Lussier, 2001). On évoquait alors le manque de clarté des énoncés, la longueur du bulletin, l'absence de repères quantitatifs et donc « objectifs », l'impossibilité de comparer les élèves entre eux... Du même souffle, sont ignorés les efforts pour éviter le recours à une symbolique numérique trompeuse, pour donner un portrait complet et nuancé du rendement de l'élève et pour le situer en regard des attentes du programme. Dans ce contexte d'un retour vers les pratiques antérieures, une question se pose : A-t-on respecté les valeurs instrumentales mises en avant dans la Politique?

En s'éloignant des contenus des programmes, le bulletin ne respecte plus la cohérence prônée dans le document. Qui plus est, la cohérence devrait aussi se vérifier dans les effets que peuvent avoir les pratiques évaluatives sur les pratiques pédagogiques et, à plus long terme, sur le curriculum lui-même (Frederiksen et Collins, 1989). En effet, l'évaluation n'est jamais un phénomène neutre : il s'établit une interaction entre l'évaluation des apprentissages et l'enseignement-apprentissage, et il est important que les changements induits par la mise en place de dispositifs d'évaluation respectent les objectifs et les orientations des programmes. La valeur de cohérence est de ce point de vue liée aux conséquences des pratiques évaluatives et il faut se demander, d'une part, si la mise en place de nouvelles pratiques conduit à des réinterprétations des programmes et, d'autre part, comment celle-ci façonne nos valeurs éducatives. On voit donc que la valeur de cohérence impose une version élargie au sens où l'entend Messick (2000).

En imposant une arithmétique discutable, le bulletin ne respecte plus la rigueur attendue de la part de professionnels. Quant à la transparence, même si c'est au nom

de cette valeur que le nouveau bulletin a été introduit, on peut se demander si elle ne tient pas uniquement au maintien de références solidement ancrées dans la tradition scolaire. Qui plus est, la Politique avait fait une place à l'équité comme valeur fondamentale pour traduire l'engagement de l'école publique à assurer la réussite pour le plus grand nombre dans la perspective d'une éducation inclusive où la différenciation devient une stratégie privilégiée. En remplaçant la moyenne de groupe dans le bulletin, on peut se demander si la table n'est pas mise pour une évaluation qui favorise la réussite des meilleurs et stigmatise ceux qui éprouvent des difficultés.

Ce qui est certain, c'est que l'évolution récente du dossier de l'évaluation des apprentissages au Québec nous rappelle que l'évaluation n'est pas à l'abri des tensions sociales et qu'en matière d'éducation les décisions politiques heurtent parfois nos idéaux.

Références bibliographiques

- ALLAL, L. (1991). *Vers une pratique de l'évaluation formative : matériel de formation continue des enseignants*. Bruxelles : De Boeck Université.
- ALLAL, L. (2007). Régulations des apprentissages : orientations conceptuelles pour la recherche et la pratique en éducation. Dans L. Allal et L. Mottier Lopez (dir.), *Régulation des apprentissages en situation scolaire et en formation* (p. 7-23). Bruxelles : De Boeck.
- ANDRADE, H. L. (2009). Summing up and moving forward. Key challenges and future directions for research and development in formative assessment. Dans H. L. Andrade et G. J. Cizek (dir.), *Handbook of Formative Assessment* (p. 344-351). New York : Routledge.
- BLACK, P. et WILIAM, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.
- BLACK, P. et WILIAM, D. (2012). Assessment for learning in the classroom. Dans J. Gardner (dir.), *Assessment and Learning* (p. 11-33). Londres : SAGE.
- BLOOM, B. S. (1969). Some theoretical issues relating to educational evaluation. *Educational Evaluation: New Roles, New Means. The 63rd Yearbook of the National Society for the Study of Education* (part 2), 26-50.
- BONNIOL, J.-J. et VIAL, M. (1997). *Les modèles de l'évaluation. Textes fondateurs avec commentaires*. Bruxelles : De Boeck Université.
- CARETTE, V. (2007). L'évaluation au service de la gestion des paradoxes liés à la notion de compétence. *Mesure et évaluation en éducation*, 30(2), 49-71.

- CARPENTIER, A. (2012). Les approches et les stratégies gouvernementales de mise en œuvre des politiques éducatives. *Éducation et francophonie*, 40(1), 12-31.
- CONSEIL SUPÉRIEUR DE L'ÉDUCATION (1982). L'évaluation des apprentissages : « Ça compte-tu? » Québec : Gouvernement du Québec.
- CRAHAY, M. (2006). Dangers, incertitudes et incomplétude de la logique de la compétence en éducation. *Revue française de pédagogie*, 154 (La construction des politiques d'éducation : de nouveaux rapports entre science et politique), 97-110.
- CROOKS, T., KANE, M. et COHEN, A. (1996). Threats to the valid use of assessment. *Assessment in Education*, 3(3), 265-285.
- DIERENDONCK, C. et FAGNANT, A. (2010). Quelques réflexions autour des épreuves d'évaluation développées dans le cadre de l'approche par compétences. *Bulletin de l'ADMEE-Europe* 1, 5-20.
- FREDERIKSEN, J. R. et COLLINS, R. (1989). A system approach to educational testing. *Educational Research*, 9, 27-42.
- FULLAN, M. (2007). *Leading in a Culture of Change*. San Francisco, CA : Jossey-Bass.
- HOWLETT, M., RAMESH, M. et PERL, A. (2009). *Studying Public Policy. Policy Cycles and Policy Subsystems* (3^e éd.). Canada : Oxford University Press.
- JORRO, A. (2006). L'éthos de l'évaluateur : entre imaginaire et posture. Dans G. Figari et L. Mottier Lopez (dir.), *Recherche sur l'évaluation en éducation* (p. 67-75). Paris : L'Harmattan.
- LAVEAULT, D. (2009). L'évaluation en classe : des politiques aux pratiques. *Mesure et évaluation en éducation*, 32(3), 4-22.
- LESSARD, C., KAMANZI, P. et LAROCHELLE, M. (2008). Les politiques et les pratiques en éducation : un couplage problématique. Une perspective anglo-américaine. *Carrefours de l'éducation*, 1(25), 155-194.
- LINN, R. L., BAKER, E. L. et DUNBAR, S. B. (1991). Complex, performance-based assessment. Expectations and validation criteria. *Educational Researcher*, 20, 15-21.
- LUSSIER, D. (2001). Le bulletin scolaire au carrefour des pressions sociales et des mythes qui l'entourent. *Vie pédagogique*, 120, 35-37.
- MESSICK, S. (2000). Consequences of test interpretation and use. The fusion of validity and values in psychological assessment. Dans R. D. Goffin et E. Helmes (dir.), *Values in Psychological Assessment* (p. 3-20). New York : Springer Science + Business Media.
- MINISTÈRE DE L'ÉDUCATION DU QUÉBEC (1977). *Réaffirmer l'école. Rapport du Groupe de travail sur la réforme du curriculum*. Québec : Gouvernement du Québec.

- MINISTÈRE DE L'ÉDUCATION DU QUÉBEC (1979). *L'école québécoise: énoncé de politique et plan d'action*. Québec: Gouvernement du Québec.
- MINISTÈRE DE L'ÉDUCATION DU QUÉBEC (1981). *Politique générale d'évaluation des apprentissages*. Québec: Gouvernement du Québec.
- MINISTÈRE DE L'ÉDUCATION DU QUÉBEC (2002). *L'évaluation des apprentissages au préscolaire et au primaire*. Québec: Gouvernement du Québec.
- MINISTÈRE DE L'ÉDUCATION DU QUÉBEC (2003). *Politique d'évaluation des apprentissages*. Québec: Gouvernement du Québec.
- MINISTÈRE DE L'ÉDUCATION, DU LOISIR ET DU SPORT (2007). *Échelles des niveaux de compétence: enseignement secondaire, 1^{er} cycle*. Québec: Gouvernement du Québec.
- MINISTÈRE DE L'ÉDUCATION, DU LOISIR ET DU SPORT (2009a). *Échelles des niveaux de compétence: enseignement primaire*. Québec: Gouvernement du Québec.
- MINISTÈRE DE L'ÉDUCATION, DU LOISIR ET DU SPORT (2009b). *Échelles des niveaux de compétence: enseignement secondaire, 2^e cycle (3^e éd.)*. Québec: Gouvernement du Québec.
- MINISTÈRE DE L'ÉDUCATION, DU LOISIR ET DU SPORT (2013). *Guide de gestion de la sanction des études et des épreuves ministérielles*. Québec: Gouvernement du Québec.
- MORRISSETTE, J. (2010). Un panorama de la recherche sur l'évaluation formative des apprentissages. *Mesure et évaluation en éducation*, 33(3), 1-28.
- MOTTIER LOPEZ, L. (2012). *La régulation des apprentissages en classe*. Bruxelles: De Boeck.
- NEWTON, P. (2007). Clarifying the purposes of educational assessment. *Assessment in Education*, 14(2), 149-170.
- PERRENOUD, P. (1998). *L'évaluation des élèves: de la fabrication de l'excellence à la régulation des apprentissages*. Bruxelles: De Boeck.
- RAWLS, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- REY, B. (2014). Compétence et évaluation en milieu scolaire: une relation complexe. Dans C. Dierendonck, E. Loarer et B. Rey (dir.), *L'évaluation des compétences en milieu scolaire et en milieu professionnel* (p. 8-30). Bruxelles: De Boeck.
- ROGERS, T. et al. (1994). *Principes d'équité pour l'évaluation des apprentissages au Canada*. Edmonton, AL: CRAME.
- SCALLON, G. (2007). *L'évaluation des apprentissages dans une approche par compétences (2^e éd.)*. Bruxelles: De Boeck.

- SCRIVEN, M. (1967). The methodology of evaluation. Dans R. W. Tyler, R. M. Gagné et M. Scriven (dir.), *Perspectives of Curriculum Evaluation* (vol. 1, p. 39-83). Chicago, IL : Rand McNally.
- SHEPARD, L. A. (2007). Formative assessment: Caveat emptor. Dans C. A. Dwyer (dir.), *The Future of Assessment: Shaping Teaching and Learning* (p. 279-303). Mahwah, NJ : Lawrence Erlbaum.
- STIGGINS, R. J. (2005). From formative assessment to assessment FOR learning: A path to success in standards-based schools. *Phi Delta Kappan*, 87(4), 324-328.

La politique d'évaluation du rendement en Ontario : un alignement qui se précise dans la persévérance et la durée

Dany LAVEAULT

Université d'Ottawa, Ontario, Canada

Louise BOURGEOIS

Université d'Ottawa, Ontario, Canada

VOLUME XLII : 3 – NUMÉRO SPÉCIAL,
AUTOMNE 2014

Revue scientifique virtuelle publiée par
l'Association canadienne d'éducation de
langue française dont la mission est la
suivante : « Par la réflexion et l'action de
son réseau pancanadien, l'ACELF exerce
son leadership en éducation pour
renforcer la vitalité des communautés
francophones ».

Directrice de la publication

Chantal Lainey, ACELF

Présidente du comité de rédaction

Lucie DeBlois,
Université Laval

Comité de rédaction

Sylvie Blain,
Université de Moncton
Lucie DeBlois,
Université Laval
Nadia Rousseau,
Université du Québec à Trois-Rivières
Paul Ruest,
Université de Saint-Boniface
Mariette Thérberge,
Université d'Ottawa

Directeur général de l'ACELF

Richard Lacombe

Conception graphique et montage

Claude Baillargeon

Responsable du site Internet

Étienne Ferron-Forget

Diffusion Érudit

www.erudit.org

Les textes signés n'engagent que
la responsabilité de leurs auteurs
et auteurs, lesquels en assument
également la révision linguistique.
De plus, afin d'attester leur recevabilité,
au regard des exigences du milieu
universitaire, tous les textes sont
arbitrés, c'est-à-dire soumis à des pairs,
selon une procédure déjà convenue.

La revue *Éducation et francophonie* est
publiée deux fois l'an grâce à
l'appui financier du ministère du
Patrimoine canadien et du Conseil
de recherches en sciences humaines
du Canada.

acelf

ASSOCIATION
CANADIENNE
D'ÉDUCATION DE
LANGUE FRANÇAISE

265, rue de la Couronne, bureau 303
Québec (Québec) G1K 6E1
Téléphone : 418 681-4661
Télécopieur : 418 681-3389
Courriel : info@acelf.ca

Dépôt légal
Bibliothèque et Archives nationales
du Québec
Bibliothèque et Archives du Canada
ISSN 1916-8659 (En ligne)
ISSN 0849-1089 (Imprimé)

Les politiques d'évaluation dans le domaine de l'éducation

Rédacteur invité :

Dany LAVEAULT

Liminaire

1 Les politiques d'évaluation en éducation. Et après?

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

15 La formule actuelle d'évaluation ministérielle au Nouveau-Brunswick. Pour qui? Pourquoi?

Jimmy BOURQUE, Université de Moncton, Nouveau-Brunswick, Canada

Mathieu LANG, Université de Moncton, Nouveau-Brunswick, Canada

Jean LABELLE, Université de Moncton, Nouveau-Brunswick, Canada

31 La politique québécoise d'évaluation des apprentissages et les pratiques évaluatives

Michel LAURIER, Université d'Ottawa, Ontario, Canada

50 La politique d'évaluation du rendement en Ontario : un alignement qui se précise dans la persévérance et la durée

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

Louise BOURGEOIS, Université d'Ottawa, Ontario, Canada

68 Une recherche documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public (1^{re} à 12^e année) des provinces de l'Ouest canadien

Jules ROCQUE, Université de Saint-Boniface, Manitoba, Canada

85 L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds? Les enseignements de l'histoire récente de l'école primaire genevoise

Lucie MOTTIER LOPEZ, Université de Genève, Suisse

102 Comparer les systèmes éducatifs francophones à travers le monde grâce au PISA : pas si simple!

Ariane BAYE, Université de Liège, Belgique

Marc DEMEUSE, Université de Mons, Belgique

Nathanaël FRIANT, Université de Mons, Belgique

La politique d'évaluation du rendement en Ontario : un alignement qui se précise dans la persévérance et la durée

Dany LAVEAULT

Université d'Ottawa, Ontario, Canada

Louise BOURGEOIS

Université d'Ottawa, Ontario, Canada

RÉSUMÉ

Donnant suite au rapport de la Commission royale sur l'éducation (CRÉ), l'Ontario a progressivement implanté les fondements d'une politique d'évaluation qui ont peu changé au cours des vingt dernières années, si ce n'est pour des ajustements visant à en actualiser le contenu et à en supprimer les principaux irritants. Jusqu'en 2010, c'est le curriculum, avec la grille d'évaluation générique, qui a constitué les assises de l'évaluation scolaire jusqu'à la publication définitive de la politique d'évaluation *Faire croître le succès*. Cette politique met l'accent sur le rôle de l'évaluation en tant que soutien d'apprentissage et adopte des positions audacieuses quant au cumul des notes et sur la question de l'évaluation des « habiletés d'apprentissage et habitudes de travail » (EDU, 2010) des élèves. Elle conduit les enseignants à planifier l'évaluation des apprentissages sur le long terme, à échantillonner des preuves d'apprentissage au moyen d'une variété de situations d'évaluation complexes

et authentiques. Le suivi de la progression des élèves au regard de la norme provinciale est assuré par l'Office de la qualité et de la responsabilité en éducation (OQRE). Son mandat consiste à fournir une rétroaction à tous les échelons du système scolaire ontarien sur leur alignement avec les visées des politiques relatives au curriculum obligatoire. À proprement parler, c'est la refonte en profondeur des programmes-cadres en Ontario qui a le plus influencé les pratiques d'évaluation des enseignants et a soutenu une mise en œuvre progressive de la politique d'évaluation actuelle.

ABSTRACT

The performance assessment policy in Ontario – aligned on perseverance and duration

Dany LAVEAULT
University of Ottawa, Ontario, Canada

Louise BOURGEOIS
University of Ottawa, Ontario, Canada

Following the report of the Royal Commission on Education (RCE), Ontario has progressively established the foundations of an assessment policy that have changed little over the past twenty years, except for adjustments to update the policy and remove the main irritants. Until 2010 and the final publication of the assessment policy *Growing Success*, the curriculum and the generic evaluation grid were the foundations of school evaluation. This policy emphasizes the role of evaluation in support of learning and adopts bold positions on the accumulation of marks and the question of assessing students' "learning skills and work habits". It leads teachers to plan learning assessment over the long term, sampling evidence of learning through a variety of complex and authentic evaluation situations. Monitoring student progress in relation to provincial standards is provided by the Education Quality and Accountability Office (EQAO). Its mandate is to provide feedback to all levels of the Ontario school system on their alignment with the objectives of policies related to the compulsory curriculum. Strictly speaking, the in-depth overhaul of framework programs in Ontario has had the most influence on teachers' evaluation practices and has supported a progressive implementation of the current assessment policy.

RESUMEN

Política de evaluación del rendimiento en Ontario: un lineamiento que se precisa con la perseverancia y la duración

Dany LAVEAULT

Universidad de Ottawa, Ontario, Canadá

Louise BOURGEOIS

Universidad de Ottawa, Ontario, Canadá

Como consecuencia del informe de la Comisión Real sobre la Educación (CRE), Ontario implantó progresivamente las bases de una política de evaluación que han cambiado un poco en el curso de los últimos veinte años, salvo algunos ajustes cuyo objetivo ha sido actualizar el contenido y suprimir los principales irritantes. Hasta 2010, fue el currículo, y su cuadro de evaluación genérica, que sirvió de base de la evaluación escolar, hasta la publicación definitiva de la política de evaluación *Faire croître le succès*. Dicha política enfatiza el rol de la evaluación en tanto que apoyo al aprendizaje y adopta posturas audaces en lo referente al cúmulo de notas y la cuestión de la evaluación de las «habilidades de aprendizaje y habilidades de trabajo» (EDU, 2010) de los alumnos. Conduce a los maestros a planificar la evaluación de los aprendizajes a largo plazo, a muestrear las pruebas de aprendizajes a partir de una variedad de situaciones de evaluación complejas y auténticas. El seguimiento de la progresión de los alumnos con relación a la norma provincial es asegurado por la Oficina de la Calidad y de la Responsabilidad en Educación (OQRE). Su mandato consiste a retroalimentar a todos los niveles del sistema escolar de Ontario sobre sus lineamientos con los objetivos de las políticas relacionadas con el currículo obligatorio. Propiamente dicho, ha sido la refundición total de los programas referenciales en Ontario lo que ha influenciado las prácticas de evaluación de los maestros y ha apoyado la operacionalización progresiva de la política de evaluación actual.

Introduction

La traduction dans les faits des intentions d'une politique nécessite des objectifs à long terme et des efforts concertés qui s'inscrivent dans la durée. C'est le cas de la politique d'évaluation de l'Ontario qui a maintenu le cap dans ses grandes lignes au cours des vingt dernières années, si ce n'est que pour des ajustements visant à en actualiser le contenu et à chercher une solution à ses principaux irritants.

Le but de cet article est de documenter l'origine de la politique d'évaluation actuelle intitulée *Faire croître le succès* (MEO, 2010), de décrire les difficultés associées

à sa mise en place progressive ainsi que les ajustements apportés depuis le début. Les deux auteurs ont privilégié une approche historique et chronologique fondée sur la consultation de l'ensemble des publications gouvernementales depuis 1994 en rapport avec cette politique. Le premier auteur œuvre dans le milieu universitaire et scolaire ontarien depuis trente ans, alors que le second possède une expérience variée du système, tant dans le milieu scolaire que gouvernemental et, plus récemment, à titre d'étudiante au doctorat.

Les auteurs ont tenté de mettre en relief les effets de cette politique, effets qui ne peuvent cependant être distingués des autres réformes importantes ayant eu cours durant la même période et dont les rapports demeurent anecdotiques et fondés sur l'expérience partagée des deux auteurs. Il a néanmoins été possible de tenir compte du rendement global du système scolaire ontarien – auquel la politique d'évaluation participe – à l'aide des résultats obtenus aux tests provinciaux et internationaux ainsi qu'au moyen d'un échantillon non exhaustif de publications, tant par des chercheurs de l'Ontario que de l'extérieur de cette province, qui se sont penchés sur la question. Enfin, étant donné le caractère récent de cette politique, les réponses aux questions portant sur la réussite de son implantation et sur son impact réel sur les pratiques en salle de classe devront attendre encore un peu.

Origine de la politique actuelle

Pour bien comprendre la politique d'évaluation du rendement des élèves en Ontario, un rappel du contexte historique est indispensable pour en expliquer le contenu, la lente émergence ainsi que sa finalisation tardive en 2010. Il est possible de faire remonter son origine à décembre 1994, alors qu'à l'issue de vastes consultations publiques la Commission royale sur l'éducation (CRÉ) déposait un rapport visant à orienter la réforme de l'éducation aux paliers élémentaire et secondaire. Des changements étaient devenus nécessaires, puisque la dernière réforme d'envergure remonte au rapport du comité provincial Hall et Dennis paru en 1968.

Parmi les recommandations de la CRÉ qui ont été mises en œuvre, mentionnons l'instauration d'un programme d'études secondaires de quatre ans au lieu de cinq, la réforme de l'ensemble des programmes-cadres de la maternelle à la 12^e année, la création de nouveaux organismes provinciaux, notamment l'Ordre des enseignantes et des enseignants de l'Ontario, et la mise sur pied de douze conseils scolaires de district de langue française. Le rapport de la CRÉ formule également onze recommandations relatives à l'évaluation de l'apprentissage qui portent entre autres sur le temps à accorder à la formation des enseignants à l'évaluation formative, du besoin pressant de ressources pédagogiques pour soutenir la fiabilité de la notation, de l'introduction d'un bulletin scolaire uniforme dans toute la province et de la mise sur pied d'un programme de testing à l'échelle de la province auquel toutes les écoles seraient tenues de participer.

Dès 1996, le ministère de l'Éducation met sur pied l'Office de la qualité et de la responsabilité en éducation (OQRE), un organisme indépendant chargé d'élaborer et

de faire passer des examens provinciaux, de communiquer les résultats au Ministère, aux conseils scolaires, aux parents d'élèves et au grand public ainsi que de formuler des recommandations pour soutenir l'amélioration continue du rendement des élèves (OQRE, 1999). Assurément, les résultats inférieurs à la moyenne nationale et internationale des élèves de l'Ontario à la Troisième étude internationale en mathématiques et en sciences (TEIMS) en 1995 ont été l'un des éléments déclencheurs de cette initiative (MEO, 1997).

La réforme du curriculum commence en 1998 avec l'élaboration et la mise en œuvre des programmes-cadres de français et de mathématiques à l'élémentaire et se poursuit jusqu'en 2002, date à laquelle on peut compter plus d'une dizaine de programmes-cadres pour le primaire et plus d'une vingtaine pour le secondaire. Cette période marque également la mise en œuvre d'un bulletin scolaire uniforme de la 1^{re} à la 12^e année (MEO, 1998, 1999b) et l'introduction d'une politique destinée aux écoles secondaires intitulée *Planification des programmes et évaluation* (MEO, 2000a), qui aborde de façon succincte certains aspects de l'évaluation certificative au secondaire, dont la façon de déterminer la note finale du bulletin. À partir de 2003, le Ministère entreprend un cycle continu de révision du curriculum tous les sept ans dans le but de maintenir les programmes-cadres pertinents et à jour (MEO, 2004a). Enfin, pour répondre aux besoins des écoles de langue française en Ontario, le Ministère élabore en 2004 une politique d'aménagement linguistique destinée à soutenir la capacité des écoles de langue française à promouvoir la langue, la culture et la réussite scolaire des élèves en milieu minoritaire (MEO, 2004c).

Émergence progressive de la politique actuelle

Au départ et jusqu'en 2010, c'est le curriculum qui influence le plus les changements dans les pratiques évaluatives sans toutefois faire office de politique d'évaluation. En effet, non seulement les nouveaux programmes-cadres proposent des modifications importantes à ce qui doit être enseigné dans la salle de classe, mais ils entraînent également des changements en matière d'évaluation des apprentissages.

Le curriculum de l'Ontario s'appuie sur deux piliers, des normes de contenu et des normes de performance, qui fournissent un cadre commun sur lequel les enseignants s'appuient pour planifier l'enseignement, l'apprentissage et l'évaluation (MEO, 2010). Les normes de contenu décrivent les connaissances et les habiletés que les élèves doivent maîtriser à la fin de chaque année d'études selon la matière. Les normes de performance, pour leur part, sont décrites dans une grille d'évaluation générique faisant partie intégrante de chaque programme-cadre. Cette grille comporte quatre catégories nommées «compétences» qui s'accompagnent de «critères» évalués au moyen d'échelles descriptives à quatre niveaux de rendement.

En plus du curriculum et de la grille d'évaluation générique, les enseignants de l'Ontario doivent continuer à prendre en considération plusieurs autres publications ministérielles ayant un rapport plus ou moins étroit avec l'évaluation :

- *Des choix qui mènent à l'action* (MEO, 1999a).
- *Guide d'utilisation du bulletin scolaire de l'Ontario* (MEO, 1998, 1999b).
- *Plan d'enseignement individualisé* (PEI), Guide (MEO, 2004b).
- *Plan d'enseignement individualisé : normes pour l'élaboration, la planification des programmes et la mise en œuvre* (MEO, 2000b).
- *Politique d'aménagement linguistique* (MEO, 2004c).

La première publication du Ministère à tenter de chapeauter ce faisceau de publications en matière d'évaluation est sans doute le document *Comment et pourquoi évaluer?* (MEO, 2006a), qui vise d'abord et avant tout à réunir les renseignements provenant de plusieurs sources et à outiller les enseignants de l'élémentaire, notamment en cherchant à illustrer de bonnes pratiques visant à adapter la grille d'évaluation générique du rendement selon différentes tâches d'apprentissage. Ce document appuie ce qui tient alors lieu de politique d'évaluation, mais n'apporte aucun changement. Il s'agit d'un ouvrage didactique et non pas d'une politique d'évaluation coordonnée. Une véritable politique, *Faire croître le succès*, mise en place en 2010, intègre et remplace les directives de l'ensemble des publications antérieures du ministère de l'Éducation se rapportant à l'évaluation. Par ailleurs, l'armature distinctive et centrale de la politique actuelle est la grille d'évaluation générique.

Grille d'évaluation générique

Le but poursuivi par la grille d'évaluation est de fournir un cadre et des repères communs à tous les enseignants de la province pour soutenir leur jugement d'évaluation et uniformiser le processus de notation (MEO, 1997). La grille d'évaluation constitue le cadre général devant servir à évaluer et à noter les productions des élèves dans toutes les matières de la 1^{re} à la 12^e année. Elle comprend quatre catégories nommées «compétences» : connaissance et compréhension, habiletés de la pensée, mise en application et communication. Ces «compétences» ne sont pas sans rappeler, sauf pour la dernière, la taxonomie des objectifs cognitifs de Bloom (1956) et, à cet égard, le choix du terme «compétence» est douteux. Enfin, ces catégories de rendement, définies au moyen de critères d'évaluation, sont évaluées selon quatre échelons appelés «niveaux de rendement». Selon cette échelle, le niveau 3 correspond à la norme provinciale et indique que l'élève est préparé adéquatement pour l'année d'études ou le cours suivant, tandis que le niveau 1 «dénote un rendement très inférieur à la norme provinciale», le niveau 2, «un rendement qui se rapproche de la norme provinciale» et le niveau 4, un rendement «supérieur à la norme provinciale» (MEO, 2010, p. 26).

Pour les enseignants de l'Ontario, la grille d'évaluation a transformé les exigences en matière d'évaluation, par exemple dans la façon d'interpréter les informations sur l'apprentissage. Bien que sa mise en application ait suscité et suscite encore des difficultés importantes en salle de classe (p. ex. fiabilité quant au choix des catégories), l'usage d'une telle grille s'appuie sur un grand nombre de recherches (p. ex.

Andrade, 2000; Arter et McTighe, 2001; Wiggins, 1998) qui démontrent que l'usage de telles échelles descriptives offre plus de précision et une rétroaction plus efficace à l'élève qu'une note chiffrée.

Le système de notation fondé sur la grille d'évaluation contraint les enseignants à adapter les critères d'évaluation et les descripteurs de la grille d'évaluation générique pour que celle-ci s'ajuste à la tâche assignée aux élèves (MEO, 1997). Comme la validité des notes issues de cette grille d'évaluation adaptée dépend de l'articulation de critères d'évaluation pertinents, du choix de descripteurs appropriés et du jugement d'évaluation de l'enseignant (Forgette-Giroux et Simon, 2004; Tierney et Simon, 2004; Wiggins, 1998), le ministère de l'Éducation entreprend l'élaboration de ressources pédagogiques pour former les enseignants à bien utiliser ces grilles. À partir de 1999, il publie des copies types, des exemples de travaux d'élèves représentatifs de chacun des niveaux de la grille d'évaluation, pour accompagner la plupart des programmes-cadres afin d'assurer une plus grande uniformité de l'évaluation et de la notation à travers la province.

Les copies types ne sont qu'un exemple parmi une panoplie de ressources pédagogiques qui ont eu pour effet de submerger les enseignants. On estime que le Ministère aurait produit depuis 1997 plus d'un millier de documents pédagogiques pour renforcer la capacité du personnel enseignant à soutenir l'apprentissage des élèves (Simbagoye, 2007). Par ailleurs, à cette époque, le Secrétariat de la littératie et de la numératie du ministère de l'Éducation est le maître d'œuvre d'un très grand nombre de ces publications. Parmi celles-ci, mentionnons la série de guides d'enseignement efficace en communication orale, en écriture, en lecture et en mathématiques (p. ex. MEO, 2008a, 2006b).

Copies types

Utilisées conjointement avec une grille d'évaluation adaptée, les copies types servent de guide pour soutenir la stabilité du jugement d'évaluation (Simon et Forgette-Giroux, 2001). Bien que de prime abord les enseignants puissent éprouver de la difficulté à respecter les dimensions de la grille d'évaluation en intégrant, par exemple, des critères personnels (Pomplun, Capps et Sundbye, 1998), le travail en collaboration avec des pairs au moyen des copies types est fortement encouragé pour amener les enseignants à partager une représentation commune des exigences du programme-cadre et de la norme provinciale et, par le fait même, soutenir leur compétence à formuler des jugements d'évaluation fiables (Klenowski et Adie, 2009). Par ailleurs, l'usage des copies types avec les parents et les élèves a pour fonction d'aider ceux-ci à mieux comprendre les critères d'évaluation et les attentes (MEO, 1999c; Laveault, Miles et Bourgeois, 2006), tandis que l'analyse des copies types avec les élèves peut également les amener à prendre conscience de leurs progrès, à s'auto-évaluer de façon plus juste et à réguler leurs apprentissages (Simon et Forgette-Giroux, 2001; Wiggins, 1998). Avec le renouvellement du curriculum tous les sept ans, toutefois, ces ressources deviennent « périmées et en contradiction avec les politiques

et pratiques en vigueur» (MEO, 2012, p. 1), si bien que le Ministère prend la décision de retirer les documents de copies types de son site Web et des sites connexes.

Le recours initial aux copies types a permis de prendre conscience de toute l'importance de se doter de moyens adéquats pour décrire les quatre niveaux de rendement de manière fiable. En effet, les niveaux de rendement ontariens ne reposent pas sur de véritables standards provinciaux qui décriraient à quels acquis ils doivent correspondre dans chaque matière et pour chaque année d'études. C'est ainsi que l'interprétation de la norme provinciale (niveau de rendement 3) et des autres normes repose en grande partie sur l'analyse que l'enseignant fait des attentes et des exigences du curriculum. L'évaluation au moyen des niveaux de rendement, qui pouvait être soutenue initialement par les copies types provinciales, repose désormais en grande partie sur le jugement professionnel de l'enseignant et sur des pratiques volontaires d'harmonisation avec d'autres enseignants (MEO, 2010), évoquées mais peu développées dans la nouvelle politique.

Harmonisation de l'évaluation

Pour renforcer la fiabilité du jugement d'évaluation, la voie proposée revient fréquemment à la collaboration entre enseignants. Le ministère de l'Éducation utilise le terme « harmonisation de l'évaluation » (MEO, 2007) pour décrire ce processus de collaboration qui implique un rapprochement entre les jugements d'évaluation des enseignants et la recherche de consensus. Wyatt-Smith, Klenowski et Gunn (2010), Klenowski et Adie (2009) ainsi que Black, Harrison, Hodgen, Marshall et Serret (2010, 2011) ont démontré que ce processus permet d'acquérir une représentation commune des critères d'évaluation, de limiter les biais qui peuvent affecter le jugement d'évaluation et, par le fait même, soutient la compétence à formuler des jugements d'évaluation fiables. Ce travail d'alignement contraint les enseignants à partager leur interprétation du programme-cadre et à échanger sur leurs niveaux d'exigence, ce qui peut aussi contribuer à développer leurs compétences professionnelles en matière d'évaluation.

Alors que le processus d'harmonisation de l'évaluation est conçu pour soutenir l'évaluation sommative (p. ex. le jugement d'évaluation), le ministère de l'Éducation a également élaboré le Parcours fondamental d'enseignement et d'apprentissage (PFEA; MEO, 2008b), un processus qui guide la collaboration entre pairs tout au long de la démarche d'évaluation aux fins de la régulation de l'enseignement et de l'apprentissage (Allal et Mottier Lopez, 2014). Ce processus intègre l'harmonisation de l'évaluation, mais sa visée principale est la décision d'évaluation, notamment la discussion et la mise à l'essai de stratégies de régulation pour soutenir l'apprentissage des élèves (Fullan, Hill et Crévola, 2006). Par conséquent, le PFEA joue deux rôles : réduire les variations de jugements d'évaluation entre enseignants et renforcer la qualité des interventions choisies pour soutenir l'apprentissage des élèves. C'est grâce à ce dialogue collégial que les enseignants en viennent à remettre en question certains aspects de leurs pratiques pédagogiques et à les améliorer.

Ce travail en communauté d'apprentissage (Fullan, 2001) qui permet d'unir les compétences de plusieurs enseignants peut se révéler une stratégie efficace pour soutenir le perfectionnement professionnel en matière d'évaluation lorsqu'il est mandaté par les politiques. Alors que certaines juridictions (p. ex. en Australie et en Nouvelle-Zélande) rendent obligatoire ce travail en communauté d'apprentissage, celui-ci demeure facultatif en Ontario. De telles conditions volontaires ne sont pas toujours facilement réalisables, notamment dans le cas de petits établissements en régions éloignées où l'enseignement est la responsabilité d'un seul titulaire par année d'études lorsqu'il ne s'agit pas de classes à niveaux multiples. De telles conditions d'évaluation, fort répandues dans le système d'éducation franco-ontarien, font en sorte que l'évaluation des niveaux de rendement au moyen de pratiques d'harmonisation soulève d'importants défis dans l'élaboration de grilles adaptées correctement alignées sur le curriculum.

Impacts de la politique d'évaluation actuelle

Selon la politique d'évaluation actuelle de l'Ontario, «le but premier de toute évaluation et de la communication du rendement est d'améliorer l'apprentissage de l'élève» (MEO, 2010, p. 6). L'évaluation soutien d'apprentissage y occupe donc une place prépondérante et l'un des huit principes directeurs de la politique actuelle insiste pour que l'enseignant utilise des pratiques qui «développent la capacité de l'élève à s'autoévaluer, à se fixer des objectifs d'apprentissage personnels et à déterminer les prochaines étapes», ce qui, à proprement parler, n'est autre que la capacité à réguler ses apprentissages (MEO, 2010, p. 6).

Considérant ces objectifs, on peut s'interroger sur les raisons pour lesquelles le ministère de l'Éducation a choisi de se départir des copies types si celles-ci jouaient un rôle utile dans la communication des résultats aux parents et aux élèves et contribuaient à développer chez les élèves les capacités à s'autoréguler et, chez les enseignants, la fiabilité de leur jugement d'évaluation. La révision des programmes-cadres tous les sept ans est sans doute l'un des facteurs importants qui expliquent ce choix. Déjà, plusieurs des copies types s'avéraient obsolètes après quelques années et n'étaient plus utilisées comme points d'ancrage des niveaux de rendement après la première ronde de révisions. Si la publication des copies types avait pu servir initialement d'instrument de perfectionnement professionnel et de modération sociale, elle pouvait s'avérer onéreuse ou insuffisante à long terme dans le contexte de la révision continue des programmes-cadres.

Il est important de noter que l'abandon du projet provincial de copies types ne signifie pas que celles-ci ne doivent plus être utilisées localement. Elles ont toujours un rôle crucial à jouer dans la communication aux élèves et à leurs parents, dans l'harmonisation des jugements d'évaluation entre enseignants et comme soutien pour l'élève au développement de l'habileté à s'autoévaluer. Il revient aux établissements et aux enseignants de se doter de copies types même si cette pratique n'est pas

toujours facile dans le contexte des petits établissements et qu'elle peut donner lieu à des variations importantes entre ceux-ci.

L'utilisation de la grille d'évaluation générique, si elle peut soulever des questions de fiabilité quant aux niveaux de rendement (1 à 4) ou quant au choix des catégories de rendement, nommées « compétences », a cependant l'avantage de contraindre les enseignants à diversifier leurs modalités d'évaluation sommative en les incitant à prendre en considération trois des « compétences » et tout le spectre des quatre niveaux de rendement dans l'élaboration de tâches et de situations d'évaluation (MEO, 2006a). Cela a pour effet qu'une évaluation sommative ne peut se fonder que sur l'évaluation des connaissances ou uniquement sur la mise en application de concepts. Une telle directive encourage la conception de tâches d'évaluation complexes permettant d'évaluer l'acquisition des connaissances et habiletés sur un continuum de progression ainsi que plusieurs aspects de la performance à la fois, notamment la capacité de l'élève à intégrer différentes notions de la matière. Une autre conséquence de cette directive est d'obliger les enseignants à planifier l'échantillonnage de leur prise de renseignements sur l'apprentissage de sorte que toutes les « compétences » soient évaluées au cours de l'année à plusieurs reprises et dans une variété de contextes.

De telles stratégies diversifiées d'évaluation favorisent une « planification à rebours » de l'évaluation (MEO, 2007). Le choix des tâches que l'enseignant utilisera pour évaluer les apprentissages se fait en parallèle avec la sélection des « compétences » qui seront évaluées dans chaque tâche. Chaque « compétence » devra avoir été évaluée par plusieurs tâches et chaque tâche devra évaluer au moins trois « compétences » (MEO, 2006a). Se voyant imposer de telles restrictions dans l'utilisation de grilles adaptées, les enseignants sont forcés de concevoir des situations d'évaluation riches et authentiques et de planifier l'évaluation sommative à l'avance.

Si la diversification des stratégies et des outils d'évaluation favorise un meilleur échantillonnage des preuves d'apprentissage des élèves, elle pose cependant de nouvelles difficultés quant au cumul des résultats et à la communication d'une note globale (un niveau de rendement) sous forme de cote ou de note en pourcentage dans le bulletin. La politique d'évaluation fournit à ce sujet plusieurs directives, pas toujours faciles à concilier entre elles dans la pratique. D'une part, la politique préconise une approche globale lorsqu'elle affirme : « Il est entendu que la note ne sera pas déterminée seulement par l'utilisation de calculs mathématiques. L'approche globale est donc à privilégier pour déterminer la note finale, car elle se prête bien à l'évaluation des tâches signifiantes, complexes et authentiques » (MEO, 2010, p. 69). D'autre part, elle introduit des niveaux intermédiaires entre les quatre niveaux de rendement pour accroître la différenciation des résultats : « Sur le bulletin scolaire de la 1^{re} à la 6^e année, on rendra compte du rendement par rapport aux attentes de chaque matière et/ou domaine d'études en attribuant une cote avec un + ou un -, le cas échéant » (MEO, 2010, p. 50). Par exemple, l'élève pourra se voir attribuer des notes intermédiaires, telles que 3+ ou 2-.

Sur le plan de la notation, la politique d'évaluation cherche un équilibre entre le besoin de différencier les acquis à des fins formatives et la nécessité, particulièrement

saillante au secondaire, de noter les résultats en multipliant les niveaux intermédiaires de rendement. Elle fournit un tableau de conversion qui établit la correspondance entre les niveaux de rendement et des cotes jusqu'en 6^e année et des intervalles de pourcentage à partir de la 7^e année (MEO, 2010, p. 51). Ce travail de conversion en pourcentage pose des difficultés de « commutativité », pas toujours bien saisies par les enseignants. Par exemple, un élève qui obtiendrait un niveau de rendement 3+ se verrait accorder une note en pourcentage entre 77 % et 79 %, mais un élève qui obtient 79 % à une tâche mérite-t-il un 3+? Tout dépend de la difficulté de la tâche, de son alignement avec les attentes du programme-cadre et si cette tâche se produit au début ou à la fin d'une période d'apprentissage. C'est sans doute la raison pour laquelle la politique suggère que la note ne soit pas « déterminée seulement par l'utilisation de calculs mathématiques, mais [qu'elle soit] aussi fondée sur le jugement professionnel de l'enseignante ou l'enseignant » (MEO, 2010, p. 51).

La coexistence de deux systèmes de notes (niveau de rendement et pourcentage) vient sans doute de ce qu'il n'existe pas de système d'examen de fin d'études secondaires en Ontario¹. Par conséquent, ce double système de notes vise à rendre plus fiable l'interprétation des exigences au secondaire, mais aussi à accroître la différenciation des résultats au moyen de notes en pourcentage à l'admission aux études postsecondaires. S'il est plus facile d'obtenir des jugements fiables d'évaluation sur une échelle à quatre niveaux, on peut se demander ce qu'il advient de la fidélité et de la validité de ces résultats scolaires une fois ces derniers convertis en pourcentage.

Sur d'autres aspects, la politique d'évaluation a pris deux positions que nous qualifierions d'« audacieuses » en ce qui concerne le cumul des notes. La première consiste à faire porter le résultat de fin d'année sur un échantillon adéquat, représentatif et récent des preuves d'apprentissage. Cela a pour effet d'accorder une pondération nécessairement plus grande aux évaluations de fin d'année et à ne pas grever le résultat final d'un élève qui aurait commencé l'année avec de faibles résultats et qui aurait accompli d'énormes progrès. La seconde consiste à dissocier l'évaluation par rapport aux attentes du curriculum, notamment les connaissances et habiletés rattachées à la matière, de celle des « habiletés d'apprentissage et habitudes de travail », par exemple l'« autonomie », la « fiabilité », l'« esprit de collaboration » et le « sens de l'initiative » (MEO, 2010, p. 67). Ces deux aspects du rendement de l'élève sont notés séparément dans le bulletin scolaire, de sorte que l'évaluation par rapport aux attentes du curriculum ne devrait pas faire intervenir des référents externes, tels que les habiletés de collaboration de l'élève ou encore la ponctualité dans la remise des travaux.

Bien qu'elle se fonde sur des principes justes d'évaluation scolaire, la nouvelle politique ne résout pas tous les problèmes liés aux pratiques d'évaluation antérieures et certains irritants persistent, notamment en rapport avec l'obligation faite aux

1. Le Test provincial de compétence linguistique (TPCL) est le seul test provincial dont la réussite est obligatoire pour l'obtention du diplôme d'études secondaires de l'Ontario. Il est administré par l'OQRE et correspond aux attentes du programme de français de 10^e année. Les élèves peuvent reprendre le test jusqu'à la 12^e année s'ils échouent la première fois.

enseignants de ne pas tenir compte de la ponctualité (nommée «fiabilité») dans la notation des travaux. Les enseignants sont en effet confrontés à des stratégies d'élèves qui peuvent pervertir ces beaux principes et les utiliser à leur avantage. C'est ainsi que certains élèves pourront retarder la remise de leurs travaux ou encore remettre un nombre limité de travaux. C'est le cas par exemple d'élèves doués, qui vont faire preuve de procrastination pendant toute l'année et entreprendre un marathon de fin d'année pour améliorer leurs résultats au moment où ceux-ci comptent le plus.

En ce qui concerne ces irritants, la politique ne fait que réaffirmer l'importance pour les conseils scolaires de travailler «en collaboration avec les écoles et leur communauté scolaire afin d'élaborer des directives qui portent sur les travaux d'évaluation de l'apprentissage non remis ou remis en retard» (MEO, 2010, p. 55). Il appartient donc aux conseils scolaires et aux écoles de trouver des stratégies pédagogiques nécessaires pour que les «habiletés d'apprentissage et habitudes de travail» fassent l'objet non seulement d'une notation distincte, mais aussi d'un suivi particulier. Cela oblige les enseignants à faire preuve de créativité et surtout à s'attaquer aux véritables problèmes plutôt que de pénaliser les élèves en s'en prenant à leurs notes, ce qui ne résout rien.

En milieu francophone minoritaire, l'assiduité n'est pas la seule habileté d'apprentissage et habitude de travail à laquelle il faut accorder une attention particulière. Dans le contexte franco-ontarien où l'anglais est souvent la langue de communication entre les élèves ou en milieu familial, il s'avère important d'évaluer si la fréquence de l'utilisation du français à l'école est satisfaisante. Cette habileté d'apprentissage se réfère à l'utilisation du français, ce qui est différent du rendement en français qui est évalué séparément au moyen de la grille d'évaluation générique.

Rôle de l'OQRE

Le suivi de la progression des élèves au regard de la norme provinciale n'est pas assuré directement par le Ministère, mais par l'Office de la qualité et de la responsabilité en éducation (OQRE), également chargé de superviser la participation de l'Ontario aux enquêtes nationales et internationales. Ce rôle ne se limite pas au suivi de la progression des élèves ni à la communication de leur niveau de rendement aux parents et aux enseignants. Les résultats au programme de tests de l'OQRE jouent un rôle important dans le pilotage à tous les niveaux du système éducatif, notamment en servant de base à la production de rapports annuels et de plans d'amélioration, tant au niveau provincial qu'au niveau des conseils scolaires et de chaque établissement.

Tous ne sont pas d'avis que l'OQRE soit fidèle aux recommandations de la CRÉ (Després, Kuhn, Ngirumpatse et Parent, 2013) et certains voudraient voir son mandat révisé. Tout comme dans de nombreux États américains, l'utilisation de tests à grande échelle appelés «tests normalisés» soulève des inquiétudes, entre autres quant à leurs incidences sur l'acquisition de compétences et de connaissances plus

larges ou quant au bien-fondé de la pression qu'ils exerceraient sur les enseignants et sur les élèves. Tout laisse croire, cependant, que ces inquiétudes ne sont pas généralisées. Selon un sondage réalisé par l'Institut d'études pédagogiques de l'Ontario (Hart, 2012), les deux tiers (67 %) des adultes de la population ontarienne qui ont répondu au sondage croient que les résultats du programme de tests de l'OQRE sont utilisés pour améliorer la qualité de l'éducation en Ontario.

Ces inquiétudes doivent aussi être tempérées par le fait que le mandat de l'OQRE ne cherche pas à noter les élèves – puisqu'une majorité de tests ne comptent pas pour l'obtention du diplôme –, mais à évaluer l'efficacité du système d'éducation, dont le rendement des élèves constitue un indicateur important. En ce sens, l'OQRE assure une forme de régulation externe en fournissant au système scolaire ontarien une rétroaction à tous les échelons du système éducatif et particulièrement au regard de son alignement avec les visées des politiques relatives au curriculum. Par exemple, les enseignants de chaque établissement sont à même de vérifier chaque année si le rendement des élèves de 3^e, 6^e, 9^e et 10^e année correspond à leur propre évaluation du rendement et ils peuvent apporter, s'il y a lieu, les correctifs nécessaires à leurs pratiques d'enseignement et d'évaluation avec la cohorte suivante. En ce sens, les résultats aux examens de l'OQRE fournissent aux enseignants une référence concrète aux objectifs des programmes-cadres et à la grille d'évaluation générique visant à améliorer l'alignement de l'enseignement et de l'évaluation sur les politiques ministérielles.

Le suivi des recommandations de la CRÉ semble avoir donné lieu à une amélioration du rendement du système éducatif ontarien, attesté tant par le rendement des élèves de l'Ontario aux résultats des tests de l'OQRE que par ceux des études internationales. Les résultats du Programme international pour le suivi des acquis des élèves (PISA) le confirment encore en 2012 même si ceux-ci ont connu de légères baisses par rapport à l'évaluation de 2003 (OQRE, 2013) : les résultats de l'Ontario ne sont surpassés que par ceux de trois pays en lecture, de sept pays en sciences et de dix pays en mathématiques.

Ces résultats, même s'ils demeurent parmi les meilleurs, masquent cependant des écarts importants entre les résultats des garçons et des filles, les garçons réussissant mieux en mathématiques et les filles en lecture.

Les écarts entre francophones et anglophones sont également inquiétants. Les résultats des élèves francophones sont moins élevés que ceux de leurs compatriotes anglophones dans toutes les matières, particulièrement en lecture et en sciences, mais aussi en mathématiques, et ce, sans que de tels écarts aient été observés au moyen des tests provinciaux de l'OQRE, faute de mise en équivalence des tests en français et en anglais. Il est donc urgent de se pencher sur la raison des écarts observés parfois entre élèves francophones et anglophones que des études trop peu nombreuses ont cherché à expliquer, mis à part celles de Landry et Allard (2002) et de Dénomme (2006). Sans possibilités réelles de comparaison entre les résultats des élèves francophones et ceux des anglophones aux tests provinciaux de l'OQRE, de telles études risquent de ne pas avoir lieu ou de demeurer incomplètes. La mise en équivalence des résultats entre les élèves francophones et les anglophones devrait

donc faire l'objet d'un suivi aussi régulier que celui de la mise en équivalence des résultats des élèves, tant anglophones que francophones, d'une année à l'autre.

Bilan de la politique d'évaluation

À proprement parler, c'est la refonte en profondeur des programmes-cadres en Ontario qui a le plus influencé les pratiques d'évaluation des enseignants et a soutenu une mise en œuvre progressive de la politique d'évaluation actuelle. Tout en prescrivant un certain nombre de paramètres, la politique d'évaluation prévoit la dévolution aux conseils scolaires et aux autorités locales de la responsabilité de la mise en œuvre et de l'adaptation de la politique aux défis et contextes particuliers de chacun. À cette fin, il faudra pouvoir compter sur des professionnels formés adéquatement tant au niveau local que provincial.

Pour faire face à ces nouveaux défis, l'Ontario mettra en place prochainement l'une des dernières recommandations de la CRÉ à ne pas avoir fait l'objet d'un suivi, soit le passage de la formation initiale des enseignants d'une durée d'un an à une durée de deux ans, après l'obtention d'un premier baccalauréat général. Dans le programme actuel de formation des maîtres, c'est moins de trois crédits universitaires qui sont alloués à l'évaluation. Considérant la complexité grandissante de l'évaluation du rendement des élèves tant dans sa visée sommative que formative, il est à souhaiter que cette prolongation de la formation contribuera à consolider les principes de la politique d'évaluation dans la pratique des enseignants. En agissant ainsi, l'Ontario se conformera aux recommandations d'un rapport récent de l'OCDE, selon lequel la formation des enseignants à l'évaluation des apprentissages a un impact positif sur la capacité de ceux-ci à réguler leur enseignement de façon constructive (OCDE, 2013). En effet, il ne suffit pas d'améliorer les politiques d'évaluation, encore faut-il que les compétences des enseignants qui auront à les mettre en pratique connaissent une progression similaire : il n'y aura pas de meilleures évaluations sans de meilleurs évaluateurs, que ce soit en Ontario ou ailleurs.

Références bibliographiques

- ALLAL, L. et MOTTIER LOPEZ, L. (2014). Teachers' professional judgment in the context of collaborative assessment practice. Dans C. Wyatt-Smith et V. Klenowski (dir.), *The Enabling Power of Assessment* (vol. 1, p. 151-165). Londres : Springer.
- ANDRADE, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57, 13-18.

- ARTER, J. et McTIGHE, J. (2001). *Scoring Rubrics in the Classroom: Using Performance Criteria for Assessing and Improving Student Performance*. Thousand Oaks, CA: Corwin Press.
- BLACK, P., HARRISON, C., HODGEN, J., MARSHALL, M. et SERRET, N. (2011). Can teachers' summative assessments produce dependable results and also enhance classroom learning? *Assessment in Education*, 18(4), 451-469.
- BLOOM, B. S. (1956). *Taxonomy of Educational Objectives*. Handbook I: The Cognitive Domain. New York: David McKay.
- COMMISSION ROYALE SUR L'ÉDUCATION (CRÉ) (1994). *Pour l'amour d'apprendre : rapport de la Commission royale sur l'éducation*. Vol. 3, *Les éducateurs, les éducatrices*. Toronto : Imprimeur de la Reine pour l'Ontario.
- DENOMMÉ, F. (2006). *Determinants of Educational Achievement of Francophone Students in Ontario* (thèse de doctorat non publiée). Université de Toronto, Ontario Institute for Studies in Education, Ontario, Canada.
- DESPRÉS, S., KUHN, S., NGIRUMPATSE, P. et PARENT, M.-J. (2013). *Responsabilité réelle ou illusion de réussite? Un appel à une révision des tests normalisés en Ontario*. Action Canada : Rapport d'un groupe de travail. Récupéré le 29 mai 2014 de http://www.actioncanada.ca/fr/wp-content/uploads/2013/02/TF3-Report_Standardized-Testing_FR.pdf
- FORGETTE-GIROUX, R. et SIMON, M. (2004). Échelle descriptive d'appréciation du rendement de l'élève. Acquis et problématiques. *McGill Journal of Education*, 39(3), 241-259.
- FULLAN, M. (2001). *Leading in a Culture of Change*. San Francisco : Jossey-Bass.
- FULLAN, M., HILL, P. et CRÉVOLA, C. (2006). *Breakthrough*. Thousand Oaks, CA: Corwin.
- HALL, E. M. et DENNIS, L. A. (1968). *Vivre et s'instruire. Le rapport du Comité provincial sur les buts et les objectifs de l'éducation dans les écoles de l'Ontario*. Toronto : Newton Publishing.
- HART, D. (2012). *Public Attitudes Toward Education in Ontario 2012. The 18th OISE Survey of Educational Issues*. Toronto : University of Toronto, Ontario Institute for Studies in Education. Récupéré le 29 mai 2014 de http://www.oise.utoronto.ca/oise/UserFiles/File/OISE%20Survey/18th_OISE_Survey/OISE%20SURVEY%2018.pdf
- KLENOWSKI, V. et ADIE, L. (2009). Moderation as judgment practice: Reconciling system level accountability and local level practice. *Curriculum Perspectives*, 29(1), 10-28.

- LANDRY, R. et ALLARD, R. (2002). *Résultats pancanadiens des élèves francophones en milieu minoritaire aux évaluations du PIRS : variables contextuelles et conséquences pédagogiques*. Moncton : Institut canadien de recherche sur les minorités linguistiques et Centre de recherche et de développement en éducation. Récupéré le 29 mai 2014 de <http://www.cmec.ca/docs/pirs-rapport-technique-voletA.pdf>
- LAVEAULT, D., MILES, C. et BOURGEOIS, L. (2006). L'utilisation de copies types pour valider les acquis. Dans P. Alves, G. Figari, P. Rodrigues et P. Valois (dir.), *L'évaluation des compétences : entre reconnaissance et validation des acquis* (p. 323-334). Bruxelles : De Boeck.
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (1997). *L'excellence en éducation : le plan de la réforme du système d'éducation, une meilleure éducation gage d'un avenir meilleur*. Toronto : Imprimeur de la Reine pour l'Ontario. Récupéré le 29 mai 2014 de <http://www.edu.gov.on.ca/fre/document/brochure/exfr.pdf>
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (1998). *Guide d'utilisation du bulletin scolaire de l'Ontario de la 1^{re} à la 8^e année*. Toronto : Imprimeur de la Reine pour l'Ontario.
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (1999a). *Des choix qui mènent à l'action. Politique régissant le programme d'orientation et de formation au cheminement de carrière dans les écoles élémentaires et secondaires de l'Ontario*. Toronto : Imprimeur de la Reine pour l'Ontario.
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (1999b). *Guide du bulletin scolaire de l'Ontario de la 9^e à la 12^e année*. Toronto : Imprimeur de la Reine pour l'Ontario.
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (1999c). *Le curriculum de l'Ontario. Copies types de la 1^{re} à la 8^e année – Écriture*. Toronto : Imprimeur de la Reine pour l'Ontario.
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2000a). *Le curriculum de l'Ontario, de la 9^e à la 12^e année : Planification des programmes et évaluation*. Toronto : Imprimeur de la Reine pour l'Ontario. Récupéré le 29 mai 2014 de <http://www.edu.gov.on.ca/fre/curriculum/secondary/progplan912curr.pdf>
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2000b). *Plan d'enseignement individualisé. Normes pour l'élaboration, la planification des programmes et la mise en œuvre*. Toronto : Imprimeur de la Reine pour l'Ontario.
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2004a). *Actualités de la mise en œuvre du curriculum. Nouvelles et informations destinées aux éducateurs de l'Ontario*. Bulletin publié par le Partenariat pour la mise en œuvre du curriculum en collaboration avec le ministère de l'Éducation.

- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2004b). *Plan d'enseignement individualisé (PEI) – Guide*. Toronto : Imprimeur de la Reine pour l'Ontario.
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2004c). *Politique d'aménagement linguistique de l'Ontario pour l'éducation en langue française*. Toronto : Imprimeur de la Reine pour l'Ontario. Récupéré le 29 mai 2014 de <http://www.edu.gov.on.ca/fre/document/policy/linguistique/linguistique.pdf>
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2006a). *Comment et pourquoi évaluer. Document d'appui en évaluation de la 1^{re} à la 8^e année*. Toronto : Imprimeur de la Reine pour l'Ontario. Récupéré le 29 mai 2014 de http://portail.cspgno.ca/proffusion/CPE_document.pdf
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2006b). *Guide d'enseignement efficace en matière de littératie, de la 4^e à la 6^e année*. Toronto : Imprimeur de la Reine pour l'Ontario. Récupéré le 29 mai 2014 de http://atelier.on.ca/edu/resources/guides/Fascicule_1-2008.pdf
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2007). *Harmonisation de l'évaluation : une collaboration pour évaluer de façon cohérente et équitable le travail des élèves*. Toronto : Imprimeur de la Reine pour l'Ontario. Récupéré le 29 mai 2014 de http://www.edu.gov.on.ca/fre/literacynumeracy/inspire/research/Teacher_Moderation_fr.pdf
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2008a). *Guide d'enseignement efficace de la communication orale, de la maternelle à la 3^e année*. Toronto : Imprimeur de la Reine pour l'Ontario. Récupéré le 29 mai 2014 de http://atelier.on.ca/edu/resources/guides/GEE_Communication_orale_M_3.pdf
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2008b). *Parcours fondamental d'enseignement et d'apprentissage : un modèle d'enseignement pour les communautés d'apprentissage professionnel*. Toronto : Imprimeur de la Reine pour l'Ontario. Récupéré le 29 mai 2014 de http://www.edu.gov.on.ca/fre/literacynumeracy/inspire/research/teaching_learning_fr.pdf
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (2010). *Faire croître le succès : évaluation et communication du rendement des élèves fréquentant les écoles de l'Ontario*. Toronto : Imprimeur de la Reine pour l'Ontario. Récupéré le 29 mai 2014 de <http://www.edu.gov.on.ca/fre/policyfunding/growSuccessfr.pdf>
- MINISTÈRE DE L'ÉDUCATION DE L'ONTARIO (MEO) (11 janvier 2012). Note de service aux directrices et directeurs de l'éducation. Récupéré le 29 mai 2014 de <http://www.edu.gov.on.ca/fre/policyfunding/memos/jan2012/CourseProfiles2012Fr.pdf>
- ORGANISATION DE COOPÉRATION ET DE DÉVELOPPEMENT ÉCONOMIQUES (OCDE) (2013). *Principaux résultats de l'Enquête PISA 2012. Ce que les élèves de 15 ans savent et ce qu'ils peuvent faire avec ce qu'ils savent*. Paris : OECD Publishing. Récupéré le 29 mai 2014 de <http://www.oecd.org/pisa/keyfindings/PISA-2012-results-overview-FR.pdf>

- ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT (OECD)
(2013). *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, Résumé en français de *OECD Reviews of Evaluation and Assessment in Education*. Paris : OECD Publishing. Récupéré le 29 mai 2014 de http://www.oecd.org/edu/school/Synergies_for_Better_Learning_Résumé.pdf
- OFFICE DE LA QUALITÉ ET DE LA RESPONSABILITÉ EN ÉDUCATION (OQRE).
(1999). *Rapport provincial de l'Ontario sur le rendement : écoles de langue française*. Toronto : Imprimeur de la Reine pour l'Ontario.
- OFFICE DE LA QUALITÉ ET DE LA RESPONSABILITÉ EN ÉDUCATION (OQRE)
(2012). *OQRE : Le programme de tests provinciaux de l'Ontario. Son histoire, son influence*, 1996-2012. Toronto : Imprimeur de la Reine pour l'Ontario.
- OFFICE DE LA QUALITÉ ET DE LA RESPONSABILITÉ EN ÉDUCATION (OQRE)
(2013). *Les résultats des élèves de l'Ontario à l'évaluation internationale du PISA se maintiennent en lecture et en sciences, mais déclinent en mathématiques*. Toronto. Récupéré le 29 mai 2014 de <http://www.newswire.ca/fr/story/1272693/les-resultats-des-eleves-de-l-ontario-a-l-evaluation-internationale-du-pisa-se-maintiennent-en-lecture-et-en-sciences-mais-declinent-en-mathematiques>
- OFFICE DE LA QUALITÉ ET DE LA RESPONSABILITÉ EN ÉDUCATION (OQRE)
(2013). *PISA 2012. Faits saillants des résultats des élèves de l'Ontario*. Toronto : OQRE. Récupéré le 29 mai 2014 de http://www.eqao.com/pdf_f/13/2012_PISA_Highlights_fr.pdf
- POMPLUN, M., CAPPS, L. et SUNDBYE, N. (1998). Criteria teachers use to score performance items. *Educational Assessment*, 5, 95-110.
- SCALLON, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Saint-Laurent : Éditions du renouveau pédagogique.
- SIMBAGOYE, A. (2007). La réforme curriculaire dans les écoles de langue française de l'Ontario. Dans L. Lafortune, M. Ettayebi et Ph. Jonnaert (dir.), *Observer les réformes en éducation* (p. 131-168). Québec : Presses de l'Université du Québec.
- SIMON, M. et FORGETTE-GIROUX, R. (2001). *Notation du rendement des élèves à partir de tâches complexes de performance : recension des écrits*. Rapport soumis au ministère de l'Éducation de l'Ontario.
- TIERNEY, R et SIMON, M. (2004). What's still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research & Evaluation*, 9(2).
- WIGGINS, G. (1998). *Educative Assessment. Designing Assessments to Inform and Improve Student Performance*. San Francisco, CA : Jossey-Bass
- WYATT-SMITH, C. M., KLENOWSKI, V. et GUNN, S. (2010). The centrality of teachers' judgment practice in assessment: A study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*, 17(1) 59-75.

Une recherche documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public (1^{re} à 12^e année) des provinces de l'Ouest canadien

Jules ROCQUE

Université de Saint-Boniface, Manitoba, Canada

VOLUME XLII : 3 – NUMÉRO SPÉCIAL,
AUTOMNE 2014

Revue scientifique virtuelle publiée par l'Association canadienne d'éducation de langue française dont la mission est la suivante : « Par la réflexion et l'action de son réseau pancanadien, l'ACELF exerce son leadership en éducation pour renforcer la vitalité des communautés francophones ».

Directrice de la publication

Chantal Lainey, ACELF

Présidente du comité de rédaction

Lucie DeBlois,
Université Laval

Comité de rédaction

Sylvie Blain,
Université de Moncton
Lucie DeBlois,
Université Laval
Nadia Rousseau,
Université du Québec à Trois-Rivières
Paul Ruest,
Université de Saint-Boniface
Mariette Thérberge,
Université d'Ottawa

Directeur général de l'ACELF

Richard Lacombe

Conception graphique et montage

Claude Baillargeon

Responsable du site Internet

Étienne Ferron-Forget

Diffusion Érudit

www.erudit.org

Les textes signés n'engagent que la responsabilité de leurs auteurs et auteurs, lesquels en assument également la révision linguistique. De plus, afin d'attester leur recevabilité, au regard des exigences du milieu universitaire, tous les textes sont arbitrés, c'est-à-dire soumis à des pairs, selon une procédure déjà convenue.

La revue *Éducation et francophonie* est publiée deux fois l'an grâce à l'appui financier du ministère du Patrimoine canadien et du Conseil de recherches en sciences humaines du Canada.

acelf

ASSOCIATION
CANADIENNE
D'ÉDUCATION DE
LANGUE FRANÇAISE

265, rue de la Couronne, bureau 303
Québec (Québec) G1K 6E1
Téléphone : 418 681-4661
Télécopieur : 418 681-3389
Courriel : info@acelf.ca

Dépôt légal
Bibliothèque et Archives nationales
du Québec
Bibliothèque et Archives du Canada
ISSN 1916-8659 (En ligne)
ISSN 0849-1089 (Imprimé)

Les politiques d'évaluation dans le domaine de l'éducation

Rédacteur invité :

Dany LAVEAULT

Liminaire

1 Les politiques d'évaluation en éducation. Et après?

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

15 La formule actuelle d'évaluation ministérielle au Nouveau-Brunswick. Pour qui? Pourquoi?

Jimmy BOURQUE, Université de Moncton, Nouveau-Brunswick, Canada

Mathieu LANG, Université de Moncton, Nouveau-Brunswick, Canada

Jean LABELLE, Université de Moncton, Nouveau-Brunswick, Canada

31 La politique québécoise d'évaluation des apprentissages et les pratiques évaluatives

Michel LAURIER, Université d'Ottawa, Ontario, Canada

50 La politique d'évaluation du rendement en Ontario : un alignement qui se précise dans la persévérance et la durée

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

Louise BOURGEOIS, Université d'Ottawa, Ontario, Canada

68 Une recherche documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public (1^{re} à 12^e année) des provinces de l'Ouest canadien

Jules ROCQUE, Université de Saint-Boniface, Manitoba, Canada

85 L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds? Les enseignements de l'histoire récente de l'école primaire genevoise

Lucie MOTTIER LOPEZ, Université de Genève, Suisse

102 Comparer les systèmes éducatifs francophones à travers le monde grâce au PISA : pas si simple!

Ariane BAYE, Université de Liège, Belgique

Marc DEMEUSE, Université de Mons, Belgique

Nathanaël FRIANT, Université de Mons, Belgique

Une recherche documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public (1^{re} à 12^e année) des provinces de l'Ouest canadien

Jules ROCQUE

Université de Saint-Boniface, Manitoba, Canada

RÉSUMÉ

Cet article propose une analyse documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public de l'Ouest canadien. Il porte sur le contexte dans lequel ces éléments ont évolué, c'est-à-dire la transparence et la responsabilité caractéristiques de la nouvelle gestion publique (New Public Management – NPM). Les initiatives d'évaluation à grande échelle qui permettent de comparer les résultats des élèves et l'élaboration de plans stratégiques pour guider les principaux acteurs n'en sont que quelques exemples. Nous brossons un portrait nuancé de la notion d'évaluation *au service de* l'apprentissage, d'évaluation *en tant qu'* apprentissage et d'évaluation *de* l'apprentissage. L'analyse documentaire fournit une description des principales politiques et pratiques d'évaluation en vigueur dans l'Ouest, notamment l'évaluation portant sur les écoles dans une

province ainsi que les évaluations nationales et internationales. Les principes directeurs des quatre provinces convergent vers diverses pratiques d'évaluation des apprentissages. L'influence des réformes aux États-Unis et en Angleterre, entre autres, se fait sentir dans chacune des provinces de l'Ouest, plus particulièrement en ce qui concerne les outils d'évaluation et les pratiques de communication des résultats au moment où les dirigeants veulent faire preuve de transparence et de responsabilité envers tous les citoyens.

ABSTRACT

Descriptive study on the policies, practices and guiding principles for assessing student learning in the public school system (Grades 1 to 12) of Western Canada

Jules ROCQUE

University of Saint-Boniface, Manitoba, Canada

This article focusses on the context in which the policies, practices and guiding principles for the assessment of student learning in Western Canada have evolved in terms of the transparency and responsibility characteristic of the New Public Management – NPM. The initiatives of large-scale assessment to allow the comparison of student results and the development of strategic plans to guide the main actors are just a few examples. The concept of evaluation is then qualified with specific focus on evaluation for learning, evaluation as learning and the evaluation of learning. The results provide a description of the main political and practical assessment principles in force in the West, including the assessment of schools – provincial evaluation, diagnostic evaluation and national and international evaluations. The guiding principles for four provinces converge on a variety of learning assessment practices. The impact of reforms at the international level is felt in each of the Western provinces, especially with regard to assessment tools and practices for reporting results, at a time when the leaders wish to demonstrate transparency and responsibility toward all citizens.

RESUMEN

Una investigación descriptiva sobre las políticas, las prácticas y los principios directores de evaluación de los alumnos del sistema escolar público (1^o al 12^o año) en el Oeste canadiense

Jules ROCQUE

Universidad de Saint-Boniface, Manitoba, Canadá

Este artículo aborda el contexto en el que las políticas, las prácticas y los principios que rigen la evaluación de los aprendizajes de los alumnos del Oeste canadiense han evolucionado, es decir, la transparencia y la responsabilidad característica de la nueva gestión pública (New Public Management – NPM). Las iniciativas a gran escala que permiten comparar los resultados de los alumnos y la elaboración de planes estratégicos para guiar a los principales actores no son sino algunos de los ejemplos. El concepto de evaluación es matizado gracias a una atención especial a la evaluación del servicio del aprendizaje, a la evaluación en tanto que aprendizaje y a la evaluación del aprendizaje. Los resultados ofrecen una descripción de las principales políticas y prácticas de evaluación vigentes en el Oeste, en particular la evaluación que se realiza sobre las escuelas: la evaluación provincial, la evaluación diagnóstico y las evaluaciones nacionales e internacionales. Los principios rectores de cuatro provincias convergen hacia diversas prácticas de evaluación de los aprendizajes. El impacto de las reformas a nivel internacional se deja sentir en cada una de las provincias del Oeste, sobre todo en lo que concierne a las herramientas de evaluación y las prácticas de comunicación de los resultados cuando los dirigentes desean demostrar transparencia y responsabilidad a la ciudadanía.

Introduction

Les politiques, les pratiques et les principes de l'évaluation des apprentissages des élèves à l'intérieur du système scolaire demeurent nombreux et variés. L'évaluation sous toutes ses formes cherche à offrir un éclairage aux principaux acteurs du système sur l'enseignement et l'apprentissage leur permettant d'intervenir en proposant des changements aux politiques et aux pratiques. L'objet de cette analyse documentaire porte sur les politiques, les pratiques et les principes directeurs de l'évaluation des apprentissages des élèves du système scolaire public (1^{re} à 12^e année) de l'Ouest canadien (Colombie-Britannique, Alberta, Saskatchewan et Manitoba). En faisant une comparaison des sites Internet et des documents officiels des ministères de l'Éducation de ces provinces, nous avons voulu faire la lumière sur la façon dont les systèmes d'éducation évaluent l'apprentissage des

élèves tout en examinant comment les grandes initiatives de réformes en éducation ont marqué ces politiques et ces pratiques.

Nous situerons d'abord la problématique dans un contexte précis en exposant quelques initiatives de réformes qui ont marqué les politiques publiques tant au niveau canadien qu'international. Puis nous présenterons le cadre conceptuel en portant une attention aux concepts de l'évaluation au service de l'apprentissage (*Assessment for Learning*), de l'évaluation en tant qu'apprentissage (*Assessment as Learning*) et de l'évaluation de l'apprentissage (*Assessment of Learning*). La méthodologie, le protocole et les résultats de la recherche documentaire suivront. L'impact des tendances mondiales sur les politiques, les pratiques et les principes d'évaluation des apprentissages des élèves de l'Ouest canadien fera l'objet de la dernière partie de l'article avant la conclusion.

Problématique et contexte

Au fil des ans, de multiples initiatives ont été lancées par les ministères de l'Éducation afin de trouver les meilleures façons d'améliorer le système scolaire. Gauthier, Michel, Colombo *et al.* (2000) parlent de différentes tendances apparues dès les années 1960, qui révèlent un grand souci social en matière d'éducation. Les principes de la nouvelle gestion publique (New Public Management – NPM) (Hood, 1991; Pollitt et Bouckaert, 2004; OCDE, 2005) comprennent la transparence et l'imputabilité, et le système d'éducation n'y échappe pas. L'ensemble des systèmes scolaires des pays membres de l'OCDE (Organisation de coopération et de développement économiques) subit donc les effets du NPM faisant en sorte que leurs initiatives d'évaluation des élèves et du système scolaire en comportent certaines caractéristiques.

Politiques et pratiques administratives publiques

Responsabilité

Jaafar et Anderson (2007) présentent la notion de responsabilité (accountability) en faisant appel à deux paradigmes : 1) l'EBA (*Economic-Bureaucratic Accountability*) et 2) l'EPA (*Ethical-Professional Accountability*). L'EBA s'inspire du monde des affaires et met l'accent sur la relation entre consommateur et producteur. Les responsabilités des acteurs demeurent clairement indiquées, ces derniers devant rendre des comptes en tant que particuliers à un groupe public, et le premier souci de ce paradigme concerne la performance et l'atteinte des résultats. L'EPA suppose un principe public selon lequel un discours démocratique existe pour le plus grand bien collectif. Ce paradigme met l'accent sur le processus plutôt que sur les résultats, tout en reconnaissant que la responsabilité de la réussite repose entre les mains de l'ensemble des acteurs.

Nous proposons maintenant un aperçu de l'évolution des politiques et des pratiques d'évaluation des apprentissages des élèves qui ont influencé celles du Canada, particulièrement aux États-Unis, en Angleterre et dans des pays francophones d'Europe.

États-Unis

L'émergence au cours des dernières années de politiques et de pratiques de gestion axées sur les résultats dans le système d'éducation s'explique principalement par la présence de certains éléments catalyseurs issus du domaine de l'administration publique. La notion de responsabilité qui s'est manifestée dans les années 1960 aux États-Unis (Crundwell, 2005; Maroy, 2013) a mené à la création d'outils ou de moyens plus formels permettant à l'ensemble des citoyens de connaître l'état de santé d'un système. L'évaluation de la lecture, à titre d'exemple, nous aide à avoir une idée de la qualité de l'enseignement et de l'apprentissage, tant à l'échelle locale (école) qu'à l'échelle régionale (district, division ou commission scolaire), nationale (Canada) ou internationale.

Dans les années 1980, les États-Unis s'inquiètent de la situation économique où les continents concurrents, notamment l'Europe et l'Asie, menacent de les dépasser sur les plans de la technologie et de l'économie. Le rapport *Nation at Risk* décrit la portée de cette menace en prônant d'importantes réformes en éducation pour améliorer les résultats. Ces réformes comprennent l'amélioration du temps d'apprentissage, la standardisation des programmes d'études et le développement d'outils de responsabilité permettant de rendre des comptes (Carnoy et Loeb, 2002; Harris et Herrington, 2006). En 2002, sous l'administration Bush, les États-Unis lancent une initiative, créant par le fait même une politique de responsabilité, en adoptant la loi *No Child Left Behind* (NCLB). C'est l'égalité d'accès à une éducation de qualité pour tous les citoyens qui est alors visée.

Angleterre et pays de Galles

En Europe, c'est du côté de l'Angleterre et du pays de Galles que les politiques de responsabilité voient le jour (Broadfoot, 2000; Dupriez, 2004; Normand et Derouet, 2011). Les citoyens élisent des représentants qui se dotent de moyens de rendre les responsables du système d'éducation redevables. Ils répondent ainsi aux critiques du public par rapport au faible taux de réussite des élèves et aux nombreux défis à relever dans les écoles, malgré les investissements importants en éducation. On assiste alors au lancement de nouvelles initiatives, dont un curriculum national, des évaluations communes pour tous les élèves et des inspections d'établissements. Ces initiatives représenteront des enjeux majeurs pour les écoles quant à leur réputation (résultats publiés dans les journaux) et à leur financement (réduction à la suite de piètres résultats).

Belgique, France et Suisse

Les pays francophones, notamment la Belgique, la France et la Suisse, ont adopté une approche dite réflexive, plutôt que celle de la responsabilité « dure » des pays de l'Amérique du Nord (Maroy, 2013). Les chercheurs de ces pays s'intéressent à différentes perspectives: le « pilotage » (De Landsheere, 1994, cité dans Maroy, p. 19), l'« évaluation des politiques éducatives » ou l'« obligation de résultats » (Demailly, 2003; Maroy, 2013). Ces approches cherchent à remettre entre les mains des principaux décideurs les données qui leur permettent de connaître la situation

de leur établissement, d'y réfléchir et de déterminer les actions à entreprendre afin d'améliorer le système. Elles visent une amélioration de nature pédagogique, une amélioration des pratiques d'enseignement du personnel professionnel grâce à une appropriation des problèmes que rencontrent les élèves. Finalement, c'est une approche diagnostique et formative qui voit le jour, laquelle s'éloigne de la collecte de données pour imposer des sanctions aux particuliers ou aux systèmes moins performants, comme le veut l'approche de responsabilité.

Canada

Certaines politiques éducatives au Canada, dont celles des provinces qui font l'objet d'une étude dans cet article, s'inspirent d'éléments du mouvement de responsabilité des États-Unis (Brassard, 2006 et 2009; Jaafar et Anderson, 2007) et du NPM. Les autorités locales, notamment les commissions scolaires, préparent leur planification stratégique et ont l'obligation de rendre des comptes au public. Elles cherchent à aligner leur plan stratégique avec les priorités établies par le ministère de l'Éducation de leur province.

Après avoir situé le lecteur dans le contexte des grands courants qui ont marqué les politiques et les pratiques dans le secteur de l'éducation en Amérique du Nord et en Europe, nous passons maintenant au cadre conceptuel, précisant quelques notions clés pour délimiter le sujet à l'étude.

Cadre conceptuel

Trois principaux concepts liés à l'évaluation des élèves sont présentés, soit l'évaluation *au service de* l'apprentissage, l'évaluation *en tant qu'*apprentissage et l'évaluation *de* l'apprentissage.

L'évaluation *au service de* l'apprentissage

L'évaluation *au service de* l'apprentissage se poursuit tout au long du processus d'apprentissage de l'élève. C'est l'enseignant qui en est responsable. Cette évaluation permet de voir ce que l'élève comprend, elle est interactive et aide l'enseignant à adapter son enseignement, à déterminer les besoins d'apprentissage spécifiques de l'élève, à sélectionner du matériel approprié et à concevoir des stratégies différenciées pour mieux répondre aux besoins de l'élève. Dans l'évaluation *au service de* l'apprentissage, l'enseignant fournit une rétroaction immédiate à l'élève et motive ce dernier en se servant « d'observations ciblées, de questions, de conversations, de questionnaires, d'évaluations informatisées, de carnets d'apprentissage ou de toute autre méthode susceptible de [lui] donner les renseignements qui [l']aideront dans [son] travail de planification et d'enseignement » (Éducation, Citoyenneté et Jeunesse Manitoba, 2006, p. 31).

L'évaluation *en tant qu'apprentissage*

L'évaluation *en tant qu'apprentissage* place l'élève au cœur du processus d'apprentissage. L'élève réfléchit à son apprentissage, se pose des questions (ex. : Qu'est-ce que je connais sur ce sujet? Quelles sont les stratégies à appliquer pour m'aider à apprendre?) et cherche à faire des ajustements pour mieux comprendre. L'évaluation *en tant qu'apprentissage* « vise à promouvoir de façon explicite la capacité des élèves à devenir les meilleurs évaluateurs de leur propre travail » (Éducation, Citoyenneté et Jeunesse Manitoba, 2006, p. 42). Bien que conçue pour rendre l'élève plus autonome, l'évaluation *en tant qu'apprentissage* exige que l'enseignant aide l'élève à se fixer des objectifs et à veiller à son propre progrès en lui donnant une rétroaction précise. Les autres élèves de la classe, les membres de la famille et de la collectivité contribuent aussi à l'évaluation *en tant qu'apprentissage* en donnant une rétroaction sur le travail de l'élève ou en offrant d'autres informations, différentes perspectives et idées sur le thème à l'étude ou la tâche à réaliser.

L'évaluation *de l'apprentissage*

L'évaluation *de l'apprentissage* cherche à fournir des preuves de la performance des élèves, à confirmer ce qu'ils savent et à voir s'ils ont atteint les objectifs précis des programmes d'études. L'évaluation *de l'apprentissage* sert, entre autres, à communiquer les résultats d'apprentissage des élèves aux principaux partenaires, notamment les parents, les autorités scolaires et même les futurs employeurs. C'est précisément ce type d'évaluation qui illustre le principe de responsabilité (*accountability*) dont il a été question plus haut. Les décideurs politiques s'inspirent souvent de ces résultats pour élaborer des initiatives de réformes. En plus de cette dimension publique de l'évaluation *de l'apprentissage*, les enseignants doivent être en mesure d'expliquer et d'interpréter avec transparence les résultats des élèves, d'articuler les objectifs fixés et de décrire le processus utilisé dans l'administration des outils d'évaluation (tests, examens, observations de la performance, etc.).

Après avoir présenté brièvement le cadre conceptuel, nous rappelons l'objectif de l'étude pour ensuite en exposer la méthodologie.

Objectif

L'objectif de cette étude est de faire la lumière sur les politiques, les pratiques et les principes directeurs de l'évaluation des apprentissages des élèves du système scolaire public (1^{re} à 12^e année) de l'Ouest canadien, qui comprend la Colombie-Britannique, l'Alberta, la Saskatchewan et le Manitoba. À cette fin, nous avons eu recours à l'analyse documentaire.

Méthodologie

La recherche documentaire consiste à trouver et à analyser diverses sources documentaires afin de s'informer sur un sujet et de répondre à une question précise. «[L]e chercheur n'exerce aucun contrôle sur la façon dont les documents ont été établis» (Grawitz, 2001, p. 573). Toutefois, le choix des éléments documentaires retenus, principalement des extraits des sites Internet et des documents officiels des ministères de l'Éducation, s'est fait en fonction des critères suivants : la pertinence du contenu par rapport à l'objet à l'étude, une attention particulière portée à des mots clés (évaluation *au service de, en tant que* et *de* l'apprentissage, politiques et pratiques d'évaluation des élèves, fondements et principes directeurs de l'évaluation, examens provinciaux) et le désir de comparer la réalité des quatre provinces de l'Ouest en matière d'évaluation de l'apprentissage des élèves.

Protocole

La collecte des données s'est déroulée au cours des mois d'automne de 2013. En prêtant attention aux critères énoncés ci-dessus, nous avons procédé à une recherche systématique dans les sites Internet des ministères de l'Éducation des quatre provinces de l'Ouest canadien. Nous nous sommes intéressés aux pratiques d'évaluation des apprentissages des élèves de la maternelle à la 12^e année. Nous nous sommes également penchés sur les principes directeurs et les politiques qui guident les pratiques d'évaluation des élèves ainsi que sur les éléments de la NPM soulignés plus bas, notamment la responsabilité et la nécessité de rendre des comptes.

Résultats

Colombie-Britannique¹

Après consultation de deux publications majeures, *The Primary Program: A Framework for Teaching*² et *Résultats scolaires: politiques et pratiques (2000)*³, nous sommes en mesure de voir les grandes lignes des politiques et pratiques en matière d'évaluation des élèves dans cette province. Le ministère de l'Éducation (Ministry of Education) souligne l'importance d'une évaluation systématique qui comprend la collecte d'informations sur les connaissances et les habiletés de chaque élève. Les méthodes et les outils d'évaluation sont nombreux : l'observation, l'autoévaluation par les élèves, les exercices pratiques au quotidien, les échantillons de travail des élèves, les tests, les projets, les portfolios.

La Colombie-Britannique gère un programme d'évaluation qui voit les élèves participer à trois types d'évaluation : 1) l'évaluation des compétences fondamentales

-
1. Pour en connaître davantage sur l'évaluation des élèves en Colombie-Britannique : <http://www.bced.gov.bc.ca/assessment/>
 2. http://www.bced.gov.bc.ca/primary_program/primary_prog.pdf
 3. http://www.bced.gov.bc.ca/classroom_assessment/reporting_student_progress_fr.pdf

(ÉCF); 2) les examens provinciaux en 10^e, 11^e et 12^e année; et 3) les évaluations nationales et internationales, qui feront l'objet d'une analyse plus loin dans l'article.

L'ÉCF, qui se fait annuellement à l'échelle de la province, sert à donner un aperçu de la façon dont les élèves acquièrent les compétences de base en lecture, en écriture et en numératie. Le principal objectif de cette évaluation sommative est d'aider la province, les écoles et les conseils de planification scolaire à évaluer le rendement des élèves et à prévoir les façons de l'améliorer. L'ÉCF est conçue et mise en œuvre par les enseignants, conformément aux compétences et aux normes de rendement énoncées dans les programmes d'études. L'ÉCF est administrée chaque année aux élèves des 4^e et 7^e années dans toutes les écoles publiques financées par le gouvernement de la province.

Les examens provinciaux des 10^e, 11^e et 12^e années font partie intégrante du programme du secondaire dans cette province. Depuis 2011, tous les élèves de la 12^e année passent les examens provinciaux dans les matières suivantes : *Language Arts*, français (langue première et langue seconde – immersion), sciences, mathématiques, sciences humaines et études des Premières Nations de la Colombie-Britannique. Cela reflète un changement de politique, car, auparavant, les élèves pouvaient choisir de passer ou non certains examens de la 12^e année en vue de l'obtention du diplôme. Comme le gouvernement offre dorénavant des bourses d'études aux élèves qui obtiennent les meilleurs résultats pour l'ensemble des programmes d'études de la 12^e année, ces évaluations sommatives sont obligatoires.

Alberta⁴

Le ministère de l'Éducation (Alberta Education) fonde ses politiques et ses pratiques provinciales d'évaluation des élèves sur le rapport de recherche de 2011 *Successful Assessment for Learning Projects*. En parcourant le rapport, nous sommes en mesure de relever certains principes qui guident l'évaluation : la définition claire des attentes; l'emploi de travaux modèles illustrant les critères précis d'évaluation mis à la disposition des élèves et l'autoévaluation entreprise par l'élève lui permettant d'examiner ses forces, ses faiblesses (besoins) et d'identifier des objectifs précis à atteindre pour son apprentissage.

L'Alberta a une vision de l'éducation et de l'évaluation qui repose sur le principe qu'il faut aider les élèves à devenir des penseurs engagés, des citoyens ayant un sens éthique et un esprit entrepreneurial (*Alberta Education – Provincial Testing*, 2014).

Au printemps 2013, le gouvernement de l'Alberta a annoncé qu'afin de mieux harmoniser les nouvelles orientations des programmes d'études (basées sur les compétences) avec les outils d'évaluation, il remplacerait les tests standardisés provinciaux par des outils diagnostiques⁵. Cette annonce semble témoigner d'un changement de cap dans cette province. Les nouvelles évaluations fourniront aux élèves, aux enseignants et aux parents des informations leur permettant de mieux gérer l'apprentissage. Les données recueillies se feront au début de l'année scolaire

4. Pour en connaître davantage sur l'évaluation des élèves en Alberta :

<http://education.alberta.ca/teachers/aisi/themes/assessment.aspx>

5. Pour en savoir davantage : <http://education.alberta.ca/francais/ipr/curriculumfr/evaluation.aspx>

et serviront de référence utile pour tous les acteurs, favorisant davantage l'apprentissage des élèves tout au long de l'année.

Malgré ce changement, les élèves de la 12^e année en Alberta doivent toujours passer des examens en vue de l'obtention du diplôme d'études secondaires en sciences naturelles, en mathématiques, en sciences humaines et dans les cours de langue (anglais et français).

La Saskatchewan⁶

En 2009, le ministère de l'Éducation de la Saskatchewan (Ministry of Education) publie *Saskatchewan Provincial Assessment – Conceptual Frameworks*, qui décrit le cadre conceptuel de l'évaluation en Saskatchewan. En 2010, une nouvelle publication voit le jour : *Renewed Curricula: Understanding Outcomes*. Le « pourquoi » de l'évaluation en vue de l'apprentissage des élèves dans cette province rejoint certains énoncés d'autres provinces de l'Ouest et déjà présentés ci-dessus. Notamment, l'évaluation vise à soutenir et à favoriser davantage l'apprentissage des élèves; elle fournit l'information diagnostique requise pour améliorer les programmes et l'enseignement et non pour mettre en évidence les lacunes ni punir les enseignants en raison des faibles résultats; elle donne un aperçu de ce que l'élève peut réaliser ainsi que ses résultats à un moment donné dans le temps; elle favorise la participation de tous les élèves au processus d'apprentissage, renforce leur confiance et les encourage à continuer à apprendre.

La Saskatchewan a mis en place un programme d'évaluation qui a pour but principal d'améliorer l'apprentissage et la réussite des élèves de la province. Le Programme des évaluations ministérielles pour l'apprentissage⁷ (PEA) comporte cinq principaux objectifs :

- 1) améliorer le niveau d'apprentissage scolaire et la réussite de tous les élèves;
- 2) renforcer la capacité des enseignants, des écoles et des divisions scolaires à utiliser les données pour éclairer la prise de décisions;
- 3) accroître les connaissances des enseignants et de l'administration en matière d'évaluation;
- 4) soutenir le développement de communautés d'apprentissage professionnelles;
- 5) renforcer la capacité des divisions scolaires à rendre des comptes au public.

Le PEA évalue le rendement des élèves dans une matière particulière et à des niveaux scolaires spécifiques tous les deux ans. En plus du PEA, les élèves de la 12^e année subissent les examens ministériels en anglais, mathématiques et sciences – biologie, chimie et physique.

6. Pour en connaître davantage sur l'évaluation des élèves en Saskatchewan : <http://ae.gov.sk.ca/provincial-policy-framework> et <http://www.education.gov.sk.ca/Renewed-Curricula-Understanding-Outcomes>

7. Pour en connaître davantage sur le PEA : <http://www.education.gov.sk.ca/AFL/francais>

Le Manitoba⁸

En 2010, le ministère de l'Éducation (Éducation et enseignement supérieur) a publié un document intitulé *Politique provinciale d'évaluation. Maternelle à 12^e année. Responsabilité, honnêteté et passage ou non au niveau d'études suivant* (Éducation Manitoba, 2010), qui décrit les principes directeurs, les politiques ainsi que les pratiques relatives à l'évaluation qui devaient entrer en vigueur dans les écoles du Manitoba à partir de septembre 2011. Le principal objectif de l'évaluation demeure le soutien et l'amélioration de l'apprentissage des élèves. La province fait la promotion de valeurs (diligence, honnêteté et équité) qui permettent aux enfants de devenir des «adultes responsables» tout en développant leur intégrité et en leur permettant de construire des «collectivités fortes et démocratiques» (Éducation Manitoba, 2010, p. 5). De plus, la politique provinciale d'évaluation vise à aider les élèves à apprendre à faire «eux-mêmes leurs travaux, à respecter les échéanciers et à donner crédit à un auteur lorsqu'ils ont emprunté de ses idées; ils doivent aussi assumer la responsabilité de leurs comportements» (Éducation Manitoba, 2010, p. 6).

Au Manitoba, le Ministère évalue les apprentissages des élèves des 3^e, 7^e et 8^e années. Ces évaluations s'inscrivent dans un processus continu afin de fournir des renseignements aux parents sur les compétences de base de leur enfant en lecture et en notions de calcul au début de la 3^e année. En plus, elles permettent aux enseignants d'exercer leur jugement professionnel en faisant appel à une évaluation diagnostique, d'offrir une rétroaction précise à l'élève, d'engager le dialogue avec les parents sur la performance de l'élève à l'égard des compétences de base⁹.

Au Manitoba, la province administre également à la fin de chaque semestre des tests basés sur les normes en 12^e année pour les matières scolaires suivantes : anglais, français et mathématiques.

Principes directeurs – Synthèse des quatre provinces de l'Ouest

Toutes les provinces de l'Ouest possèdent une documentation qui fait référence à des principes directeurs ou à un cadre philosophique donné sur l'évaluation *au service de, en tant que et de* l'apprentissage des élèves, comme cela a été mentionné précédemment. Le ministère de l'Éducation de la Colombie-Britannique a publié un document intitulé *Résultats scolaires : politiques et pratiques*¹⁰, qui dresse les principes directeurs et les grandes lignes des pratiques en matière d'évaluation. L'Alberta fonde ses politiques et ses pratiques provinciales d'évaluation des élèves sur certains principes qui se trouvent dans le rapport de recherche de 2011 *Successful Assessment for Learning Projects from AISI Cycle 3* (Townsend, Adams et White, 2011). Ce projet fait partie de l'initiative du PARSA¹¹ (Programme d'amélioration du rendement

8. Pour en connaître davantage sur l'évaluation des élèves au Manitoba :

<http://www.edu.gov.mb.ca/m12/eval/index.html>

9. Pour en connaître davantage : http://www.edu.gov.mb.ca/m12/eval/eval_7_8.html

10. http://www.bced.gov.bc.ca/classroom_assessment/reporting_student_progress_fr.pdf

11. Pour consulter le cadre de référence du PARSA : <http://education.alberta.ca/media/677227/questions.pdf>

scolaire) lancée en 1999. La Saskatchewan a publié un document intitulé *Saskatchewan Provincial Assessment – Conceptual Frameworks*¹² en 2009 qui décrit le cadre conceptuel de l'évaluation des élèves. Enfin, le ministère de l'Éducation du Manitoba a fait paraître en 2010 un document intitulé *Politique provinciale d'évaluation. Maternelle à 12^e année. Responsabilité, honnêteté et passage ou non au niveau d'études suivant*¹³ (Éducation Manitoba, 2010) décrivant les principes directeurs, les politiques ainsi que les pratiques relatives à l'évaluation des élèves.

Le tableau 1 donne un aperçu de quelques principes directeurs et leur présence ou non dans les publications des ministères de l'Éducation des quatre provinces de l'Ouest. Les concepts liés à l'évaluation des élèves contenus dans le cadre conceptuel soutiennent certains de ces principes directeurs, comme le précise le tableau.

Nous constatons que toutes les provinces ont des exemples de pratiques d'évaluation :

- *au service de l'apprentissage* : principe *b*, outil diagnostique;
- *de l'apprentissage* : principes *c*, examens provinciaux, et *h* responsabilité.

Deux provinces ont explicité des principes se rapprochant des éléments de l'évaluation *en tant qu'apprentissage*, notamment l'Alberta et le Manitoba : principe *d*, autoévaluation avec objectifs à atteindre.

Tableau 1. **Principes directeurs et pratiques – Évaluation dans les provinces de l'Ouest**

Principes/Provinces (Éléments du cadre conceptuel)	Colombie-Britannique	Alberta	Saskatchewan	Manitoba
a) Diverses évaluations utilisées pour déterminer l'apprentissage des élèves	X	X	X	X
b) Outil d'évaluation des compétences fondamentales et diagnostic à l'échelle de la province (Évaluation <i>au service de l'apprentissage</i>)	4 ^e et 7 ^e années	3 ^e , 6 ^e et 9 ^e années	M à 3 ^e année	3 ^e , 7 ^e et 8 ^e années
c) Examens provinciaux pour déterminer l'obtention du diplôme d'études secondaires (Évaluation <i>de l'apprentissage</i>)	X	X	X	X
d) Présence d'autoévaluation des élèves en précisant des objectifs à atteindre (Évaluation <i>en tant qu'apprentissage</i>)		X		X
e) Orientation/optique entrepreneuriale avec accent mis sur le développement des compétences de l'élève		X		
f) Énoncé sur l'équité et la réussite de tous les élèves	X	X	X	X
g) Dimension de citoyen/citoyenne responsable	X	X	X	X
h) Aspect de responsabilité / imputabilité envers les citoyens et citoyennes de la province (Évaluation <i>en tant qu'apprentissage</i>)	X	X	X	X

Avis aux lecteurs : Toute étude comportant des limites, l'auteur tient à souligner qu'il n'a pas été possible d'analyser toutes les publications ministérielles des quatre provinces de l'Ouest. Il se peut donc que certaines pratiques ou certains principes directeurs en évaluation soient présents dans l'une ou l'autre des provinces même s'ils ne sont pas indiqués dans le tableau 1.

12. Pour en connaître davantage :

<http://www.education.gov.sk.ca/Default.aspx?DN=bcd61715-1d02-423e-87ba-55196b38554e>

13. Pour consulter le cadre de référence du PARSA : <http://education.alberta.ca/media/677227/questions.pdf>

De plus, les quatre provinces de l'Ouest se rejoignent au chapitre des principes directeurs portant sur l'équité et la réussite de tous les élèves (principe f) et la dimension de la citoyenneté (principe g). Ces éléments communs peuvent s'expliquer par le fait que, même si l'éducation est de compétence provinciale dans la Constitution canadienne, les provinces de l'Ouest travaillent souvent en étroite collaboration pour l'élaboration de programmes d'études et la mise en œuvre de stratégies ou d'approches pédagogiques et d'évaluation. Nous pouvons citer le Protocole de l'Ouest et du Nord canadiens (PONC) en mathématiques (2006¹⁴) et en français (2012¹⁵) à titre d'exemple.

Les évaluations nationales et internationales

Les ministères de l'Éducation des quatre provinces de l'Ouest veulent pouvoir comparer le rendement de leurs élèves avec celui de l'ensemble des élèves d'ailleurs, tant sur la scène nationale qu'internationale. En portant une attention particulière aux compétences de base, notamment en mathématiques, en lecture et en sciences, les quatre provinces participent à diverses évaluations. Nous voyons au tableau 2 quelles évaluations nationales et internationales se font dans chacune des provinces de l'Ouest canadien. Ces pratiques illustrent le concept de l'évaluation *de* l'apprentissage tel que présenté dans le cadre conceptuel. Elles permettent de comparer les résultats d'évaluation sommative d'élèves précis à différents niveaux (local, régional, national et international) afin de savoir si les objectifs ciblés dans les programmes d'études sont atteints. Deux des principaux objectifs de ce type d'évaluation sont la communication des résultats au public (responsabilité) et la formulation de politiques et de pratiques d'évaluation.

Tableau 2. **Évaluations nationales et internationales (année en vigueur) / âge / matières – Ouest canadien**

PIRS : Programme d'indicateurs du rendement scolaire
 PPCE : Programme pancanadien d'évaluation
 PIRLS : Programme international de recherche en lecture scolaire
 PISA : Programme international pour le suivi des acquis des élèves
 TIMSS : Trends in International Mathematics and Science Study
 M : mathématiques
 L : lecture
 S : sciences

Province	Évaluations nationales		Évaluations internationales		
	PIRS (1993) 13 et 16 ans M-L-S	PPCE (2007) 13 et 16 ans M-L-S	PIRLS (1995) 9,5 ans L	TIMSS (1995) 9,5 et 13 ans M-S	PISA (1997) 15 ans M-L-S
Colombie-Britannique	✓	✓	✓	✓	✓
Alberta	✓	✓	✓	✓	✓
Saskatchewan	✓	✓	✓		✓
Manitoba	✓	✓	✓		✓

14. Pour en connaître davantage sur le PONC – mathématiques:

http://wncp.ca/media/39903/cadrecommun_06.pdf

15. Pour en connaître davantage sur le PONC – français:

https://www.wncp.ca/media/44901/table_des_mati%C3%A8res.pdf

Sauf pour l'évaluation internationale TIMSS¹⁶, l'ensemble des quatre provinces de l'Ouest participe aux diverses évaluations nationales et internationales. Seules la Colombie-Britannique et l'Alberta¹⁷ participent à l'enquête TIMSS (Trends in International Mathematics and Science Study), qui est organisée par l'IEA (International Association for the Evaluation of Educational Achievement).

Impact des tendances en éducation sur l'ouest canadien

Les tendances et les différentes initiatives de réforme en éducation sur la scène nationale et internationale ont exercé une influence sur les lois, les politiques, les pratiques et les principes directeurs en matière d'évaluation des élèves dans les quatre provinces de l'Ouest. À titre d'exemple, le NPM se manifeste en Colombie-Britannique avec l'adoption en 2002 des changements à la Loi scolaire¹⁸ (articles 8, 79 et 171) (*Accountability Framework*¹⁹), obligeant ainsi les acteurs du système scolaire à rendre des comptes. L'Alberta adopte une approche de responsabilité en 2006-2007 en modifiant son cadre de financement²⁰. Par le fait même, elle érige trois piliers (la flexibilité, la responsabilité et le financement) pour assurer un financement équitable tout en permettant aux districts scolaires d'identifier leurs priorités locales dans un cadre de responsabilité²¹. La province veut assurer une plus grande transparence et améliorer la communication des résultats des élèves au grand public. La Saskatchewan n'y échappe pas non plus. Le gouvernement crée une direction générale des politiques stratégiques (*Strategic Policy Branch*²²) dont le mandat principal est d'effectuer des recherches et de faire des analyses afin de fournir des conseils et d'assurer la coordination entre tous les secteurs, départements et partenaires de l'éducation, veillant ainsi à l'avancement des priorités gouvernementales et assurant l'atteinte des objectifs fixés. Le Manitoba subit aussi l'influence du NPM. Toutefois, contrairement aux trois autres provinces de l'Ouest, le gouvernement manitobain, élu majoritairement sous la bannière du Nouveau Parti démocratique (NPD) en 1999, 2003, 2007 et 2011, a effectué des changements importants par rapport à l'évaluation sommative des élèves dès 2003 en éliminant les tests basés sur les normes en 3^e année, tout en rendant ceux des 6^e et 9^e années facultatifs. Cela s'éloigne du courant de l'évaluation à grande échelle du NPM (Owens, 2003).

16. Pour en savoir davantage sur l'enquête TIMSS : <http://www.erc.ie/?p=221>

17. La Colombie-Britannique, l'Alberta, l'Ontario et le Québec participent à l'enquête TIMSS en raison de la forte proportion de la population canadienne qui se trouve dans ces provinces, permettant ainsi des analyses comparatives sur la scène internationale (benchmarking jurisdictions) : <http://timss.bc.edu/>.

18. Pour prendre connaissance de la Loi scolaire : <http://www.bced.gov.bc.ca/legislation/schoollaw/revisedstatutescontents.pdf>

19. Pour en savoir davantage sur l'encadrement de la redevabilité : <http://www2.gov.bc.ca/gov/topic.page?id=DCC1C4815C65494E98F53969B67DA3B4&title=Accountability%20Framework>

20. Pour en connaître davantage sur le cadre de financement : <http://education.alberta.ca/media/482898/framework.pdf>

21. Pour en savoir davantage sur ce pilier : <http://education.alberta.ca/admin/funding/accountability.aspx>

22. Pour en connaître davantage : <http://www.education.gov.sk.ca/Policy-Evaluation-Legislative-Services>

Conclusion

Le présent article a voulu examiner l'impact de différents courants nationaux et internationaux sur les politiques, les pratiques et les principes directeurs de l'évaluation des apprentissages des élèves des quatre provinces de l'Ouest. Les responsables de l'éducation, et plus particulièrement les élus, les directions générales, les directions d'école et les enseignants, doivent faire preuve d'une plus grande responsabilité en s'assurant de disposer d'outils variés pour vérifier le progrès des élèves, pour les comparer avec d'autres – tant sur le plan national qu'international – et pour faire preuve de transparence en communiquant les résultats au grand public. Dans un monde où les échanges commerciaux entre nations et la mobilité des travailleurs s'accroissent, chaque pays cherche à avoir un système d'éducation de qualité qui lui permette de former une main-d'œuvre capable de relever les défis d'un marché de plus en plus concurrentiel. L'Ouest canadien n'échappe pas aux tendances mondiales du NPM. Les citoyens exigent que les fonds publics soient gérés judicieusement. Ils veulent que les écoles atteignent leur principal objectif qui est d'inciter chaque élève à s'engager dans le développement de ses compétences, ce qui lui permettra, en tant que citoyen responsable, de contribuer à une société prospère et à une économie durable.

L'évaluation demeure un exercice complexe qui veut suivre le progrès de l'élève tout au long de son parcours. L'enseignant peut ainsi adapter ses pratiques, offrir une rétroaction précise et constante à l'élève tout en maintenant un dialogue avec ce dernier, illustrant ainsi les principes de l'évaluation *au service de* l'apprentissage. L'autoévaluation place l'élève au cœur du processus d'évaluation. Celui-ci réfléchit à son apprentissage, se pose des questions et demeure à l'écoute des gens autour de lui qui lui offrent une rétroaction sur le thème à l'étude et son progrès. C'est l'évaluation *en tant qu'apprentissage*. L'évaluation *de* l'apprentissage cherche à fournir des preuves de la performance des élèves et constitue toujours une dimension essentielle des pratiques d'évaluation. Les futurs employeurs ainsi que les autorités scolaires et les parents veulent pouvoir comparer les élèves avec d'autres. Les enseignants doivent être en mesure d'interpréter les résultats lorsqu'ils procèdent à l'évaluation *de* l'apprentissage.

La recherche documentaire de la présente étude nous a permis de constater la présence d'une variété d'outils en place dans les quatre provinces de l'Ouest ainsi que d'une diversité de politiques et de principes qui guident l'évaluation des élèves. Les outils diagnostiques sont utilisés par l'ensemble des provinces, tout comme les examens provinciaux, pour déterminer l'octroi des diplômes. Toutes les provinces expriment un énoncé sur l'équité et la réussite de tous les élèves. Elles cherchent à rassurer le public: leurs élèves réussissent en se comparant avec les évaluations nationales et internationales et les autorités font preuve de responsabilité et de transparence en faisant connaître les résultats des élèves dans les principales matières, soit les mathématiques, la lecture et les sciences.

Nous terminons en reconnaissant que les limites de cette recherche se situent, entre autres, au niveau de la réflexion sur la valeur éducative des différentes activités

d'évaluation relevées. Il reste aussi à savoir si les politiques, les pratiques et les principes directeurs en matière d'évaluation des élèves dans les provinces de l'Ouest canadien contribuent ou non à améliorer les systèmes d'éducation et l'apprentissage des élèves.

Références bibliographiques

- ALBERTA EDUCATION (2013). *Assessment for Learning*. Récupéré de <http://education.alberta.ca/teachers/aisi/themes/assessment.aspx>
- ALBERTA EDUCATION (2014). *Provincial Testing – Student Learning Assessments*. Récupéré de <http://education.alberta.ca/admin/testing/student-learning-assessments.aspx>
- BRASSARD, A. (2006). L'évolution de la gestion scolaire depuis les années 1950. *Le point en administration scolaire*, 8(4), 22-27.
- BRASSARD, A. (2009). Le projet de loi 88 : ses implications sur l'autonomie de l'établissement et sur les relations entre les échelons du système. *Le point en administration scolaire*, 11(3), 11-14.
- BROADFOOT, P. (2000). Un nouveau mode de régulation dans un système décentralisé : l'État évaluateur. *Revue française de pédagogie*, 130, 43-55.
- CARNOY, M. et LOEB, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305-331.
- CRUNDWELL, R. M. (2005). Alternative strategies for large scale student assessment in Canada. Is value-added assessment one possible answer? *Canadian Journal of Educational Administration and Policy*, 41, 1-21.
- DEMAILLY, L. (2003). L'évaluation de l'action éducative comme apprentissage et négociation. *Revue française de pédagogie*, 142, 115-129.
- DUPRIEZ, V. (2004). La place de l'évaluation comme ressource pour le pilotage des systèmes scolaires : état des lieux en Belgique francophone et en Angleterre. *Les Cahiers de recherche en éducation et formation*, 35, 2-18.
- ÉDUCATION, CITOYENNETÉ ET JEUNESSE MANITOBA (2006). *Repenser l'évaluation en classe en fonction des buts visés : l'évaluation au service de l'apprentissage, l'évaluation en tant qu'apprentissage, l'évaluation de l'apprentissage*, 2^e éd. Winnipeg, MB : Gouvernement du Manitoba.

- ÉDUCATION MANITOBA (2010). *Politique provinciale d'évaluation. Maternelle à 12^e année – Responsabilité, honnêteté et passage ou non au niveau d'études suivant*. Winnipeg, MB : Gouvernement du Manitoba. Récupéré de http://www.edu.gov.mb.ca/m12/frpub/me/politique_m12/docs/politique.pdf
- GAUTHIER, P., MICHEL, A., COLOMBO, G. *et al.* (2000). L'évaluation des systèmes éducatifs. *Revue internationale d'éducation de Sèvres*, 26, 15-129.
- GRAWITZ, M. (2001). *Méthodes des sciences sociales*. Paris : Dalloz.
- HARRIS, D. N. et HERRINGTON, C. D. (2006). Accountability, standards and the growing achievement gap. Lessons for the past half-century. *American Journal of Education*, 112, 209-239.
- HOOD, C. (1991). A public management for all seasons? *Public Administration*, 69, 1-19.
- JAAFAR, S. B. et ANDERSON, S. (2007). Policy Trends in Accountability for Educational Management and Services in Canada. *The Alberta Journal of Educational Research*, 53(2), 207-227.
- MAROY, C. (dir.). (2013). *L'école à l'épreuve de la performance. Les politiques de régulation par les résultats*. Bruxelles : De Boeck.
- MINISTRY OF EDUCATION – BRITISH COLUMBIA (2000). *The Primary Program – A Framework for Teaching*. Victoria, BC : Ministry of Education. Récupéré de http://www.bced.gov.bc.ca/primary_program/primary_prog.pdf
- MINISTRY OF EDUCATION – SASKATCHEWAN (2013). *Plan for 2013-2014*. Récupéré de <http://www.finance.gov.sk.ca/PlanningAndReporting/2013-14/EducationPlan1314.pdf>
- NORMAND, R. et DEROUET, J.-L. (2011). Évaluation, développement professionnel et organisation scolaire. *Revue française de pédagogie*, 174, 5-20.
- OCDE (2005). *Moderniser l'État. La route à suivre*. Paris : OCDE.
- OWENS, D. (2003). *Educational Accountability in Manitoba*. Winnipeg, MB : Frontier Centre for Public Policy. Récupéré de <http://www.policy.ca/policy-directory/Detailed/Educational-Accountability-in-Manitoba-528.html>
- POLLITT, C. et BOUCKAERT, G. (2004). *Public Management Reform. A Comparative Analysis*. Oxford : Oxford University Press.
- TOWNSEND, D., ADAMS, P. et WHITE, R. (2011). *Successful Assessment for Learning Projects From AISI Cycle 3. Rapport de recherche 2010 sur les projets de l'initiative Alberta Initiative for School Improvement*. Lethbridge, AB : University of Lethbridge. Récupéré de http://education.alberta.ca/media/6412192/research_review_assessment_2010.pdf

L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds? Les enseignements de l'histoire récente de l'école primaire genevoise

Lucie MOTTIER LOPEZ

Université de Genève, Suisse

VOLUME XLII : 3 – NUMÉRO SPÉCIAL,
AUTOMNE 2014

Revue scientifique virtuelle publiée par l'Association canadienne d'éducation de langue française dont la mission est la suivante : « Par la réflexion et l'action de son réseau pancanadien, l'ACELF exerce son leadership en éducation pour renforcer la vitalité des communautés francophones ».

Directrice de la publication

Chantal Lainey, ACELF

Présidente du comité de rédaction

Lucie DeBlois,
Université Laval

Comité de rédaction

Sylvie Blain,
Université de Moncton
Lucie DeBlois,
Université Laval
Nadia Rousseau,
Université du Québec à Trois-Rivières
Paul Ruest,
Université de Saint-Boniface
Mariette Thérberge,
Université d'Ottawa

Directeur général de l'ACELF

Richard Lacombe

Conception graphique et montage

Claude Baillargeon

Responsable du site Internet

Étienne Ferron-Forget

Diffusion Érudit

www.erudit.org

Les textes signés n'engagent que la responsabilité de leurs auteurs et auteurs, lesquels en assument également la révision linguistique. De plus, afin d'attester leur recevabilité, au regard des exigences du milieu universitaire, tous les textes sont arbitrés, c'est-à-dire soumis à des pairs, selon une procédure déjà convenue.

La revue *Éducation et francophonie* est publiée deux fois l'an grâce à l'appui financier du ministère du Patrimoine canadien et du Conseil de recherches en sciences humaines du Canada.

acelf

ASSOCIATION
CANADIENNE
D'ÉDUCATION DE
LANGUE FRANÇAISE

265, rue de la Couronne, bureau 303
Québec (Québec) G1K 6E1
Téléphone : 418 681-4661
Télécopieur : 418 681-3389
Courriel : info@acelf.ca

Dépôt légal
Bibliothèque et Archives nationales
du Québec
Bibliothèque et Archives du Canada
ISSN 1916-8659 (En ligne)
ISSN 0849-1089 (Imprimé)

Les politiques d'évaluation dans le domaine de l'éducation

Rédacteur invité :

Dany LAVEAULT

Liminaire

1 Les politiques d'évaluation en éducation. Et après?

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

15 La formule actuelle d'évaluation ministérielle au Nouveau-Brunswick. Pour qui? Pourquoi?

Jimmy BOURQUE, Université de Moncton, Nouveau-Brunswick, Canada

Mathieu LANG, Université de Moncton, Nouveau-Brunswick, Canada

Jean LABELLE, Université de Moncton, Nouveau-Brunswick, Canada

31 La politique québécoise d'évaluation des apprentissages et les pratiques évaluatives

Michel LAURIER, Université d'Ottawa, Ontario, Canada

50 La politique d'évaluation du rendement en Ontario : un alignement qui se précise dans la persévérance et la durée

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

Louise BOURGEOIS, Université d'Ottawa, Ontario, Canada

68 Une recherche documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public (1^{re} à 12^e année) des provinces de l'Ouest canadien

Jules ROCQUE, Université de Saint-Boniface, Manitoba, Canada

85 L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds? Les enseignements de l'histoire récente de l'école primaire genevoise

Lucie MOTTIER LOPEZ, Université de Genève, Suisse

102 Comparer les systèmes éducatifs francophones à travers le monde grâce au PISA : pas si simple!

Ariane BAYE, Université de Liège, Belgique

Marc DEMEUSE, Université de Mons, Belgique

Nathanaël FRIANT, Université de Mons, Belgique

L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds? Les enseignements de l'histoire récente de l'école primaire genevoise

Lucie MOTTIER LOPEZ

Université de Genève, Suisse

RÉSUMÉ

Comment les fonctions formative et certificative de l'évaluation des apprentissages des élèves ont-elles été traitées dans les réformes en Suisse romande? Quels enseignements tirer des expériences? L'article présente un aperçu des réformes dans les cantons romands ainsi que les grandes lignes de la politique actuelle d'harmonisation de la scolarité obligatoire en Suisse. Il expose ensuite les résultats d'une étude documentaire des offres de formation professionnelle continue en évaluation, de 1976 à 2013, destinées aux enseignants de l'école primaire genevoise. Les thématiques privilégiées, leur transformation, voire leur disparition, sont interprétées à la lumière des grandes phases de la réforme de l'école primaire genevoise et de ses visées. Les résultats de l'étude donnent à voir des déséquilibres successifs entre les fonctions formative et certificative de l'évaluation, ainsi que des éléments d'incompréhension et de débats sociaux qui en ont résulté. À la lumière des expériences, il apparaît qu'un enjeu crucial est de penser un modèle d'une évaluation pédagogique intégrative et

systémique parvenant à trouver un équilibre entre ses différentes fonctions qui, toutes, devraient tendre vers une même finalité de justice et de professionnalité.

ABSTRACT

Will educational assessment finally stand on its own two feet? Recent history lessons at the Geneva elementary school

Lucie MOTTIER LOPEZ
University of Geneva, Switzerland

How were formative and summative functions of student learning dealt with in the reforms of French-speaking Switzerland? What lessons were learned from the experience? This article presents an overview of reforms in the French-speaking cantons and an outline of the current harmonization policy for compulsory schooling in Switzerland. It then presents the results of a literature review of offers of professional development evaluation courses from 1976 to 2013 for Geneva's elementary school teachers. The preferred themes and their transformation or disappearance are interpreted in the light of major phases of the Geneva elementary school reform and its aims. The results of the study show successive imbalances between the formative and summative functions of evaluation along with elements of misunderstanding and resulting social debates. In light of experience, it appears that a crucial issue is to come up with an integrative and systemic evaluation model that finds a balance between the various functions, which should be all directed toward the same goal of justice and professionalism.

RESUMEN

¿La evaluación pedagógica va finalmente marchar con los dos pies? La enseñanza de la historia reciente en la escuela primaria ginebrina

Lucie MOTTIER LOPEZ
Universidad de Genève, Suisse

¿Cómo han sido tratadas las funciones formativas y certificativas de la evaluación de los aprendizajes de los alumnos en las reformas en Suiza francófona? ¿Qué lecciones extraer de dichas experiencias? El artículo presenta una idea aproximada de las reformas en los cantones francófonos así como las grandes líneas de la política actual de armonización de la escolaridad obligatoria en Suiza. En seguida expone los resultados en un estudio documental de las propuestas de formación profesional continua en evaluación, de 1976 a 2013, destinadas a los maestros de la escuela primaria ginebrina. Los temas privilegiados, su transformación, véase su desaparición,

se interprètent avec référence à ses grandes phases de la réforme de l'école primaire genevoise et de ses buts. Les résultats de l'étude permettent percevoir les déséquilibres successifs entre les fonctions formatives et certificatives de l'évaluation, ainsi que des éléments de incompréhension et de débats sociaux qui ont provoqué. À la lumière de ces expériences, on constate qu'un défi crucial est concevoir un modèle de l'évaluation pédagogique interactive et systématique capable de réaliser l'équilibre entre ses différentes fonctions qui, ensemble, devraient tendre vers une finalité commune de justice et de professionnalité.

Introduction

Comment les fonctions formatives et certificatives de l'évaluation des apprentissages des élèves ont-elles été traitées dans les réformes scolaires qui ont vu le jour dans les cantons de Suisse romande? Comment comprendre les mouvements de résistance qui ont émergé face à l'évaluation formative et le plébiscite en faveur des notes chiffrées à l'école primaire? Quels enseignements tirer des expériences?

Pour tenter de donner quelques éléments de réponse à ces questions, l'article commence par présenter les grandes lignes du contexte des réformes en Suisse romande depuis les années 1990, puis élargit à la politique actuelle d'harmonisation de la scolarité obligatoire sur le plan helvétique. Avec en arrière-fond ce contexte de politique éducative, la deuxième partie de l'article interroge plus spécifiquement l'évaluation formative et l'évaluation certificative des apprentissages à travers une étude documentaire des offres de formation continue de l'école primaire genevoise entre 1976 et 2013. Le postulat est que la formation continue représente un espace «interface» potentiellement révélateur des enjeux de transformation des systèmes éducatifs et des pratiques pédagogiques. Quelle est la place accordée à l'évaluation des apprentissages des élèves au fil des années, avant, pendant et après la réforme de l'école primaire genevoise? Quelles sont les orientations thématiques privilégiées? Comment les comprendre au regard des enjeux de la réforme genevoise, des débats sociaux qui ont émergé, des décisions politiques?

Les résultats de l'étude documentaire donnent notamment à voir les confusions qui se sont développées entre les évaluations formatives et certificatives des apprentissages des élèves, ainsi que les importants déséquilibres entre ces deux fonctions, pourtant toutes deux constitutives du système scolaire. La troisième partie de l'article débat de l'importance de l'équilibre formatif/certificatif pour permettre au système de fonctionner, y compris vis-à-vis des enjeux sociaux que l'évaluation pédagogique doit assumer. La conclusion de l'article esquisse, à la lumière des expériences, la perspective d'une évaluation intégrative et systémique parvenant à trouver un équilibre entre les différentes fonctions pédagogiques et sociales qui la constituent.

Un mouvement de réformes en évaluation : pratiques préconisées

En Suisse, comme dans d'autres pays, une succession de réformes ont marqué le paysage scolaire depuis plus de vingt ans. Cette partie présente les grandes lignes de ces réformes, d'abord cantonales, afin de mieux comprendre l'arrière-fond historique et politique des pratiques d'évaluation scolaire actuelles. Elle expose ensuite les principaux éléments de la politique éducative, aujourd'hui en Suisse, qui, par le moyen d'un concordat, institue un certain cadre d'harmonisation entre les cantons suisses. Rappelons, en effet, que le fonctionnement fédéraliste helvétique octroie la souveraineté aux 26 cantons, chacun ayant une loi scolaire, un règlement d'application et des directives en matière d'évaluation des apprentissages des élèves qui lui sont propres.

Des projets cantonaux de réformes manifestant une volonté de promouvoir une évaluation formative dans les classes

En matière d'évaluation pédagogique, des mouvements d'innovation se sont manifestés en Suisse romande dès la fin des années 1980, en lien notamment avec les travaux scientifiques issus de la « pédagogie par objectifs » (p. ex. Mager, 1972) et avec ceux relevant de la « pédagogie de la maîtrise » (p. ex. Bloom, Hastings et Madaus, 1971). Les premiers ont mis l'accent sur l'importance de définir des objectifs de formation (ou d'apprentissage) et des critères¹ d'évaluation; les deuxièmes ont distingué une fonction formative de l'évaluation par rapport à l'évaluation sommative². Rappelons, à la suite d'Allal (2008), que l'évaluation formative vise à fournir des informations qui servent à adapter les activités didactiques pour favoriser la progression des élèves vers les objectifs de formation. Quant à l'évaluation sommative, elle a pour fonction de certifier les compétences et les connaissances acquises par l'élève à la fin d'une étape intermédiaire ou finale de formation. Pour notre part, nous parlerons dans ce dernier cas d'évaluation certificative (Mottier Lopez et Laveault, 2008).

En promouvant l'évaluation formative, les autorités scolaires des cantons romands souhaitaient encourager une évaluation plus qualitative, sans passer par l'établissement systématique de notes chiffrées³, c'est-à-dire une évaluation qui se met au service de la régulation des apprentissages des élèves et d'une différenciation pédagogique au sein de la classe. Concomitamment, et de façon mutuellement

-
1. Par opposition à la référence dite normative, qui se fonde sur la comparaison inter-individuelle pour établir la note / le score de l'élève. À noter qu'une confusion « historique » s'observe parfois ici avec ce double mouvement entre l'émergence de l'évaluation formative et celle de la référence critériée. En effet, certains auteurs, encore aujourd'hui, distinguent « évaluation formative » et « évaluation normative », la première désignant la *fonction* de l'évaluation (soutien à l'apprentissage de l'élève) et l'autre, le *cadre de référence* à l'évaluation. Cette confusion malheureuse amène à considérer que toute évaluation qui n'est pas formative serait normative, ce qui n'est évidemment pas le cas : l'évaluation sommative/certificative peut être critériée.
 2. À partir de la proposition de Scriven dans le cadre de l'évaluation de programmes et dispositifs de formation.
 3. En Suisse, la « note » exprime le résultat d'une évaluation certificative sous la forme d'un nombre de 1 à 6 (le 1 étant le plus mauvais résultat, le 6 le meilleur). Dans le canton de Genève, l'usage de demi-notes et le calcul de moyennes au dixième d'un ensemble de notes à l'école primaire ont été réintroduits à l'automne 2013 après l'élection de 2006 dont l'article parle ensuite.

complémentaire, d'autres orientations innovantes ont été envisagées par les politiques éducatives. Les finalités partagées étaient de lutter contre l'échec scolaire et de rendre l'école suisse plus démocratique, moins inégalitaire, notamment à la suite du rapport de Hutmacher (1993) constatant la persistance du redoublement pour des raisons souvent arbitraires (p. ex. la représentation que se fait l'enseignant des exigences supposées du collègue qui prendra en charge ses élèves l'année suivante).

Des réformes restées sans effet réel sur un fonctionnement méritocratique de l'école

Un bilan des réformes qui ont eu lieu dans les cantons de Berne, Genève et Vaud a été établi dans Gilliéron Giroud et Ntamakiliro (2010). Un ensemble de mesures y ont été instituées, dont la suppression des notes chiffrées au cours des premières années de la scolarité obligatoire, l'institutionnalisation de la transmission régulière d'informations aux parents des élèves, la réduction du nombre d'évaluations certificatives pour donner plus de temps et de place à des démarches d'évaluation formative, la prise en compte d'informations complémentaires aux notes chiffrées lors des décisions d'orientation (ou évaluation pronostique; Allal, 2008). Ces décisions s'inscrivent dans une idéologie de « justice corrective » (Crahay, 2013) qui cherche à réduire les inégalités de départ entre les enfants en traitant les différences individuelles par des mesures de discrimination positive. L'évaluation formative s'inscrit pleinement dans cette visée lorsque, sur la base des informations produites, l'enseignant accorde un supplément d'attention aux élèves qui éprouvent des difficultés ou ajuste ses dispositifs didactiques et pédagogiques afin qu'ils soient le mieux adaptés possible aux caractéristiques différentes de ses élèves.

Pourtant, et malgré ces mesures, les systèmes scolaires et de formation en Suisse continuent de fonctionner sur des bases essentiellement méritocratiques : attribution légitime de notes supérieures aux élèves les plus performants ou méritants; orientation précoce des élèves en fin d'école primaire⁴ vers des filières hiérarchisées dans l'enseignement secondaire; accès conditionné aux formations « prestigieuses » par les résultats scolaires, entre autres. Est-ce une des raisons? Toujours est-il que le constat semble sans appel : « Les réformes ambitieuses engagées en Suisse romande, tout comme en Belgique, en France et au Québec, n'ont pas eu les effets escomptés, loin de là » (Gilliéron Giroud et Ntamakiliro, 2010, p. 1). Nombre d'obstacles et de résistances ont surgi, montrant combien l'évaluation scolaire est un point particulièrement sensible des systèmes scolaires aux yeux des acteurs sociaux concernés (parents, enseignants, formateurs, cadres des systèmes éducatifs, employeurs, politiques). Qu'en est-il alors de la situation politique actuelle?

4. À cette période, les élèves ont 12 ans en moyenne à la fin de l'école primaire, excepté dans le canton de Vaud qui a déjà institué ce passage de l'école primaire à l'école secondaire à l'âge de 10 ans en moyenne. Une sélection précoce, donc. Aujourd'hui, tous les cantons romands se sont alignés sur un passage dans le secondaire après huit années d'école primaire obligatoire (deux années anciennement appelées « école enfantine » + six années primaires subséquentes).

Une volonté d'harmonisation de la scolarité obligatoire au niveau fédéral, mais laissant libres les évaluations internes des apprentissages des élèves

Sur un plan intercantonal, un projet politique a été élaboré au début des années 2000 dans le but d'harmoniser les structures de la scolarité obligatoire en Suisse ainsi que les objectifs de formation relatifs à chaque ordre d'enseignement. Il s'agit du «concordat HarmoS⁵». Après une phase de consultation en 2006, le projet a été soumis en 2007 à un vote populaire auprès des 26 cantons suisses. Le concordat est entré en vigueur le 1^{er} août 2009, le nombre minimum de cantons requis ayant été atteint (au moins dix)⁶. Une période transitoire de six ans a été décidée pour sa mise en place.

En adhérant au concordat, les cantons se sont engagés à harmoniser :

- la durée de l'enseignement obligatoire (en tout 11 années) et ses ordres d'enseignement primaire (8 années – 4 à 12 ans) et secondaire inférieur (3 années – 13 à 15 ans). Les deux années de l'école enfantine destinées aux élèves de 4 à 6 ans deviennent obligatoires, faisant désormais partie intégrante de l'enseignement primaire;
- les contenus de standards de formation à des fins de pilotage des systèmes éducatifs suisses. Associées à ces standards, des épreuves externes communes aux régions linguistiques sont envisagées;
- les objectifs de formation et les attentes fondamentales pour le passage d'un ordre d'enseignement à un autre, par le moyen d'un plan d'études commun aux différentes régions linguistiques, respectant les standards de formation définis.

Pour ce qui relève plus spécifiquement de l'évaluation des apprentissages des élèves, le concordat HarmoS délimite les bases d'une politique d'évaluation *externe*, dans une perspective essentiellement de pilotage, en résonance avec la logique des enquêtes internationales, tel le Programme international pour le suivi des acquis des élèves (PISA). Mais pour ce qui est de l'évaluation *interne* des apprentissages des élèves, c'est-à-dire des évaluations conçues par les enseignants ou équipes d'enseignants, le concordat ne prescrit pas une politique commune aux cantons. Ces derniers restent libres de décider ce qui relève, par exemple, des systèmes de notation et usages de moyennes arithmétiques pour déterminer les promotions et orientations des élèves, des degrés scolaires dans lesquels sont introduites les notes chiffrées, des outils formels de l'évaluation certificative/pronostique et sa communication aux parents des élèves et autres partenaires, de la part prise par l'évaluation formative dans les pratiques en salle de classe, de l'usage obligé ou non d'un dossier d'apprentissage ou portfolio, etc.

C'est au regard de ce contexte de réformes cantonales et de ce projet politique sur le plan fédéral que la partie suivante examine la question des pratiques d'évaluation

5. Voir le site officiel: <http://www.edk.ch/dyn/11737.php>

6. Actuellement, 22 cantons ont ratifié l'accord. En 2009, ils étaient au nombre de 11. Voir: <http://www.edk.ch/dyn/12536.php>

en classe, plus particulièrement dans le canton de Genève à l'école primaire, et celle des orientations voulues pour elles.

Les pratiques d'évaluation des enseignants genevois considérées à travers les offres de formation continue

Dans une étude précédente (Mottier Lopez et Tessaro, 2010), à partir d'une analyse approfondie de documents produits dans les années 2000 et suivantes, dans le contexte de l'école primaire genevoise et de la Suisse romande, nous avons constaté le rôle que la *formation professionnelle continue* avait joué dans l'élaboration et l'évolution des cadres de prescription à propos de l'évaluation des apprentissages des élèves. En effet, le contexte de la formation continue s'était avéré être un espace à l'*interface* entre les pratiques d'évaluation effectives des enseignants et les prescriptions contraintes par les choix politiques mais qui, dans une période de réforme et post-réforme, visaient aussi à institutionnaliser de « bonnes pratiques », objectivées et formalisées dans l'interaction entre enseignants et formateurs. Ce mouvement a été possible, car les formateurs étaient appelés par l'institution (plus précisément par la direction de l'enseignement primaire) à contribuer à la rédaction de « directives » qui, au milieu des années 2000, visaient à harmoniser les pratiques d'évaluation des enseignants à la suite d'une période intense de réforme.

Dans le prolongement de ces constats, il nous a paru intéressant de poursuivre l'étude en élargissant l'empan temporel, partant des premières offres de formation continue formalisées dans des documents officiels jusqu'à une période récente. Pour ce faire, nous avons effectué une analyse systématique des axes de toutes les formations continues offertes aux enseignants du primaire dans les « catalogues⁷ » officiels de 1976 à 2013 : organisation structurelle des plans de formation, rubriques de catégorisation des unités de formation, titre de celles-ci et descriptif⁸. Les offres appartenant à ces catalogues se caractérisent par le fait qu'elles sont élaborées par les responsables et formateurs mandatés par le Département de l'instruction publique, au regard des besoins liés à l'introduction d'un moyen d'enseignement (ou manuel scolaire), à une nouvelle orientation pédagogique, à des innovations à promouvoir, à un perfectionnement continu dans un domaine disciplinaire, par exemple. Les formateurs engagés dans la formation continue, dans la période qui nous intéresse car des changements sont en cours, sont le plus souvent des enseignants détachés de leur classe, à titre temporaire et parfois à temps partiel. Ce sont des acteurs « interface » entre les demandes institutionnelles (inscrites dans une politique descendante des offres de formation continue) et les préoccupations du terrain à partir desquelles ils ont la latitude de proposer des contenus de formation (dans une logique d'offres qui devient, dans ce cas, ascendante).

7. Terme utilisé sur le terrain pour désigner le document qui propose les offres de formation pour chaque année scolaire.

8. Nous remercions le Département de l'instruction publique du canton de Genève qui nous a ouvert ses archives, nous permettant ainsi de réaliser cette étude documentaire.

L'étude que nous présentons ci-après porte sur des formations qui ont toutes été dispensées, la plupart une seule fois dans l'année scolaire concernée, à des groupes pouvant varier grosso modo de 10 à 30 enseignants. Nous ne disposons pas des données statistiques informant du nombre exact de participants à chaque unité de formation. Cela ne nuit cependant pas à notre objectif, qui est ici d'exposer une analyse du *contenu thématique* de ces offres de formation au regard des périodes qui ont caractérisé la réforme de l'école primaire genevoise, afin d'en discuter par rapport aux orientations des projets voulus pour les pratiques d'évaluation des apprentissages des élèves.

Les grands axes de la réforme de l'école primaire genevoise

Afin d'effectuer cette mise en perspective, commençons par préciser que la «Rénovation», nom de la réforme de l'école primaire genevoise, a eu lieu entre les années 1994 et 2002 autour de trois axes principaux: 1) individualiser les parcours de formation des élèves en créant des cycles d'apprentissage de plusieurs années; 2) développer des projets d'école et une culture de la collaboration; 3) placer les élèves au cœur de l'action pédagogique (Direction de l'enseignement primaire, 1994). L'évaluation des apprentissages est alors un des axes importants d'innovation en vue de renforcer une évaluation au service des progressions souples et différenciées des élèves, au service aussi de l'information⁹ aux partenaires concernés par les progressions des élèves, et en limitant les étapes de certification et les décisions pronostiques (promotion, redoublement, orientation). La politique d'implantation de la réforme a consisté à engager d'abord une phase d'«exploration intensive» entre 1994 et 1998, qui a concerné une dizaine d'écoles volontaires, puis une phase d'«extension progressive» entre 1999 et 2002.

Avant l'implantation de la Rénovation de l'école primaire genevoise

Il ressort de notre analyse qu'entre 1976 et 1988 aucun contenu n'abordait un axe explicite de formation continue (appelée «perfectionnement» à cette époque) sur la question de l'évaluation pédagogique. C'est dans le catalogue de l'année scolaire 1989-1990 qu'une première proposition apparaît, sous l'intitulé *Livrets scolaires: mises de notes ou appréciation, rédaction de commentaires*. Puis, au cours des quatre années suivantes, juste avant le démarrage de la première phase de réforme, 14 offres en matière d'évaluation sont proposées, montrant l'importance accordée à cet axe dans les années au cours desquelles le projet de réforme s'est élaboré dans différentes commissions et groupes de travail avant sa mise en œuvre effective. Les offres portent sur l'évaluation formative et certificative de façon relativement équilibrée.

9. Ainsi est apparue dans le contexte genevois une «nouvelle» fonction de l'évaluation: l'évaluation dite informative, essentiellement destinée aux parents des élèves et n'entrant pas directement dans une logique de certification des apprentissages réalisés.

Pendant la première phase de la Rénovation : exploration intensive (1994-1998)

Notre analyse de contenu, conjuguée à une appréhension historique, montre d'une façon générale des « pics » en ce qui concerne les offres de formation sur l'évaluation pédagogique. Dans la période d'« exploration intensive » de la Rénovation, l'année 1995-1996 en est un exemple. Le tableau 1 présente les contenus proposés dans le catalogue de cette année, que nous avons catégorisés selon les deux approches générales qui organisent les contenus de formation initiale et continue des enseignants à Genève : l'approche transversale qui désigne les processus qui « traversent » les disciplines scolaires (évaluation, processus d'apprentissage, gestion de classe, métier d'élève, relation famille-école, etc.), et l'approche disciplinaire qui entre par les différentes didactiques des disciplines scolaires.

Tableau 1. Offres de formation professionnelle continue de l'année 1995-1996

Année	Approche transversale (8)	Approche disciplinaire (2)
1995-1996	- Connaissances de base en évaluation - Une même activité pour tous intégrant l'évaluation formative - Vers l'autonomie de l'élève par le développement de l'auto-évaluation - Comportement de l'élève : quel apprentissage, quelle évaluation? - Articulons nos pratiques d'évaluation entre divisions! - Livret scolaire 1P-2P : quelle évaluation de la progression? - Critères de réussite : observation d'élèves estimés en difficulté* - Objectifs 1E-2E : analyse de pratiques*	- Évaluer les apprentis lecteurs (2E-1P) - Évaluer l'orthographe dans les productions personnelles

* Signifie que l'évaluation n'apparaît pas directement dans le titre, mais qu'elle est présentée comme un contenu, parmi d'autres, dans le descriptif de l'unité de formation.
E = école enfantine; P = école primaire.

Ainsi, pour l'année 1995-1996, dix unités sont proposées, certainement révélatrices des besoins en formation professionnelle continue sur l'évaluation pédagogique, soit d'une façon générale (*Connaissances de base en évaluation*), soit au regard des différentes fonctions que les autorités scolaires souhaitent désormais lui faire assumer : en matière de soutien à l'apprentissage des élèves et de leur autonomie; au regard des outils institués de certification, tel le livret scolaire. Les enjeux propres aux ordres d'enseignement (ou « divisions ») se détachent. Seulement deux unités sont plus spécifiquement destinées à une évaluation de contenus disciplinaires en français, l'évaluation des apprentissages étant appréhendée à cette époque essentiellement dans une approche transversale.

Concomitamment à la première phase de la Rénovation, une réorganisation de la formation continue (dès 1996-1997)

Notre analyse documentaire montre que la politique de transformation de l'école primaire genevoise s'accompagne aussi d'une réforme de la formation professionnelle des enseignants en exercice. Appelée « Perfectionnement » de 1976 à 1996, la formation est désormais qualifiée de « continue ». Dans l'avant-propos du catalogue de 1996-1997, le directeur chargé du secteur du personnel s'adresse aux enseignants en titrant son message : « La formation continue : une priorité pour le présent et l'avenir de l'école primaire genevoise. » Il rappelle que « la formation continue anticipe et accompagne l'évolution de l'école et de la profession » et annonce une réorganisation ultérieure à partir d'un nouveau concept de formation continue. Le but annoncé est « la consolidation des compétences des enseignants et la transformation des pratiques de classe [pour permettre] d'augmenter les qualifications des élèves » (Avant-propos du catalogue, 1996-1997).

La nouvelle organisation distingue désormais des projets de formation individuels et collectifs, qui débouchent sur deux modalités principales encore en vigueur actuellement :

1. Des formations collectives soutenant la mise en œuvre d'un « projet d'établissement » (de façon cohérente avec l'axe 2 de la Rénovation; voir plus haut), au regard des besoins exprimés par les enseignants. Quatorze périodes, sur temps scolaire, sont dévolues à cette modalité qui exige une concertation entre un groupe d'enseignants ou tout le corps enseignant d'une même école et les responsables de la formation. Cette modalité s'inscrit donc dans une logique de formation qui se veut ascendante dans la définition des contenus, c'est-à-dire partant des préoccupations du terrain.
2. Des « cours » que chaque enseignant peut choisir dans le « catalogue » revu chaque année scolaire par les responsables et formateurs de la formation continue. Toujours sur temps scolaire, 14 périodes sont également disponibles pour cette modalité de formation, en plus de celles consacrées au projet d'établissement. Les contenus sont définis ici dans la logique présentée plus haut.

On retiendra encore que dès 1996, c'est-à-dire au milieu de la première phase d'implantation de la Rénovation, la formation continue s'organise en « domaines de compétences de référence » associés à des « compétences plus spécifiques ». Les catalogues ne sont plus structurés par des entrées thématiques comme c'était le cas auparavant (exemples de rubriques : « Évaluation », « Pédagogie générale »). Dans la nouvelle organisation, l'évaluation pédagogique est considérée comme une « compétence spécifique » du domaine « Gérer la progression des apprentissages ». On constate que seule l'évaluation formative figure explicitement dans le référentiel de compétences professionnelles. L'évaluation certificative n'est *jamais* nommée, comme si elle ne représentait pas une compétence professionnelle, elle aussi à développer. Nous reviendrons plus loin sur cette question. Notons que cette organisation curriculaire n'est aujourd'hui plus appliquée depuis la mise en place du concordat

HarmoS, ce dernier devenant la référence pour désigner les principaux domaines de formation professionnelle.

Pendant la deuxième phase de la Rénovation : extension progressive (1998-2002)

Notre étude documentaire ne porte que sur les cours proposés dans les catalogues officiels, sachant qu'il existe désormais aussi les projets collectifs de formation qui offrent des espaces potentiels de réflexion et de développement professionnel. Notre analyse dégage un deuxième pic d'offres en matière d'évaluation en 1999-2000, correspondant à la première année de la phase d'extension progressive de la Rénovation de l'école primaire genevoise. Le tableau 2 présente les unités de formation proposées.

Tableau 2. Offres de formation professionnelle continue de l'année 1999-2000

Année	Approche transversale (3)	Approche disciplinaire (6)
1999-2000	- Évaluation : objectifs-noyaux, carnets simplifiés et portfolio de l'élève - L'évaluation à l'école enfantine et les objectifs préscolaires - Objectifs-noyaux pour la division élémentaire*	- Évaluation des mathématiques : quelle démarche avec les nouveaux moyens? - Comprendre, produire, évaluer le texte argumentatif - Produire, évaluer un récit d'aventures - Apprendre, produire, évaluer le texte descriptif : une activité interdisciplinaire - Connaissances de base pour la numération pour la division élémentaire - Résolution de problèmes : de la pratique d'enseignement à une réflexion sur l'évaluation

* Signifie que l'évaluation n'apparaît pas directement dans le titre, mais qu'elle est présentée comme un contenu, parmi d'autres, dans le descriptif de l'unité de formation.

On observe, comparativement à la période précédente, qu'un accent plus fort est mis sur une *entrée disciplinaire* de l'évaluation, en français et en mathématiques, en lien notamment avec l'introduction de nouveaux moyens didactiques. La plupart des activités disciplinaires concernées (rédiger un texte argumentatif, écrire un récit d'aventures, résoudre des problèmes mathématiques) sont liées à l'apprentissage chez les élèves de ce que l'on pourrait appeler des « compétences ». Nous y reviendrons plus loin. Les *descriptifs* des unités de formation concernées, donc au-delà des titres de ces dernières, montrent que l'évaluation est abordée quasi exclusivement au regard de sa fonction formative, de façon cohérente avec l'orientation du nouveau référentiel de compétences professionnelles. Rappelons que celui-ci semble ne pas considérer de façon explicite¹⁰ que l'évaluation certificative doit faire partie, au même titre que l'évaluation formative, des compétences à approfondir en formation

10. De façon explicite, parce que le référentiel en question joue des points de suspension et laisse ainsi entendre qu'il ne se veut pas exhaustif.

continue. Ce choix est révélateur, à notre sens, de la politique éducative de l'époque qui, souhaitant renforcer la valeur éducative de l'évaluation, l'attribuait exclusivement à l'évaluation formative. De Ketele (2006) observe de façon générale que l'évaluation certificative, vue comme le symbole d'une «évaluation-sanction», dénoncée en raison de l'échec scolaire et des inégalités sociales qu'elle contribue à produire (p. ex. Perrenoud, 1998), est alors victime de «réductionnisme historique». On peut penser que c'est également le cas pour l'école primaire genevoise, à la lumière du référentiel de compétences professionnelles et des offres de formation continue qui tendent pour la plupart à ignorer cette fonction de l'évaluation sur laquelle, pourtant, le système scolaire s'appuie pour gérer les flux et les distributions d'élèves au regard des exigences des différentes divisions et filières de formation.

Dans une *approche transversale*, le thème du *portfolio* apparaît pour la première fois en 1999-2000, visant à outiller les enseignants dans leurs démarches d'évaluation des progressions d'apprentissage, de même que les élèves pour des démarches d'auto-évaluation et de réflexion métacognitive. Ce thème sera récurrent tout au long des années de fin d'implantation de la réforme et des quelques années suivantes. En 1999-2000, le portfolio est associé à un autre outil d'évaluation et de communication, le carnet scolaire (précédemment appelé «livret scolaire», il a changé de nom avec la phase d'extension). Il s'agit de la seule unité de formation qui thématise l'évaluation certificative dans sa relation à l'évaluation formative (par le moyen du portfolio) ainsi qu'à des «objectifs-noyaux» (voir le tableau 2). Dans le prolongement des propositions de Meirieu (1989), l'objectif-noyau visait à désigner :

un apprentissage central, autour duquel les autres s'organisent [...] la vertu didactique d'un objectif-noyau est sa capacité d'organiser en réseau un ensemble de savoirs ou de compétences complémentaires, de lui donner une structure et une cohérence (Perrenoud, 2000, p. 23).

Dans le cadre de la Rénovation, un plan d'études pour l'école primaire a été réécrit, structuré autour de cette notion, et il devient logiquement un objet de formation continue proposé par le Département de l'instruction publique. Par rapport au mouvement émergent sur le plan international de «l'approche par compétences» et, avec celle-ci, de «l'évaluation des compétences», notre analyse montre que ces thématiques n'apparaissent pas de façon explicite dans l'ensemble des catalogues que nous avons examinés de 1976 à 2013. De fait, l'idée de compétence en milieu scolaire est controversée à Genève (voir Dolz et Ollagnier, 1999), expliquant peut-être cette absence qu'il faut toutefois relativiser. En effet, d'autres choix conceptuels ont été faits pour dépasser une pédagogie organisée autour d'objectifs spécifiques «atomisés» en faveur de la conception de savoirs conçus en réseau intégré et fonctionnel face à une famille de situations (Allal, 1999). C'est précisément le cas de la notion d'objectif-noyau (appelé aujourd'hui «objectif d'apprentissage» dans le plan d'études associé au concordat HarmoS) et de «situation complexe» attachée à l'élaboration de «séquences didactiques» (p. ex. Allal *et al.*, 2001). Ces choix conceptuels pour parler des compétences et de leur évaluation s'observent dans les descriptifs de formation continue.

Dans les années post-réforme (2002-2006)

Plus généralement et au-delà de la dernière phase de la Rénovation, notre analyse thématique montre que l'intégration de l'évaluation dans la réflexion didactique continue de se faire de façon de plus en plus marquée. Ainsi, dans l'année 2004-2005 par exemple, six offres s'inscrivent dans une approche transversale et onze autres sont proposées au regard de savoirs disciplinaires particuliers (lecture, écriture, vocabulaire, grammaire, orthographe, etc.). Le syntagme « enseignement/apprentissage/évaluation » est très présent dans la documentation, signe de la conception d'une évaluation pédagogique qui fait système avec l'enseignement et l'apprentissage.

En bref, il est à retenir que jamais les enseignants de l'école primaire genevoise n'avaient eu autant d'offres de formation continue en matière d'évaluation pédagogique, plus spécifiquement dans sa fonction formative, que dans les années 2000-2006, c'est-à-dire en fin d'implantation de la réforme scolaire et juste après cette phase.

Vers un équilibre formatif/certificatif?

Par le moyen de la réforme, l'évaluation formative a ainsi été fortement soutenue par les autorités scolaires (instituée dans la loi scolaire et le règlement d'application) et par des offres de formation continue. Il a été demandé aux enseignants de développer des pratiques d'évaluation formative de la progression des élèves, dans un contexte également de transformation des objets à évaluer en faveur d'apprentissages plus complexes et au regard aussi des spécificités des disciplines scolaires. Était-ce une focalisation trop forte sur l'évaluation formative aux yeux des acteurs sociaux concernés? Toujours est-il que, dès le milieu des années 2000, c'est-à-dire à peine quelques années après la généralisation de la réforme, un climat social conflictuel se développe en faveur de la réintroduction des notes chiffrées à l'école primaire genevoise.

Un retour du sommatif qui marque tant les projets que la réalité

Même si elle ne représentait pas un objet important de formation professionnelle continue comme on l'a vu plus haut, l'évaluation certificative n'a jamais disparu des classes de l'enseignement primaire, instituée dans la loi scolaire genevoise et devant obligatoirement s'exercer en fin de cycles. Au cycle élémentaire (élèves de 4 à 8 ans), l'évaluation certificative était *qualitative* et formalisée dans un « carnet scolaire », comme dans la plupart des cantons romands. Au cycle moyen (élèves de 8 à 12 ans), deux systèmes de notation ont coexisté dans le prolongement de la réforme. Les écoles engagées dans la Rénovation ont demandé à pouvoir remplacer les notes par des appréciations, alors que les autres écoles ont continué de fonctionner de manière plus traditionnelle avec des notes chiffrées. Mais, comme l'observe Soussi (2010), « en parallèle toutefois se développe une opposition entre enseignants ainsi qu'entre parents : les partisans des notes et des moyennes [...] versus ceux qui défendent une évaluation formative » (p. 61). Ainsi émerge un important hiatus : l'évaluation certificative qualitative est confondue à l'évaluation formative, laissant

alors croire qu'elle avait disparu des classes. La fonction certificative d'un *carnet scolaire sans notes* dans les premières années de l'école primaire n'a pas été comprise par les acteurs sociaux, y compris chez certains enseignants.

En 2005, l'institution scolaire, espérant résoudre le conflit, décide d'unifier les pratiques d'évaluation certificative en réintroduisant les notes dans toutes les écoles dès la 3^e année primaire (début du cycle moyen), mais en maintenant une structure en cycles d'apprentissage pluriannuels. Cette mesure ne suffira pas. Une consultation populaire en 2006 tranchera en faveur des notes et d'un découpage en degrés annuels, amenant le conseiller d'État de l'époque, responsable du Département de l'instruction publique, à également réintroduire les moyennes de notes pour décider les promotions/redoublements à la fin de chaque année scolaire. Dès 2007, un nouveau «livret scolaire» (nom repris d'avant la réforme, marquant ainsi institutionnellement la décision) est implanté :

Des notes sont attribuées par trimestre et la moyenne réintroduite dès la 3^e année dans trois disciplines (français 1, c'est-à-dire lecture et production écrite, français 2, c'est-à-dire structuration de la langue et mathématiques), l'allemand venant compléter ce bilan dès la 4^e année [...] En 5^e s'ajoutent les sciences de la nature et en 6^e les sciences humaines (histoire et géographie). Toutefois, pour le passage dans le secondaire I, seules les deux notes de français et les mathématiques sont prises en compte pour répartir les élèves (Soussi, 2010, p. 62).

Au regard de ce retour à des pratiques d'évaluation chiffrée «continue» (terme utilisé par les autorités scolaires), que l'on peut voir comme un retour de balancier à la suite certainement d'une reconnaissance sociale insuffisante de l'évaluation certificative dans la réforme, quels sont les enjeux par rapport à l'évaluation formative? Quelle place celle-ci garde-t-elle?

L'enjeu majeur : vers une évaluation pédagogique marchant sur ses deux pieds

La consultation populaire de 2006 n'a pas engagé une modification des textes réglementaires : l'évaluation formative reste identifiée comme une des fonctions à mobiliser dans les pratiques pédagogiques en salle de classe. Pourtant, notre analyse des catalogues de formation continue dégage un constat préoccupant : à l'exception d'une seule unité de formation sur le portfolio proposée en 2009-2010 (*Le portfolio, un outil au service des apprentissages*), plus aucun thème ne porte désormais explicitement sur l'évaluation formative. Les offres en évaluation se sont raréfiées et leur contenu s'est réorienté : *Faire le point sur ses pratiques d'évaluation en lien avec le Plan d'études romand* (associé au concordat HarmoS ratifié par le canton de Genève), *Corriger : quoi, pourquoi, comment?*, *Évaluer et informer*, *Le bulletin*¹¹ *scolaire au cycle 1*. Évidemment, cela ne présuppose pas des échanges effectifs développés pendant les formations.

11. Encore un nouveau terme introduit par l'institution et qui remplace celui de «livret scolaire», pour signifier la création d'un nouveau support administratif associé, cette fois-ci, au projet HarmoS.

Nous éviterons la conclusion hâtive de penser que les enseignants ne pratiquent désormais plus d'évaluation formative destinée à soutenir les progressions d'apprentissage des élèves et à produire des informations utiles pour organiser des dispositifs d'enseignement différencié. On notera simplement ici que l'évaluation formative ne semble plus représenter une priorité dans les projets politique et éducatif de l'école primaire genevoise à la suite du vote populaire en faveur de l'évaluation certificative chiffrée.

La limite principale de notre étude est de n'avoir analysé que des documents, sans recueillir de données sur les pratiques d'évaluation effectives en salle de classe. L'étude est également partielle, puisqu'elle n'a pas pris en considération les objets de développement professionnel retenus dans les projets d'établissement (voir plus haut). Nous pensons cependant que notre analyse documentaire est suffisamment révélatrice des oppositions réductrices qui ont été faites entre les fonctions formative et certificative de l'évaluation des apprentissages des élèves. L'enseignement que l'on peut retirer de cette histoire récente de l'école primaire genevoise est qu'il convient de penser ensemble les différentes fonctions de l'évaluation scolaire, constitutives d'un même système malgré les tensions possiblement paradoxales générées entre elles. Autrement dit, l'évaluation doit assumer ses fonctions pédagogiques, mais également distributives, dans une société qui a choisi de conditionner et de limiter l'accès aux différentes formations qui ont des valeurs sociales et marchandes hiérarchisées entre elles. Les acteurs sociaux, dont les familles des élèves, ne se sont pas trompés en craignant de ne plus avoir suffisamment de prise sur la fonction sociale de l'évaluation scolaire, même si, de notre point de vue, la centration sur les notes chiffrées est essentiellement symbolique (bien que créant des effets tangibles sur les pratiques enseignantes et sur les trajectoires des élèves). Un risque actuel est que l'évaluation formative devienne le parent pauvre ou, métaphoriquement dit, le pied bot du système. Et, on l'a compris, l'évaluation scolaire a besoin de ses deux pieds pour ne pas trébucher.

Conclusion : vers un modèle de l'évaluation pédagogique intégrative et systémique

Une perspective à développer, à notre sens, est donc la modélisation d'une évaluation pédagogique plus intégrative et systémique entre ces deux fonctions principales, formative et certificative, et d'autres fonctions plus ciblées. Sur un plan pédagogique, il est évidemment nécessaire de distinguer les buts de l'évaluation des apprentissages des élèves au regard des décisions qui en résultent (soutien à l'apprentissage des élèves, différenciation des dispositifs d'enseignement, certification des acquis des élèves, orientation, sélection). Mais il faut cesser de les penser en des termes moralistes : il y aurait une « bonne » évaluation (évaluation formative) et une « mauvaise » évaluation (évaluation certificative). Comme l'ont souligné Mottier Lopez et Laveault (2008), « rien n'empêche *a priori* une évaluation certificative de jouer un rôle de support à l'apprentissage, que ce soit au niveau cognitif ou motivationnel,

tout en reconnaissant évidemment sa fonction première d'attestation des apprentissages réalisés» (p. 18). Pour Hadji (1997), l'évaluation ne couvre jamais une seule et unique fonction. Quels que soient la forme et le but de l'évaluation, il faudrait que ce soit «la volonté d'aider qui en définitive installe l'activité évaluative dans un registre formatif» (p. 19).

Sur un plan social, les travaux des sociologues ont établi à juste titre les enjeux et impacts forts de l'évaluation certificative sur les parcours scolaires des élèves, sur l'accès aux formations et l'entrée dans le monde du travail, sur les inégalités sociales produites. Si, en effet, l'évaluation certificative formalise à travers les résultats qu'elle transmet les réussites et les échecs scolaires, ces derniers se «fabriquent» (au sens de Perrenoud, 1998) déjà dans le quotidien des activités en classe, par exemple dans l'adéquation d'un feed-back et d'une régulation entreprise, dans la prise en compte des caractéristiques des élèves pour concevoir les dispositifs d'enseignement; plus généralement, au regard des normes et des pratiques de chaque microculture de classe qui offre des opportunités différentes d'apprentissage (Mottier Lopez, 2008) et au regard aussi des structures des systèmes éducatifs qui promeuvent plus ou moins des mesures précoces de sélection. Autrement dit, lutter contre l'échec scolaire et les inégalités implique des valeurs et des actions qui concernent l'ensemble du système éducatif, sous toutes ses facettes, et pas seulement l'évaluation scolaire.

Un défi est de rendre légitimes, selon les politiques éducatives et les acteurs de terrain, des cadres de pensée et d'action fondés sur une *éthique de l'évaluation des apprentissages* des élèves, toutes ses fonctions devant tendre vers une finalité commune qui est celle d'œuvrer pour une évaluation la plus juste possible. Promouvoir le jugement professionnel des enseignants, en tant que «capacité d'agir dans des situations où le caractère limité du savoir professionnel ou de nos connaissances requiert un jugement éclairé» (Laveault, 2012), semble une voie prometteuse à poursuivre. Cette conception, qui s'oppose en effet à l'idée que l'on puisse évaluer les apprentissages des élèves et prendre des décisions à partir seulement de résultats chiffrés, vise résolument à ancrer l'évaluation pédagogique dans un cadre de professionnalité réflexif et critique (Laveault, 2008).

Références bibliographiques

- ALLAL, L. (1999). Acquisition et évaluation des compétences en situation scolaire. Dans J. Dolz et E. Ollagnier (dir.), *L'énigme de la compétence en éducation* (p. 77-94). Bruxelles : De Boeck.
- ALLAL, L. (2008). Évaluation des apprentissages. Dans A. van Zanten (dir.), *Dictionnaire de l'éducation* (p. 311-314). Paris : Presses universitaires de France.

- ALLAL, L., BETRIX KÖHLER, D., RIEBEN, L., ROUILLER, Y., SAADA-ROBERT, M. et WEGMULLER, E. (2001). *Apprendre l'orthographe en produisant des textes*. Fribourg: Éditions universitaires.
- BLOOM, B. S., HASTING, J. T. et MADAUS, G. F. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill.
- CRAHAY, M. (dir.) (2013). *L'école peut-elle être juste et efficace? De l'égalité des chances à l'égalité des acquis* (2^e éd. revue et actualisée). Bruxelles: De Boeck.
- DE KETELE, J.-M. (2006). La recherche en évaluation: propos synthétiques et prospectifs. *Mesure et évaluation en éducation*, 29(1), 99-118.
- DIRECTION DE L'ENSEIGNEMENT PRIMAIRE (1994). *Les trois axes de rénovation de l'école primaire genevoise*. Genève: Département de l'instruction publique.
- DOLZ, J. et OLLAGNIER, E. (dir.) (1999). *L'énigme de la compétence en éducation*. Bruxelles: De Boeck.
- GILLIERON GIROUD, P. et NTAMAKILIRO, L. (dir.) (2010). *Réformer l'évaluation scolaire: mission impossible?* Berne: Peter Lang.
- HUTMACHER, W. (1993). *Quand la réalité résiste à la lutte contre l'échec scolaire*. Genève: Service de la recherche sociologique, Cahier no 36, 3-165.
- LAVEAULT, D. (2008). Le jugement professionnel: foyer de tensions et de synergies nouvelles en évaluation scolaire. *Revue suisse des sciences de l'éducation*, 30, 483-500.
- LAVEAULT, D. (2012, juin). *Le jugement professionnel de l'enseignant. Quel impact sur la profession?* Conférence au 32^e colloque de l'AQPC, Gatineau, Canada.
- MAGER, R. F. (1977). *Comment définir des objectifs pédagogiques?* Paris: Dunod.
- MEIRIEU, p. (1989). *Apprendre... oui, mais comment?* (4^e éd.). Paris: ESF.
- MOTTIER LOPEZ, L. (2008). *Apprentissage situé: la microculture de classe*. Berne: Peter Lang.
- MOTTIER LOPEZ, L. et LAVEAULT, D. (2008). L'évaluation des apprentissages en contexte scolaire: développements, enjeux et controverses. *Mesure et évaluation en éducation*, 3, 5-34.
- MOTTIER LOPEZ, L. et TESSARO, W. (2010). Évaluation des compétences à l'école primaire genevoise: entre prescriptions et pratiques. *Mesure et évaluation en éducation*, 33(3), 29-53.
- PERRENOUD, P. (1998). L'évaluation des élèves. *De la fabrication de l'excellence à la régulation des apprentissages*. Bruxelles: De Boeck.
- PERRENOUD, P. (2000). Du bon usage des objectifs de formation dans un cycle d'apprentissage pluriannuel. *Éducateur*, 5, 19-24.
- SOUSSI, A. (2010). Les réformes du système d'évaluation dans le canton de Genève: bilan et perspectives. Dans P. Gillieron Giroud et L. Ntamakiliro (dir.), *Réformer l'évaluation scolaire: mission impossible?* (p. 59-77). Berne: Peter Lang.

Comparer les systèmes éducatifs francophones à travers le monde grâce au PISA : pas si simple!

Ariane BAYE

Université de Liège, Belgique

Marc DEMEUSE

Université de Mons, Belgique

Nathanaël FRIANT

Université de Mons, Belgique

VOLUME XLII : 3 – NUMÉRO SPÉCIAL,
AUTOMNE 2014

Revue scientifique virtuelle publiée par
l'Association canadienne d'éducation de
langue française dont la mission est la
suivante : « Par la réflexion et l'action de
son réseau pancanadien, l'ACELF exerce
son leadership en éducation pour
renforcer la vitalité des communautés
francophones ».

Directrice de la publication

Chantal Lainey, ACELF

Présidente du comité de rédaction

Lucie DeBlois,
Université Laval

Comité de rédaction

Sylvie Blain,
Université de Moncton
Lucie DeBlois,
Université Laval
Nadia Rousseau,
Université du Québec à Trois-Rivières
Paul Ruest,
Université de Saint-Boniface
Mariette Thérberge,
Université d'Ottawa

Directeur général de l'ACELF

Richard Lacombe

Conception graphique et montage

Claude Baillargeon

Responsable du site Internet

Étienne Ferron-Forget

Diffusion Érudit

www.erudit.org

Les textes signés n'engagent que
la responsabilité de leurs auteurs
et auteurs, lesquels en assument
également la révision linguistique.
De plus, afin d'attester leur recevabilité,
au regard des exigences du milieu
universitaire, tous les textes sont
arbitrés, c'est-à-dire soumis à des pairs,
selon une procédure déjà convenue.

La revue *Éducation et francophonie* est
publiée deux fois l'an grâce à
l'appui financier du ministère du
Patrimoine canadien et du Conseil
de recherches en sciences humaines
du Canada.

acelf

ASSOCIATION
CANADIENNE
D'ÉDUCATION DE
LANGUE FRANÇAISE

265, rue de la Couronne, bureau 303
Québec (Québec) G1K 6E1
Téléphone : 418 681-4661
Télécopieur : 418 681-3389
Courriel : info@acelf.ca

Dépôt légal
Bibliothèque et Archives nationales
du Québec
Bibliothèque et Archives du Canada
ISSN 1916-8659 (En ligne)
ISSN 0849-1089 (Imprimé)

Les politiques d'évaluation dans le domaine de l'éducation

Rédacteur invité :

Dany LAVEAULT

Liminaire

1 Les politiques d'évaluation en éducation. Et après?

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

15 La formule actuelle d'évaluation ministérielle au Nouveau-Brunswick. Pour qui? Pourquoi?

Jimmy BOURQUE, Université de Moncton, Nouveau-Brunswick, Canada

Mathieu LANG, Université de Moncton, Nouveau-Brunswick, Canada

Jean LABELLE, Université de Moncton, Nouveau-Brunswick, Canada

31 La politique québécoise d'évaluation des apprentissages et les pratiques évaluatives

Michel LAURIER, Université d'Ottawa, Ontario, Canada

50 La politique d'évaluation du rendement en Ontario : un alignement qui se précise dans la persévérance et la durée

Dany LAVEAULT, Université d'Ottawa, Ontario, Canada

Louise BOURGEOIS, Université d'Ottawa, Ontario, Canada

68 Une recherche documentaire sur les politiques, les pratiques et les principes directeurs d'évaluation des élèves du système scolaire public (1^{re} à 12^e année) des provinces de l'Ouest canadien

Jules ROCQUE, Université de Saint-Boniface, Manitoba, Canada

85 L'évaluation pédagogique va-t-elle enfin marcher sur ses deux pieds? Les enseignements de l'histoire récente de l'école primaire genevoise

Lucie MOTTIER LOPEZ, Université de Genève, Suisse

102 Comparer les systèmes éducatifs francophones à travers le monde grâce au PISA : pas si simple!

Ariane BAYE, Université de Liège, Belgique

Marc DEMEUSE, Université de Mons, Belgique

Nathanaël FRIANT, Université de Mons, Belgique

Comparer les systèmes éducatifs francophones à travers le monde grâce au PISA : pas si simple!

Ariane BAYE

Université de Liège, Belgique

Marc DEMEUSE

Université de Mons, Belgique

Nathanaël FRIANT

Université de Mons, Belgique

RÉSUMÉ

Le Programme international de suivi des acquis des élèves (PISA) permet de comparer les résultats des élèves de 15 ans issus d'un nombre grandissant de pays, principalement industrialisés. Cette évaluation récurrente des acquis des élèves fait l'objet de nombreuses publications. Cependant, si des pays francophones y participent avec des résultats plus ou moins favorables, il existe encore très peu d'écrits portant sur la comparaison de pays ayant le français en partage.

À cela, plusieurs raisons. Le premier problème réside dans la définition de ce qu'il faut entendre par « pays francophones ». Si plusieurs pays participant au PISA appartiennent à la francophonie politique, peu d'entre eux ont effectivement administré les tests en français et utilisent cette langue comme langue d'enseignement. Les pays d'Afrique subsaharienne, à l'opposé, ne participent pas au PISA. Il est donc

impossible d'opérer des comparaisons entre tous les pays où le français est effectivement la ou l'une des langues d'enseignement. Il faut aussi pouvoir identifier les résultats des établissements où le français est effectivement pratiqué au sein des pays qui utilisent plusieurs langues d'enseignement et comparer des situations comparables, en dehors de cette variable. De ce point de vue, les publications internationales ne fournissent pas toujours les informations, qu'il faut dès lors rechercher dans les rapports nationaux. Enfin, il faut sans doute tenir compte de la situation des élèves qui, bien que scolarisés en français, n'utilisent pas cette langue à la maison, ce qui complique encore les comparaisons.

ABSTRACT

Comparing Francophone educational systems worldwide through PISA – not easy!

Ariane BAYE
University of Liège, Belgium

Marc DEMEUSE
University of Mons, Belgium

Nathanaël FRIANT
University of Mons, Belgium

The Program for International Student Assessment (PISA) compares the results of 15-year-old students from a growing number of countries, mainly industrialized. This recurring evaluation of student achievement has been the subject of numerous publications. However, if Francophone countries participate with more or less favourable results, there is still very little literature comparing countries that share the French language.

There are several reasons for this. The first problem lies in the definition of what is meant by “Francophone countries”. Although several countries participating in PISA belong to the political Francophonie, few of them have effectively administered tests in French and use it as the language of instruction. The countries of Sub-Saharan Africa, in contrast, do not participate in PISA. It is therefore impossible to make comparisons among all the countries where French is actually the, or one of the, languages of instruction. It is also necessary to identify the performance of schools where French is actually practiced in the countries using several languages of instruction and compare similar situations outside of this variable. From this perspective, international publications do not always provide information, which must then be sought in national reports. The situation of students educated in French but not using the language at home, further complicating comparisons, should probably also be taken into account.

RESUMEN

Comparer los sistemas educativos francófonos a través del mundo gracia a PISA: ¿no es tan simple!

Ariane BAYE
Université de Lieja, Bélgica

Marc DEMEUSE
Universidad de Mons, Bélgica

Nathanaël FRIANT
Universidad de Mons, Bélgica

El programa internacional de seguimiento de conocimientos de los alumnos (PISA) permite comprender los resultados de los alumnos de 15 egresados de un creciente número de países, principalmente industrializados. Esta evaluación recurrente de conocimientos de los alumnos ha dado lugar a muchas publicaciones. Sin embargo si bien los países francófonos participan con resultados más o menos favorables, aun existen pocos escritos sobre la comparación de países que comparten el francés. Esto se debe a varias razones. El primer problema reside en la definición de lo que debe entenderse como «país francófono». Bien que varios países participantes a PISA forman parte de la francofonía política, de entre ellos muy pocos realizan efectivamente pruebas en francés y lo utilizan como lengua de enseñanza. Los países de África subsahariana, al contrario, no participan en el PISA. Es pues imposible realizar comparaciones entre los países en donde el francés es efectivamente la lengua o una de las lenguas de enseñanza. Es necesario poder identificar los resultados de los establecimientos en donde el francés se practica efectivamente en los países en donde se utilizan varias lenguas en la enseñanza y comparar situaciones comparables, más allá de esta variable. Desde este punto de vista, las publicaciones internacionales no siempre ofrecen la información necesaria, la cual hay que buscarla en los informes nacionales. Finalmente, hay que tener en cuenta la situación de los alumnos que, aunque se escolarizan en francés, no utilizan esta lengua en el hogar, lo que complica más aun las comparaciones.

Introduction

Le Programme pour le suivi des acquis des élèves (PISA) mis au point par l'Organisation de coopération et de développement économiques (OCDE) à la fin des années 1990 permet, depuis 2000, de récolter, tous les trois ans des données comparables sur les performances des élèves de 15 ans en lecture, en mathématique et en sciences dans un nombre de plus en plus grand de pays. Le groupe des pays participants s'est progressivement élargi, bien au-delà des seuls membres de l'OCDE, depuis la première campagne de test. Ces informations permettent de dresser les palmarès médiatisés des systèmes éducatifs, mais les chercheurs en éducation exploitent également ces données de manière à mieux comprendre le fonctionnement des divers systèmes (Demeuse, 2012). Il est néanmoins assez difficile de bien prendre en compte, d'une part, les éléments distinctifs et, d'autre part, les éléments communs à l'ensemble des systèmes éducatifs pour tenter d'expliquer les différences de performances moyennes, mais aussi les plus ou moins grandes disparités de résultats qui existent au sein d'un même système. Chaque système éducatif se distingue en effet de tous les autres, de manière plus ou moins affirmée, mais il partage aussi certaines caractéristiques, comme la langue d'enseignement, avec certains systèmes.

De quoi parle-t-on lorsqu'on se réfère à la Francophonie?

Une recherche documentaire mettant en rapport les résultats des différentes campagnes du PISA et les systèmes éducatifs francophones ramène très peu de documents. L'un de ceux-ci, produit en Suisse à l'issue de la première enquête menée en 2000, se réfère exclusivement à des systèmes éducatifs appartenant à quatre pays industrialisés: la Communauté française de Belgique, la France, le Québec et la Suisse romande (Soussi, Broi, Moreau et Wirthner, 2004). L'intérêt de cette étude est d'interroger quatre systèmes, sachant que la Suisse romande compte plusieurs cantons pouvant être considérés comme autant de systèmes éducatifs à part entière, qui partagent une même langue d'enseignement: le français. Ces quatre systèmes offrent l'avantage, pour les personnes intéressées par l'éducation comparée, de présenter des résultats assez contrastés, comme les résumant les auteurs de l'étude: «Le Québec obtient des résultats très élevés avec peu d'écarts entre les élèves, la Suisse romande a des résultats relativement bons, la France est légèrement au-dessus de la moyenne avec très peu d'écart entre élèves et la Communauté française de Belgique présente des résultats médiocres et très dispersés» (p. 3). La Francophonie dépasse cependant les frontières de ces quatre pays ou régions.

En effet, la Francophonie, si l'on considère son organisation politique, à travers l'Organisation internationale de la Francophonie (OIF), compte 77 États et gouvernements, dont 57 membres et 20 observateurs¹. Certains pays ayant une forte

1. Source: site de l'OIF (<http://www.francophonie.org/>).

présence francophone, comme l'Algérie, ne sont cependant ni membres ni même observateurs de l'OIE. A contrario, certains membres connaissent, de fait, une présence francophone très limitée sur leur territoire. Parmi les pays membres (M) de l'OIE, treize ont participé au moins une fois au PISA depuis la première évaluation en 2000², contre onze parmi les pays observateurs (O)³. Onze pays également, à savoir l'Autriche (O), la Belgique (M), le Canada (M), la France (M), la Grèce (M), la Hongrie (O), le Luxembourg (M), le Monténégro (O), la Pologne (O), la République tchèque (O) et la Suisse (O), ont participé aux cinq enquêtes (2000, 2003, 2006, 2009 et 2012). En dehors de la Tunisie, qui participe au PISA depuis 2003, aucun des pays de l'Afrique francophone n'a participé à cette enquête, mise en œuvre par l'OCDE pour ses membres et les pays partenaires volontaires.

La Conférence des ministres de l'Éducation des États et gouvernements de la Francophonie (CONFEMEN), qui regroupe un sous-ensemble des membres de l'OIE et qui compte actuellement 44 membres⁴, est sans doute un périmètre plus adéquat lorsqu'il s'agit de parler d'enseignement en langue française. Le Programme d'analyse des systèmes éducatifs (PASEC), lancé en 1991 lors de la 43e session ministérielle de la CONFEMEN à Djibouti, permet de fournir un certain nombre d'informations sur les systèmes éducatifs qui ne participent pas au PISA, notamment en Afrique subsaharienne⁵. Malheureusement, la méthodologie, les populations et les objectifs du PASEC sont très différents de l'enquête de l'OCDE. Les pays francophones industrialisés ne participent pas au PASEC, ce qui rend aujourd'hui les comparaisons entre les deux enquêtes complètement impossibles, même si l'on peut percevoir certaines convergences liées à une mondialisation de plus en plus grande du secteur de l'éducation (Cusso, 2007) et l'intérêt de l'OCDE pour l'évaluation des acquis des élèves dans les pays en développement⁶.

L'existence de pays ou régions bilingues permet, en théorie, de s'intéresser, toutes choses étant par ailleurs égales, aux performances des élèves fréquentant des écoles francophones ou d'autres écoles au sein de contextes scolaires et sociaux identiques. C'est par exemple le cas dans certaines provinces canadiennes, comme le Nouveau-Brunswick. Cependant, dans son rapport à propos de l'enquête PISA 2009, le ministère de l'Éducation et du Développement de la petite enfance (2012) de cette province canadienne précise, par exemple, que «le secteur francophone montre 77 %

2. L'Albanie, la Belgique (dont la Fédération Wallonie-Bruxelles), la Bulgarie, le Canada (dont le Nouveau-Brunswick et le Québec), Chypre, la France, la Grèce, le Luxembourg, le Qatar, la Roumanie, la Suisse, la Tunisie et le Vietnam.
3. L'Autriche, la Croatie, l'Estonie, la Hongrie, la Lettonie, la Lituanie, le Monténégro, la Pologne, la République tchèque, la Serbie et la Slovaquie.
4. Source: site de la CONFEMEN (<http://www.confemen.org/>).
5. En 2014, dix pays participeront au PASEC: le Bénin, le Burkina Faso, le Burundi, le Cameroun, la Côte d'Ivoire, le Congo, le Niger, le Sénégal, le Tchad et le Togo.
6. Ainsi, J. Malpel, coordinateur du PASEC, précise dans un entretien sur le site de la CONFEMEN (<http://www.confemen.org/3869/entretien-avec-m-jacques-malpel-coordonnateur-du-pasec-2/>): «Nous sommes en contact également avec l'OCDE qui, dans le cadre de son programme PISA pour le Développement, souhaite développer un programme de mesure des acquis scolaires pour les jeunes dans les pays en développement. Nous avons le devoir et l'intérêt de travailler avec ce type d'initiatives afin que le PASEC soit reconnu dans le concert international et que nous puissions apprendre des autres.»

de ses écoles dans des zones rurales, comparativement à 52 % pour le secteur anglophone. La population du Nouveau-Brunswick francophone est donc plus rurale que celle du Canada et du Nouveau-Brunswick anglophone». Il ajoute «qu'une vaste majorité d'écoles du Nouveau-Brunswick francophone indique que moins de 10 % de leurs élèves ont une langue maternelle différente de celle du test. Ce pourcentage est de 47 % pour le Canada et de 34 % pour le secteur anglophone» (p. 118). Il n'est donc pas possible de comparer les résultats des élèves de ces deux secteurs sans être extrêmement prudent, sans quoi le risque est grand d'attribuer à une seule variable, la langue d'enseignement, toutes les différences de rendement qui peuvent s'expliquer par d'autres variables, plus complexes à cerner ou à collecter. Cet appel à la prudence est encore plus important dans le cas de pays qui comportent cette fois des régions distinctes quant à l'utilisation des langues, comme la Belgique ou la Suisse.

Le PISA, une épreuve développée en deux langues

Le français bénéficie d'un statut particulier dans le PISA. L'OCDE est en effet une organisation internationale bilingue dont le siège est situé à Paris. Dès le premier cycle de l'étude, deux versions «sources» du test ont été produites, l'une en anglais, l'autre en français. Ces deux versions internationales du test font l'objet d'une vérification serrée en vue d'établir la meilleure correspondance possible entre elles. Les pays sont invités à traduire le test dans leur(s) langue(s) nationale(s) à partir de ces deux versions sources. Se référer aux deux versions sources du test s'avère toujours utile dans un processus de traduction, car cette opération permet de lever des ambiguïtés dues à l'idiosyncrasie d'une seule langue source. Les difficultés de traduction dues aux particularités sémantiques ou syntaxiques d'une des versions sources ont en effet été résolues dans l'autre version source. La qualité de la procédure de traduction à partir de deux versions sources a d'ailleurs été documentée. Les versions nationales du test PISA 2003 réalisées à partir des deux versions sources (soit par une double traduction à partir de l'anglais et du français, soit par une double traduction de l'anglais avec une vérification approfondie à partir de la version française) sont d'une qualité nettement supérieure aux traductions réalisées uniquement à partir de l'anglais (OECD, 2005a).

Le statut bilingue de l'OCDE offre également aux pays qui utilisent la version francophone du test une version d'une grande qualité qui ne doit plus être traduite, mais simplement adaptée aux particularités linguistiques/sémantiques/éducatives des pays francophones participant au PISA.

Quels résultats dans les pays francophones?

Les trois tableaux suivants présentent, en synthèse, les résultats moyens (et les erreurs liées à ces estimations, entre parenthèses⁷) des différents États membres de

7. Dans le PISA, l'erreur de mesure est estimée par une méthode de rééchantillonnage à répliques équilibrées (Balanced Repeated Replication method), telle qu'adaptée par Fay («PISA uses the Fay method with a factor of 0.5», OECD, 2005b, p. 50).

l'OIF lors des cinq campagnes de test du PISA, entre 2000 et 2012. Pour interpréter ces résultats, il faut savoir que la moyenne des pays de l'OCDE est fixée à 500 et l'écart type à 100 chaque fois qu'un domaine est le domaine principal de l'évaluation pour la première fois (2000 pour la lecture, 2003 pour la culture mathématique et 2006 pour la culture scientifique). Seuls les membres de l'OIF qui ont participé à au moins une campagne de test figurent dans les tableaux. Les résultats, établis à partir des publications internationales réalisées par l'OCDE (2001, 2004, 2007, 2011a, 2011b, 2014; OECD/UNESCO-UIS, 2003) à l'issue de chaque campagne, sont présentés séparément pour la lecture (tableau 1), la culture mathématique (tableau 2) et la culture scientifique (tableau 3). La Fédération Wallonie-Bruxelles figure dans les tableaux de manière autonome, sous la ligne consacrée à la Belgique, car elle fait l'objet d'une publication dans chaque rapport international, contrairement aux résultats des différents cantons suisses ou des provinces canadiennes, dont le Québec et le Nouveau-Brunswick, tous deux membres de l'OIF. Il en est de même pour d'autres pays où l'éducation relève en tout ou en partie d'entités ou de niveaux de pouvoir différents, comme dans le cas des États-Unis d'Amérique. Progressivement cependant, au fil des parutions, certaines entités, comme l'Espagne ou l'Italie, obtiennent de figurer dans le rapport international de l'OCDE, ce qui présente une certaine logique dans la mesure où cette étude à large échelle vise à soutenir les décideurs dans leurs tâches.

Tableau 1. Résultats moyens des pays appartenant à l'OIF à la partie « lecture » du PISA entre 2000 et 2012 (les erreurs d'estimation sont indiquées entre parenthèses)

	2000	2003	2006	2009	2012
Albanie	349 (3,3)	---	---	385 (4,0)	394 (3,2)
Autriche	507 (2,4)	491 (3,8)	490 (4,1)	470 (2,9)	490 (2,8)
Belgique *	507 (3,6)	507 (2,6)	501 (3,0)	506 (2,3)	509 (2,3)
Fédération Wallonie-Bruxelles *	476 (7,2)	477 (5,0)	473 (5,0)	490 (4,2)	497 (3,9)
Bulgarie	430 (4,9)		402 (6,9)	429 (6,7)	436 (6,0)
Canada *	534 (1,6)	528 (1,7)	527 (2,4)	524 (1,5)	523 (1,9)
Chypre	---	---	---	---	449 (1,2)
Croatie	---	---	477 (2,8)	476 (2,9)	485 (3,3)
Émirats arabes unis	---	---	---	---	442 (2,5)
Estonie	---	---	501 (2,9)	501 (2,6)	516 (2,0)
France *	505 (2,7)	496 (2,7)	488 (4,1)	496 (3,4)	505 (2,8)
Grèce	474 (5,0)	472 (4,1)	460 (4,0)	483 (4,3)	477 (3,3)
Hongrie	480 (4,0)	482 (2,5)	482 (3,3)	494 (3,2)	488 (3,2)
Lettonie	458 (5,3)	491 (3,7)	479 (3,7)	484 (3,0)	---
Lituanie	---	---	470 (3,0)	468 (2,4)	477 (2,5)
Luxembourg *	441 (1,6)	479 (1,5)	479 (1,3)	472 (1,3)	488 (1,5)
Monténégro	---	---	392 (1,2)	408 (1,7)	422 (1,2)
Pologne	479 (4,5)	497 (2,9)	508 (2,8)	500 (2,6)	518 (3,1)
Qatar	---	---	---	---	388 (0,8)
République tchèque	492 (2,4)	489 (3,5)	483 (4,2)	478 (2,9)	493 (2,9)
Roumanie	---	---	508 (2,8)	424 (4,1)	438 (4,0)
Serbie	---	412 (3,6)	401 (3,5)	442 (2,4)	446 (3,4)
Slovaquie	---	469 (3,1)	466 (3,1)	477 (2,5)	463 (4,2)
Slovénie	---	---	494 (1,0)	483 (1,0)	481 (1,2)
Suisse *	494 (4,3)	499 (3,3)	499 (3,1)	501 (2,4)	509 (2,6)
Tunisie	---	375 (2,8)	380 (4,0)	404 (2,9)	404 (4,5)
Vietnam	---	---	---	---	508 (4,4)
OCDE	500 (0,6)	494 (0,6)	492 (0,6)	493 (0,5)	496 (0,5)

* Pays ayant totalement ou partiellement administré les tests du PISA en français.

Tableau 2. Résultats moyens des pays appartenant à l'OIF à la partie « culture mathématique » du PISA entre 2000 et 2012 (les erreurs d'estimation sont indiquées entre parenthèses)

	2000	2003	2006	2009	2012
Albanie	381 (3,1)	---	---	377 (4,0)	394 (2,0)
Autriche	515 (2,5)	506 (3,3)	505 (3,7)	496 (2,7)	506 (2,7)
Belgique *	520 (3,9)	529 (2,3)	520 (3,0)	515 (2,3)	515 (2,1)
Fédération Wallonie-Bruxelles *	491 (7,2)	498 (4,3)	490 (5,2)	488 (3,9)	493 (2,9)
Bulgarie	430 (5,7)	---	413 (6,1)	428 (5,9)	439 (4,0)
Canada *	533 (1,4)	532 (1,8)	527 (2,0)	527 (1,6)	518 (1,8)
Chypre	---	---	---	---	440 (1,1)
Croatie	---	---	467 (2,4)	460 (3,1)	471 (3,5)
Émirats arabes unis	---	---	---	---	434 (2,4)
Estonie	---	---	515 (2,7)	512 (2,6)	521 (2,0)
France *	517 (2,7)	511 (2,5)	496 (3,2)	497 (3,1)	495 (2,5)
Grèce	447 (5,6)	445 (3,9)	459 (3,0)	466 (3,9)	453 (2,5)
Hongrie	488 (4,0)	490 (2,8)	491 (2,9)	490 (3,5)	477 (3,2)
Lettonie	463 (4,5)	483 (3,7)	486 (3,0)	482 (3,1)	---
Lituanie	---	---	486 (2,9)	477 (2,6)	479 (2,6)
Luxembourg *	446 (2,0)	493 (1,0)	490 (1,1)	489 (1,2)	490 (1,1)
Monténégro	---	---	399 (1,4)	403 (2,0)	410 (1,1)
Pologne	470 (5,5)	490 (2,5)	495 (2,4)	495 (2,8)	518 (3,6)
Qatar	---	---	---	---	376 (0,8)
République tchèque	498 (2,8)	516 (3,5)	510 (3,6)	493 (2,8)	499 (2,9)
Roumanie	---	---	415 (4,2)	427 (3,4)	445 (3,8)
Serbie	---	437 (3,8)	435 (3,5)	442 (2,9)	449 (3,4)
Slovaquie	---	498 (3,3)	492 (2,8)	497 (3,1)	482 (3,4)
Slovénie	---	---	504 (1,0)	501 (1,2)	501 (1,2)
Suisse *	529 (4,4)	527 (3,4)	530 (3,2)	534 (3,3)	531 (3,0)
Tunisie	---	359 (2,5)	365 (4,0)	371 (3,0)	388 (3,9)
Vietnam	---	---	---	---	511 (4,8)
OCDE	500 (0,7)	500 (0,6)	498 (0,5)	496 (0,5)	494 (0,5)

* Pays ayant totalement ou partiellement administré les tests du PISA en français.

Tableau 3. Résultats moyens des pays appartenant à l'OIF à la partie « culture scientifique » du PISA entre 2000 et 2012 (les erreurs d'estimation sont indiquées entre parenthèses)

	2000	2003	2006	2009	2012
Albanie	376 (2,9)	---	---	391 (3,9)	397 (2,4)
Autriche	519 (2,6)	490 (4,3)	511 (3,9)	494 (3,2)	506 (2,7)
Belgique *	496 (4,3)	509 (3,6)	510 (2,5)	507 (2,5)	505 (2,2)
Fédération Wallonie-Bruxelles *	467 (8,7)	483 (4,6)	486 (4,3)	482 (4,2)	487 (3,3)
Bulgarie	448 (4,6)	---	434 (6,1)	439 (5,9)	446 (4,8)
Canada *	529 (1,6)	527 (2,3)	534 (2,0)	529 (1,6)	525 (1,9)
Chypre	---	---	---	---	438 (1,2)
Croatie	---	---	493 (2,4)	486 (2,8)	491 (3,1)
Émirats arabes unis	---	---	---	---	448 (2,8)
Estonie	---	---	531 (2,5)	528 (2,7)	541 (1,9)
France *	500 (3,2)	511 (4,1)	495 (3,4)	498 (3,6)	499 (2,6)
Grèce	461 (4,9)	487 (4,8)	473 (3,2)	470 (4,0)	467 (3,1)
Hongrie	496 (4,2)	503 (3,3)	504 (2,7)	503 (3,1)	494 (2,9)
Lettonie	460 (5,6)	487 (5,1)	490 (3,0)	494 (3,1)	---
Lituanie	---	---	488 (2,8)	491 (2,9)	496 (2,6)
Luxembourg *	443 (2,3)	489 (2,5)	486 (1,1)	484 (1,2)	491 (1,3)
Monténégro	---	---	412 (1,1)	401 (2,0)	410 (1,1)
Pologne	483 (5,1)	501 (3,2)	498 (2,3)	508 (2,4)	526 (3,1)
Qatar	---	---	---	---	384 (0,7)
République tchèque	511 (2,4)	526 (4,3)	513 (3,5)	500 (3,0)	508 (3,8)
Roumanie	---	---	418 (4,2)	428 (3,4)	439 (3,3)
Serbie	---	434 (3,7)	436 (3,0)	443 (2,4)	445 (3,4)
Slovaquie	---	502 (4,3)	488 (2,6)	490 (3,0)	471 (3,6)
Slovénie	---	---	519 (1,1)	512 (1,1)	514 (1,3)
Suisse *	496 (4,4)	518 (5,0)	512 (3,2)	517 (2,8)	515 (2,7)
Tunisie	---	380 (2,7)	386 (3,0)	401 (2,7)	398 (3,5)
Vietnam	---	---	---	---	528 (4,3)
OCDE	500 (0,7)	503 (0,7)	500 (0,5)	501 (0,5)	501 (0,5)

* Pays ayant totalement ou partiellement administré les tests du PISA en français.

Le test PISA est administré dans la langue d'enseignement de l'établissement scolaire. Dans les pays multilingues, autant de versions linguistiques que de langues d'enseignement en vigueur ont été développées. Dans le cas de langues très minoritaires, la question d'inclure ces écoles a été discutée en termes d'échantillonnage (Adams et Wu, 2002, p. 60). Comme déjà évoqué, les pays recourant à une version francophone du PISA sont relativement peu nombreux. Dans la version 2009, cinq pays ont en effet administré, à des échantillons d'ampleur variée, l'épreuve PISA en français (tableau 4) : la Belgique, la Suisse, la France, le Luxembourg et le Canada, soit cinq pays totalement ou partiellement francophones.

Tableau 4. **Effectif des échantillons auxquels le PISA 2009 a été administré en français et proportion de la population des élèves de 15 ans auxquels cette version est destinée**

Pays	Nombre d'élèves ayant passé le test en français	% de la population d'élèves francophones de 15 ans
Belgique	3 148	44,04
Suisse	4 266	25,26
France	4 298	100
Luxembourg	1 035	22,87
Canada	5 619	22,66

En 2009, lorsqu'on compare les résultats aux différentes sous-échelles de compréhension dans les cinq pays ayant fait passer le test PISA en français, on peut noter que les résultats tendent à y être inférieurs sur l'échelle qui évalue la capacité à accéder et à retrouver de l'information. Au Canada, en Suisse et au Luxembourg, les scores pour cette sous-échelle sont significativement inférieurs à une ou deux autres sous-échelles de compréhension en lecture (tableau 5). Les questions les plus simples de cette composante du test requièrent de localiser une information explicitement présentée dans le texte, mais les questions les plus complexes de cette sous-échelle demandent que l'élève retrouve et combine des informations situées à différents endroits du stimulus⁸. Les résultats pour la sous-échelle «Accéder et retrouver» sont intéressants à examiner de près dans une optique prospective, car ce type de compétence est largement mobilisé dans des environnements comme Internet, où tout lecteur-navigateur est régulièrement appelé à faire converger des informations glanées au fil de ses recherches⁹.

8. Le terme «stimulus» est employé ici, car la nature du document sur lequel les élèves sont interrogés peut être très variable. Il peut s'agir aussi bien d'un texte continu que d'un tableau ou d'un schéma ou, encore, d'une combinaison de plusieurs types de documents.
9. À cet égard, notons que depuis 2009 un test optionnel permet aux pays d'évaluer les compétences des élèves de 15 ans sur ordinateur. D'abord optionnel, le test sur support électronique est devenu obligatoire en 2015 et il est amené à remplacer le test papier-crayon.

En 2009, les scores des francophones¹⁰ belges et luxembourgeois (tableau 5) sont inférieurs aux résultats internationaux (moyenne de l'OCDE: 495 points) sur cette sous-échelle. Les résultats des Français sont comparables à la moyenne internationale, et ceux des Suisses et Canadiens francophones sont supérieurs.

Tableau 5. **Résultats des élèves ayant passé les tests PISA 2000 et PISA 2009 en français aux trois sous-échelles de littératie** (les erreurs d'estimation sont indiquées entre parenthèses)

Pays	Accéder et retrouver		Intégrer et interpréter		Réfléchir et évaluer	
	2000	2009	2000	2009	2000	2009
Belgique	476,4 (7,87)	483,9 (3,98)	481,6 (5,83)	489,6 (4,50)	465,6 (9,14)	491,6 (4,67)
Canada	526 (n. d.)	510,1 (3,56)	537 (n. d.)	516,3 (3,32)	533 (n. d.)	520,7 (3,30)
Suisse	517,9 (6,42)	516,4 (3,00)	516,3 (6,17)	525,9 (3,06)	500,0 (6,10)	525,1 (3,39)
France	514,9 (2,48)	491,6 (3,78)	505,6 (2,71)	497,2 (3,56)	496,2 (2,86)	495,2 (3,43)
Luxembourg	---	447,6 (3,31)	---	455,6 (3,05)	---	449,6 (3,16)
OCDE	498 (0,70)	495 (0,50)	501 (0,60)	493 (0,50)	502 (0,70)	494 (0,50)

n. d. : non disponible.

Source : Soussi, Broi, Moreau et Wirthner, 2004, p. 18 pour le Canada francophone (pour PISA 2000); calcul des auteurs pour les autres pays.

Les résultats aux différentes sous-échelles de compétences ont parfois nettement évolué depuis la première opération centrée sur la lecture. En 2000, les pays francophones européens se caractérisaient par des scores plus faibles (voire très faibles dans le cas de la Belgique) sur l'échelle « Réfléchir et évaluer ». Cette sous-échelle évalue la capacité des élèves à faire des liens entre les textes et les connaissances extérieures. Il s'agit par exemple de faire appel à des connaissances personnelles pour former une hypothèse cohérente avec les informations fournies par un texte. Certaines questions demandent également à l'élève de se positionner face aux choix d'un auteur. On demande ainsi aux élèves de donner leur avis sur les choix formels ou de contenu effectués par un auteur. Pour expliquer la relative faiblesse des pays francophones européens en 2000, faiblesse qui n'apparaît pas au Canada, on peut faire l'hypothèse d'une tradition de « respect » par rapport aux textes et aux auteurs qui confère à ces derniers un statut d'« intouchables ». Dans ces pays,

10. Nous utilisons dans ce texte le terme « francophone » à propos des élèves, bien qu'il s'agisse d'une simplification un peu rapide. Les élèves sont en fait des jeunes de 15 ans qui sont scolarisés en français et qui ont reçu une version francophone du test, en raison de la langue de scolarisation. Il peut naturellement s'agir d'élèves dont la langue maternelle n'est pas le français ou qui fréquentent, comme au Luxembourg, un système éducatif multilingue qui a soit opté pour une version francophone du test, soit laissé le choix aux élèves de la langue de test.

on peut supposer que, même si l'analyse textuelle est fréquente en classe, on demande rarement aux élèves de fournir une réponse personnelle par rapport à un texte, réponse impliquant à la fois un positionnement personnel et un étayage à partir des informations du texte.

Or, dans le domaine de la lecture, la compréhension est aujourd'hui définie comme un processus dynamique impliquant le lecteur, le texte et le contexte (Campbell, Kelly, Mullis, Martin et Sainsbury, 2001; Mullis, Kennedy, Martin et Sainsbury, 2006; Rosenblatt, 1978, 2004; Wilkinson et Son, 2011). Pour enseigner la compréhension, il faut guider les élèves dans cette dynamique de construction de sens qui les amène à voir que la signification ne réside pas uniquement dans le texte, mais est bien une construction opérée dans l'interaction texte-lecteur dans un contexte donné. « Le sens ne réside pas clé-sur-porte "dans" le texte ou "dans" l'esprit du lecteur, mais se manifeste ou commence à exister pendant la transaction entre le texte et le lecteur¹¹ » (Rosenblatt, 2004, p. 1063). Pour Rosenblatt, qui a proposé la théorie transactionnelle, un élément clé pour développer des activités de lecture en classe est la notion de « cross-fertilization », c'est-à-dire le renforcement mutuel des processus de lecture et d'écriture, et les échanges entre l'enseignant et l'élève, et entre les élèves entre eux, échanges au cours desquels les élèves partagent leurs « réponses » à partir d'un même texte. À partir de ces expériences collectives, les élèves « vont apprendre comment leur interprétation des mêmes signes diffère [...] et peuvent prendre conscience et être critiques de leurs propres processus de lecteurs¹² » (Rosenblatt, 2004, p. 1084).

Les données empiriques qui permettraient de mettre en rapport les évolutions des scores des francophones aux différentes sous-échelles de PISA entre 2000 et 2009 et les pratiques d'enseignement de la lecture ne sont malheureusement pas disponibles. Tout au plus est-il possible d'évoquer dans le cas des Belges francophones le « choc » qu'a constitué la contre-performance au test PISA 2000 dans le domaine de la lecture, à la suite duquel la lecture est devenue un centre de préoccupation pour de nombreux acteurs de l'enseignement. Dans le même temps, le pilotage du système éducatif a été renforcé et les évaluations externes ont été systématisées. La présence régulière dans les évaluations externes de questions relevant de la compétence « Réfléchir et évaluer », telle que définie dans le PISA, a sans doute contribué au développement de ce type de questionnement sur les textes dans l'enseignement/évaluation de la lecture en Belgique francophone.

Comme l'ont précisé Soussi, Broi, Moreau et Wirthner dès 2004, en considérant les résultats de l'enquête PISA 2000, il est « illusoire de croire qu'il suffit de prendre comme modèle ce qui se fait au Québec pour améliorer d'autres systèmes scolaires. Il n'en demeure pas moins qu'aujourd'hui la tendance des nouveaux plans d'études élaborés ces dernières années va dans le sens du modèle québécois : définition de

11. Notre traduction de: "The meaning does not reside ready-made 'in' the text or 'in' the reader's mind but happens or comes into being during the transaction between reader and text".

12. Notre traduction de: "will learn how their evocation from the same signs differ [...] and can become aware of, and critical of, their own processes as readers".

compétences (socles de compétences) et évaluation multiples présentées dans les programmes...» (p. 70).

Les facteurs motivationnels présentent généralement de bonnes corrélations avec les performances en lecture au niveau intranational et peuvent être considérés comme des leviers pédagogiques mobilisables au sein des classes. L'indice « Plaisir de lire », construit à partir d'items présentés sous forme d'échelle de Likert, mesure les attitudes des élèves à l'égard du livre et de la lecture (tableau 6). Les propriétés psychométriques de l'indice ne sont pas satisfaisantes sur le plan de la comparabilité internationale. En effet, les questions servant à la création de l'indice sont présentées aux élèves sous forme d'échelles de Likert. Or, il est aujourd'hui établi que ces échelles sont sujettes à des biais en fonction des styles de réponses (Kyllonen et Bertling, 2014; Yang, Harkness, Chin et Villar, 2010). Ces biais culturels, qui incluent la propension à l'acquiescence, la tendance à se positionner aux extrémités d'une échelle ou les phénomènes de désirabilité sociale, empêchent de comparer entre eux les indices moyens des différents pays. Par contre, le caractère cyclique du PISA permet de mesurer l'évolution des populations d'élèves de 15 ans dans chaque pays sur ce type d'indice. En termes d'attitudes à l'égard de la lecture, on observe une évolution assez semblable à celle décrite à propos des résultats cognitifs des élèves. Les élèves français, belges et canadiens testés en français y progressent, et les élèves suisses maintiennent leurs résultats.

Tableau 6. **Évolution sur l'échelle « Plaisir de lire »¹³ des élèves ayant passé le test PISA en français en 2000 et 2009**

Pays	Plaisir de lire 2000		Plaisir de lire 2009	
	Moyenne	Erreur	Moyenne	Erreur
Belgique	-0,14	0,03	-0,02	0,03
Canada	-0,08	0,02	0,07	0,02
Suisse	0,12	0,05	0,15	0,05
France	-0,06	0,02	0,01	0,03
Luxembourg	n. d.	n. d.	-0,07	0,03
OCDE	0,00	0,00	0,00	0,00

n. d. : non disponible.

Source : Calculs des auteurs pour la Belgique, la Suisse, la France et le Luxembourg. Calculs réalisés par l'équipe de P. Brochu (Conseil des ministres de l'Éducation) pour le Canada.

13. L'indice est établi à partir d'une estimation pondérée des réponses les plus vraisemblables (Weighted maximum likelihood estimate, WLE) (Warm, 1989). La moyenne de l'indice est fixée à 0 pour l'ensemble des pays de l'OCDE, l'écart type est fixé à 1. Une moyenne positive au niveau d'un pays signifie qu'en moyenne dans ce pays les élèves font preuve d'attitudes plus positives qu'en moyenne au niveau international.

Qu'advient-il des élèves qui ne parlent pas le français à la maison dans les systèmes éducatifs francophones?

Il est souvent assez difficile de comparer les systèmes éducatifs par rapport au seul aspect de la langue d'enseignement. Ainsi, comme l'ont montré Monseur et Demeuse (2004), les résultats pour la lecture, du moins lors de la première enquête PISA, permettent de distinguer les élèves autochtones¹⁴ au sens de PISA des élèves de première génération¹⁵ ou des élèves allochtones¹⁶. Cependant, la distinction n'est pas toujours à l'avantage du groupe des élèves autochtones. Cela s'explique notamment par la nature des processus migratoires. Ainsi, dans les pays qui connaissent une immigration régulée par les besoins économiques nationaux, les élèves allochtones peuvent se montrer les plus performants s'ils parlent la langue d'enseignement à la maison (y compris même par rapport aux élèves autochtones qui parlent également la langue d'enseignement à la maison). C'est notamment le cas dans des pays anglophones comme la Nouvelle-Zélande, l'Australie ou les États-Unis d'Amérique, et même au Canada où la situation linguistique est plus complexe. Par contre, les pays francophones européens (France, Suisse, Luxembourg et Belgique) ne connaissent en principe pas ce type d'immigration¹⁷ et le statut d'immigré, même associé à l'usage de la langue d'enseignement à la maison, reste toujours défavorable par rapport aux résultats en lecture selon PISA 2000. Cette différence de sélectivité à l'entrée du territoire est aussi soulignée par l'OCDE (2004, p. 180), alors même que ce sont les mathématiques qui constituent l'objet principal de cette évaluation en 2003 :

Certains pays accueillent de fortes proportions d'immigrants chaque année et se montrent peu sélectifs à l'entrée du territoire. D'autres pays en revanche appliquent des politiques d'immigration nettement plus sélectives, ce qui limite les flux d'immigration. En outre, la priorité accordée au statut professionnel et au statut économique, social et culturel par rapport aux autres critères appliqués pour autoriser l'entrée sur le territoire des immigrants ou pour décider de leur naturalisation varie selon les pays. C'est ce qui explique pourquoi les populations immigrées appartiennent à des milieux plus favorisés dans certains pays que dans d'autres.

Ce n'est donc pas tant la proportion des immigrants qui détermine les résultats au PISA que leur statut socioéconomique et culturel, de même que leur bonne maîtrise de la langue d'enseignement.

14. Élèves nés dans le pays de l'évaluation dont au moins un des parents est né dans le pays.

15. Élèves nés dans le pays de l'évaluation dont les parents sont nés à l'étranger.

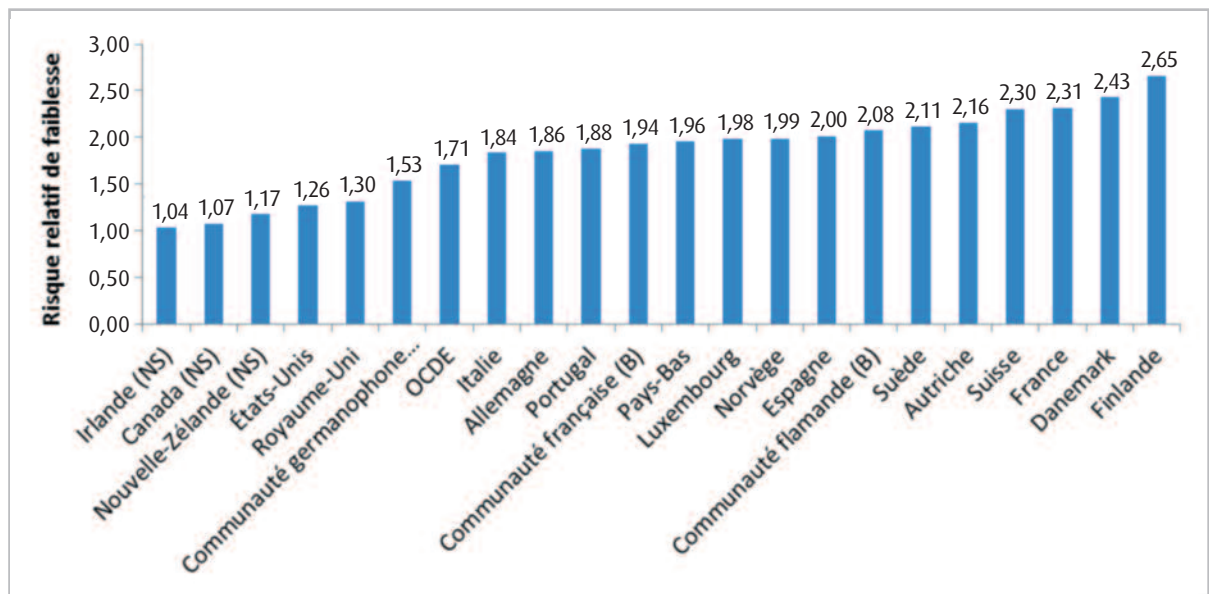
16. Élèves nés à l'étranger dont les parents sont nés à l'étranger.

17. La nature des migrations s'est transformée progressivement en Europe. Par exemple, en Belgique, « après le premier choc pétrolier, la décision prise en 1974 de mettre un terme à l'immigration de travail avait clairement conduit à une diminution du nombre d'immigrations étrangères, permettant même d'enregistrer un solde migratoire négatif pour la population étrangère en 1982 et 1983 ». Après cette date, l'immigration a pu reprendre sous deux formes différentes: la première est liée au « regroupement familial » et la seconde à l'immigration intra-européenne. À ces deux formes d'immigration s'en est ajoutée une troisième, liée à partir de 1989 à l'augmentation spectaculaire des demandeurs d'asile (Centre interfédéral pour l'égalité des chances/Centre fédéral migration, 2012, p. 8). Depuis les premières migrations, en dehors peut-être des migrations intra-européennes, y compris lorsqu'il s'agissait d'immigration de travail, ce sont surtout des populations moins qualifiées qui ont été accueillies, contrairement aux politiques d'accueil sélectif, par exemple, au Canada.

À côté de cet effet lié au statut des immigrés, certains systèmes semblent mieux que d'autres permettre aux élèves ne parlant pas la langue d'enseignement à la maison de la maîtriser en réduisant les écarts de performances entre groupes d'élèves. Sans être des pays complètement francophones, «la Belgique et la Suisse figurent toujours [après prise en compte du niveau socioéconomique et éducatif des élèves] parmi ceux qui affichent les plus grandes disparités de score entre les élèves issus de l'immigration et les élèves autochtones» (p. 183).

La figure 1 exprime en termes de risque relatif la relation entre statut d'immigré et faiblesse scolaire (ici, en mathématiques). Il est frappant de constater que dans les pays anglo-saxons mentionnés ci-dessous, les immigrés n'encourent pas un risque plus élevé de faire partie du quartile des élèves les plus faibles que les natifs (différence non significative en Irlande, au Canada et Nouvelle-Zélande), ou que ce risque est modéré par rapport aux autres pays (différences significatives, mais faibles aux États-Unis et au Royaume-Uni). Pour les pays européens partiellement francophones, on note qu'en Belgique francophone et au Luxembourg les étudiants d'origine immigrée ont presque deux fois plus de risques que les natifs d'être en situation de grande faiblesse en mathématiques. Ce risque est plus aigu encore en Suisse et en France.

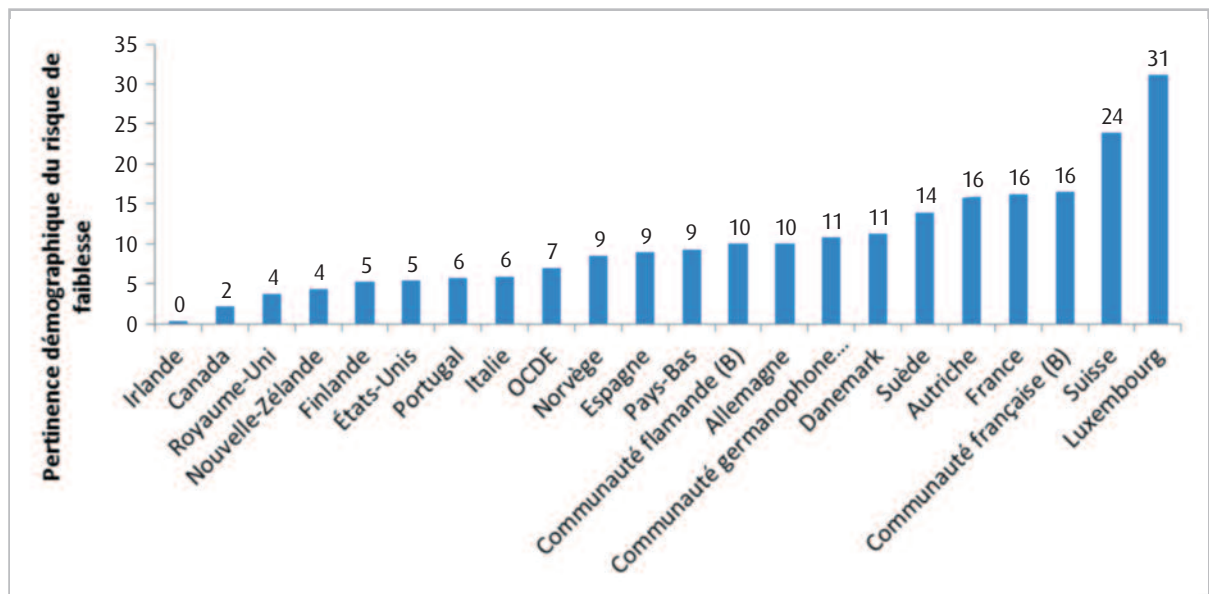
Figure 1. **Risque relatif pour les immigrés de faire partie du quartile inférieur des performances en mathématiques – PISA 2012 (OECD, 2013).** NS: non significatif.



La notion de risque relatif ne tient cependant pas compte de l'importance quantitative du groupe à l'étude. En Finlande, par exemple, les élèves d'origine immigrée sont un sous-groupe particulièrement à risque (figure 1). Cependant, ils constituent

une proportion relativement peu importante (3%) de la population des élèves de 15 ans. Le risque est donc élevé, mais le nombre de personnes concernées est plus limité que dans d'autres pays. C'est la raison pour laquelle l'OCDE a calculé la « pertinence démographique » de la vulnérabilité accrue des élèves immigrés (figure 2). «La pertinence démographique correspond à la réduction de la prévalence dans l'ensemble de la population qui résulterait de la réduction de la prévalence dans un sous-groupe (les élèves issus de l'immigration, par exemple) au même niveau que la prévalence dans le groupe non vulnérable (les élèves autochtones, en l'espèce)» (OCDE, 2011c, p. 488).

Figure 2. **Pertinence démographique du risque de faiblesse en mathématique des élèves d'origine immigrée – PISA 2012 (OECD, 2013).**



Ainsi, si des politiques éducatives parvenaient à ramener le niveau de risque de faiblesse en mathématiques qu'encourent les élèves immigrés au niveau du risque encouru par les élèves autochtones, la prévalence du risque au niveau de l'ensemble des élèves de 15 ans diminuerait plus qu'ailleurs dans les pays européens où le français est une des langues d'enseignement.

Conclusion

Les comparaisons entre systèmes éducatifs ne sont pas aisées, même si on se limite aux systèmes éducatifs qui scolarisent leurs élèves en français. Ces systèmes, qui ont été évoqués jusqu'ici, ne sont en effet pas identiques quant à l'organisation de leur enseignement et à l'utilisation du français dans le contexte scolaire. Le Luxembourg enseigne à tous les élèves dans trois langues différentes (français, allemand et luxembourgeois), dans des proportions variables, selon le niveau d'enseignement et les matières enseignées (Martin, Dierendonck, Meyers et Noesen, 2008)¹⁸. La France offre un système uniquement en français. La Belgique connaît en fait trois systèmes éducatifs, organisés de manière autonome, dans une seule langue (en dehors des possibilités d'immersion linguistique), sur des territoires parfois distincts, parfois identiques, comme à Bruxelles. Cette situation existe aussi au Canada, à travers ses différentes provinces (Churchill, 2003). La Suisse applique un principe de territorialité qui fixe la langue d'enseignement par commune lorsque le canton est plurilingue (à l'exception de Berne, la capitale fédérale).

Si l'on élargit l'étude à l'ensemble des pays qui s'identifient comme francophones, à l'intérieur de l'OIE, il devient très complexe d'effectuer des analyses robustes, d'autant qu'une majorité de pays ayant effectivement en usage le français au sein de leurs établissements scolaires, en particulier les pays francophones d'Afrique subsaharienne, ne participent pas (encore) au PISA, pour des raisons essentiellement financières.

Une analyse plus poussée serait donc bienvenue de manière à mieux apprécier, par exemple, la place qui est faite aux élèves dont le français n'est pas la langue maternelle, notamment en matière de renforcement, ou les disparités qui peuvent exister entre les différents groupes linguistiques non francophones dans chaque système. Cette analyse serait avantageusement enrichie par la prise en compte de certains pays d'Afrique où le français est langue d'enseignement uniquement, que celui-ci soit introduit de manière très précoce ou plus tardive, après une première étape en langue(s) nationale(s). Pour réaliser ce type d'étude, il faudrait pouvoir mettre en rapport les résultats obtenus au PASEC et au PISA, ce qui est aujourd'hui encore impossible, en raison des méthodes et des objectifs très différents de ces deux études à large échelle.

18. Voir à ce sujet, pour plus de précisions, le site du ministère luxembourgeois de l'Éducation nationale, de l'Enfance et de la Jeunesse (<http://www.men.public.lu/fr/systeme-educatif/langues-ecole-luxembourgeoise/index.html>) consulté le 21 août 2014.

Références bibliographiques

- ADAMS, R. et WU, M. (dir.) (2002). *PISA 2000 Technical Report*. Paris : OCDE.
- CAMPBELL, J. R., KELLY, D. L., MULLIS, I. V. S., MARTIN, M. O. et SAINSBURY, M. (2001). *Framework and Specifications for PIRLS Assessment 2001* (2^e éd.). Chestnut Hill, MA : Boston College.
- CENTRE INTERFÉDÉRAL POUR L'ÉGALITÉ DES CHANCES/CENTRE FÉDÉRAL MIGRATION (2012). *Migrations et populations issues de l'immigration en Belgique*. Chapitre 2. Les flux d'immigration et d'émigration depuis et vers la Belgique. Bruxelles : Centre interfédéral pour l'égalité des chances/Centre fédéral migration. Récupéré le 21 août 2014 de http://www.diversite.be/diversiteit/files/File/migratiedossier/Ch2.FluxMigrationLegale.STATDEM.dec2012.FR_SANS%20RENVOIS.pdf
- CHURCHILL, S. (2003). *L'enseignement des langues et l'identité civique canadienne face à la pluralité des identités des Canadiens. Guide pour l'élaboration des politiques linguistiques éducatives en Europe. De la diversité linguistique à l'éducation plurilingue*. Strasbourg : Conseil de l'Europe, Division des Politiques linguistiques.
- CUSO, R. (2007). SAQMEQ, PASEC, PISA... : les enquêtes sur les acquis des élèves et les politiques de mondialisation. Dans F. Cros et A. Bon (dir.), *Les évaluations au niveau international : impacts, contradictions et incertitudes* (p. 83-95). Sèvres : CIEP.
- DEMEUSE, M. (2012). Comparer les systèmes éducatifs européens. Une entreprise complexe mais indispensable. *Diversité (Ville École Intégration)*, 169 (juillet), 39-44.
- KYLLONEN, P. C. et BERTLING, J. P. (2014). Innovative questionnaire assessment methods to increase cross-country comparability. Dans L. Rutkowski, M. von Davier et D. Rutkowski (dir.), *Handbook of International Large-Scale Assessment Data Analysis*. Londres : Chapman and Hall/CRC.
- MARTIN, R., DIERENDONCK, C., MEYERS, C. et NOESEN, M. (2008). *La place de l'école dans la société luxembourgeoise de demain. Vers de nouveaux modèles de fonctionnement du système éducatif*. Bruxelles : De Boeck.
- MINISTÈRE DE L'ÉDUCATION DU NOUVEAU-BRUNSWICK (2012). PISA 2009. *Résultats de la performance des élèves du Nouveau-Brunswick francophone*. Fredericton : Ministère de l'Éducation du Nouveau-Brunswick, Direction de la mesure et de l'évaluation.
- MONSEUR, C. et DEMEUSE, M. (2004). Quelques réflexions méthodologiques à propos des enquêtes internationales dans le domaine de l'éducation. *Politiques d'éducation et de formation : Analyses et comparaisons internationales*, 11, 37-54.

- MULLIS, I. V. S., KENNEDY, A. M., MARTIN, M. O. et SAINSBURY, M. (2006). *PIRLS 2006 Assessment Framework and Specifications* (2^e éd). Chestnut Hill, MA: Boston College.
- OCDE (2001). *Connaissances et compétences, des atouts pour la vie. Premiers résultats de PISA 2000*. Paris: OCDE.
- OCDE (2004). *Apprendre aujourd'hui, réussir demain. Premiers résultats de PISA 2003*. Paris: OCDE.
- OCDE (2007). *Les compétences en sciences, un atout pour réussir. Volume 2: Données*. Paris: OCDE.
- OCDE (2011a). *Résultats du PISA 2009: Savoirs et savoir-faire des élèves. Performance des élèves en compréhension de l'écrit, en mathématiques et en sciences* (vol. I). Paris: OCDE.
- OCDE (2011b). *Résultats du PISA 2009: Élèves en ligne. Technologies numériques et performance* (vol. VI). Paris: OCDE.
- OCDE (2011c). *Regards sur l'éducation 2011: Les indicateurs de l'OCDE*. Paris: OCDE.
- OCDE (2014). *Résultats du PISA 2012: Savoirs et savoir-faire des élèves. Performance des élèves en mathématiques, en compréhension de l'écrit et en sciences* (vol. I). Paris: OCDE.
- OECD/UNESCO-UIS (2003). *Literacy Skills for the World of Tomorrow: Further Results from PISA 2000*. Paris: OECD.
- OECD (2005a). *PISA 2003 Technical Report*. Paris: OECD.
- OECD (2005b). *PISA 2003 Data Analysis Manual*. Paris: OECD.
- OECD (2013). *PISA 2012 Results: Excellence through Equity. Giving Every Student the Chance to Succeed* (vol. II). Paris: OECD.
- ROSENBLATT, L. M. (1978). *The Reader, the Text, the Poem: The Transactional Theory of the Literacy Work*. Carbondale, IL: Southern Illinois University Press.
- ROSENBLATT, L. M. (2004). The transactional theory of reading and writing. Dans R. B. Ruddel et N. J. Unrau (dir.), *Theoretical Models and Processes of Reading* (5^e éd.). Newark, DE: International Reading Association.
- SOUSSI, A., BROI, A. M., MOREAU, J. et WIRTHNER, M. (2004). *PISA 2000: La littérature dans quatre pays francophones. Les résultats des jeunes de 15 ans en compréhension de l'écrit*. Neuchâtel, Suisse: Institut de recherche et de documentation pédagogique (IRDP).
- WARM, T.A. (1989). Weighted Likelihood Estimation of Ability in Item Response Theory. *Psychometrika*, 54(3), 427-450.

- WILKINSON, I. A. G. et SON, E. H. (2011). A dialogic turn in reseach on learning and teaching to comprehend. Dans M. L. Kamil, P. D. Pearson, E. Birr Moje et P. Afflerbach (dir.), *Handbook of Reading Research* (vol. IV, p. 359-387). New York et Londres : Routledge.
- YANG, Y., HARKNESS, J. A., CHIN, T.-Y. et VILLAR, A. (2010). Response styles and cultures. Dans J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. Lyberg, P. Ph. Mohler, P.-E. Pennell et T. W. Smith (dir.), *Survey Methods in Multinational, Multiregional and Multicultural Contexts* (p. 203-233). Hoboken, NJ : Wiley & Sons.