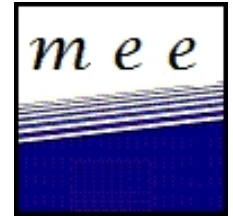


Mesure et évaluation en éducation



L'évaluation en classe Des politiques aux pratiques

Dany Laveault

Volume 32, Number 3, 2009

URI: <https://id.erudit.org/iderudit/1024929ar>

DOI: <https://doi.org/10.7202/1024929ar>

[See table of contents](#)

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)

2368-2000 (digital)

[Explore this journal](#)

Cite this article

Laveault, D. (2009). L'évaluation en classe : des politiques aux pratiques.

Mesure et évaluation en éducation, 32 (3), 1–22.

<https://doi.org/10.7202/1024929ar>

Article abstract

This paper seeks to develop a framework to study the successful implementation of educational policies and their expected outcomes on classroom assessment. Research in the field has contributed to identify several new issues and to raise important questions on how to link policies and classroom assessment practices. More precisely, this paper will try to identify the principles at play for an adequate transition from policies to classroom assessment practices and the best methods to achieve that. The paper concludes that theoretical and professional knowledge mobilization is at the basis of an appropriate junction of policies and practices. It questions the efficiency of hierarchical models of transmission between policies and practice. It also hypothesizes that it might be more advantageous to study how policies and classroom practice may work together and construct each other within a paradigm of balanced assessment and reform professionalization of educational systems.

L'évaluation en classe : des politiques aux pratiques

Dany Laveault

Université d'Ottawa

MOTS CLÉS : Évaluation en salle de classe, politique éducationnelle, réformes éducatives

Le but de cet article est de fournir un cadre de référence pour étudier l'implantation réussie des politiques et leurs effets attendus en matière d'évaluation en classe. Les recherches dans ce domaine soulèvent de nombreuses questions sur l'articulation des politiques et des pratiques en évaluation en plus d'identifier de nouvelles problématiques. Plus particulièrement, cet article cherchera à déterminer quels principes gouvernent le passage réussi des politiques aux pratiques en matière d'évaluation et comment y parvenir. L'article conclut que la mobilisation des savoirs théoriques et d'expérience est au cœur de la jonction réussie des politiques et pratiques. Il s'interroge sur l'efficacité des modèles hiérarchiques de transmission entre politiques et pratiques et émet l'hypothèse qu'il pourrait être plus profitable d'étudier comment politiques et pratiques peuvent faire œuvre commune et se construire dans un processus d'évaluation équilibrée et de professionnalisation des réformes des systèmes éducatifs.

KEY WORDS: Classroom assessment, educational policies, educational reforms

This paper seeks to develop a framework to study the successful implementation of educational policies and their expected outcomes on classroom assessment. Research in the field has contributed to identify several new issues and to raise important questions on how to link policies and classroom assessment practices. More precisely, this paper will try to identify the principles at play for an adequate transition from policies to classroom assessment practices and the best methods to achieve that. The paper concludes that theoretical and professional knowledge mobilization is at the basis of an appropriate junction of policies and practices. It questions the efficiency of hierarchical models of transmission between policies and practice. It also hypothesizes that it might be more advantageous to study how policies and classroom practice may work together and construct each other within a paradigm of balanced assessment and reform professionalization of educational systems.

PALAVRAS-CHAVE: Avaliação na sala de aula, política educativa, reformas educativas

Este artigo pretende fornecer um enquadramento para estudar a implementação bem sucedida das políticas educativas e dos efeitos pretendidos em termos de avaliação na sala de aula. Neste domínio, as investigações salientam numerosas questões sobre a articulação entre as políticas e as práticas de avaliação, bem como identificam novas problemáticas. Mais particularmente, este artigo procurará determinar quais os princípios e os métodos que garantem uma transição adequada das políticas às práticas em matéria de avaliação. O artigo conclui que é a mobilização de saberes teóricos e da experiência profissional que está na base da adequação bem sucedida entre as políticas e as práticas. Interroga-se ainda neste artigo a eficácia dos modelos hierárquicos de transmissão entre políticas e práticas e formula-se a hipótese segundo a qual poderia ser mais útil estudar como as políticas e as práticas podem trabalhar em comum, construindo-se num processo de avaliação equilibrada e de profissionalização das reformas dos sistemas educativos.

Note de l'auteur – L'auteur tient à remercier les personnes suivantes pour leurs commentaires et leurs suggestions : Linda Allal (Université de Genève), Lucie Mottier Lopez (Université de Genève) et Léopold Paquay (Université catholique de Louvain). Merci également à Léopold Paquay pour avoir appuyé et supporté la préparation de ce numéro thématique à titre de corédacteur invité. Dans ce texte, le masculin est utilisé au sens générique : il comprend aussi bien enseignant qu'enseignante. Toute correspondance peut être envoyée à Dany Laveault, Faculté d'éducation, Université d'Ottawa, Pavillon Lamoureux, bureau 474, 145, rue Jean-Jacques-Lussier, B.P. 450, succ. A, Ottawa, Ontario K1N 6N5, téléphone (613) 562-5800, poste 4118, télécopieur (613) 562-5146, ou par courriel à l'adresse suivante : [dlaveault@videotron.ca].

Au cours des dix dernières années, plusieurs réformes ont eu lieu dans les systèmes d'éducation des pays francophones : une refonte en profondeur des programmes d'étude, une diversification des modalités d'évaluation et enfin, l'apparition de standards « nationaux » et d'évaluations externes. Ces réformes ont entraîné de nombreux changements en matière d'évaluation en salle de classe, soit sur le plan des pratiques d'évaluations formative et sommative conçues et administrées par l'enseignant, seul ou avec ses collègues. À ces évaluations en salle de classe, il convient de plus en plus d'ajouter les évaluations externes préparées par les différentes juridictions scolaires qui, même si elles ne peuvent être qualifiées d'évaluation en classe, peuvent influencer les pratiques d'évaluation des enseignants et leurs choix en matière d'enseignement. Comme le font remarquer Mottier Lopez et Laveault (2008), la revue n'a pas toujours fait écho à l'émergence de ce nouveau contexte, ni à l'apparition de problématiques d'évaluation à la fois similaires et différentes dans chaque pays francophone. Ce numéro thématique, intitulé *L'évaluation en classe : des politiques aux pratiques*, cherche à dresser un bilan de l'impact qu'ont eu les changements récents et à déterminer comment les nouvelles politiques parviennent à influencer les pratiques d'évaluation dans une perspective d'amélioration.

Une telle question pour les pays francophones présente de nombreuses difficultés. La documentation en français sur le sujet est rare, dispersée et pas toujours indexée. Nous ne disposons pas, comme dans la documentation scientifique anglo-saxonne, de revues spécialisées en évaluation qui traitent de ces questions. Le présent numéro répond à un besoin en français en regroupant une série d'articles sur l'expérience particulière de la Belgique (communauté française), de la Suisse romande (cantons de Vaud, Berne et Genève), du Canada (province de Québec), du Luxembourg et de la France. Chaque article a été préparé par deux ou trois auteurs de chaque pays francophone invités à aborder les questions suivantes :

- Quelles sont les politiques nationales ou régionales en matière d'évaluation ? Comment ont-elles évolué, comment se sont-elles implantées dans le réseau scolaire, dans la pratique en salle de classe ?

- Quelle est l'importance des évaluations internationales dans l'implantation des politiques et le changement des pratiques? Existe-t-il des formes d'évaluation externes – gouvernementales – pour valider l'évaluation effectuée par l'enseignant ou pour suivre la mise en place du curriculum?
- Comment l'évaluation des apprentissages est-elle perçue par les enseignants? par les parents? par les élèves? À quoi sert-elle principalement?
- Quel impact ont les politiques nationales ou régionales sur l'évaluation en salle de classe? sur la réussite des élèves? sur la perception que les enseignants se font de l'évaluation et de leur propre rôle? Quel espace les politiques ou les directives aux enseignants laissent-elles au jugement professionnel?
- Comment les politiques sont-elles évaluées? Existe-t-il un suivi dans l'application des politiques d'évaluation? Comment les politiques d'évaluation s'inscrivent-elles dans les réformes des programmes d'étude?
- Existe-t-il des directives strictes sur la notation ou sur la préparation du livret scolaire ou bulletin? Comment politiques, évaluation en salle de classe et systèmes de notations interagissent-ils?

Bien entendu, toutes ces questions ne sont pas abordées dans chaque article. Dans plusieurs cas, les recherches ou la documentation pour y répondre sont incomplètes ou tout simplement absentes. Dans chaque pays, les changements intervenus dans le domaine de l'évaluation se situent dans la continuité du développement des politiques éducatives et se font souvent en réaction aux politiques – ou à l'absence de politiques – antérieures. De plus, la gouvernance des systèmes éducatifs varie d'un pays à l'autre. Elles sont proprement nationales, dans le cas du Luxembourg et de la France. Dans le cas de fédérations, comme la Suisse, le Canada et la Belgique, ce sont les régions, cantons ou provinces qui sont responsables du système éducatif.

Historique des réformes et situation dans les pays francophones

Dans chacun des articles, l'interprétation des pratiques actuelles nécessite un rappel du contexte historique pour bien comprendre le lien entre l'établissement de politiques et leur application en salle de classe. Cette diversité des contextes nationaux ou locaux fournit un échantillon circonstanciel de diverses modalités d'implantation des politiques et de leurs conséquences plus

ou moins grandes sur les pratiques d'évaluation et les résultats d'apprentissage. À quoi bon en effet établir des politiques, les changer ou chercher à les améliorer si leur impact sur les pratiques n'est pas suivi des effets voulus ?

La connaissance du contexte récent peut contribuer à mieux comprendre les défis posés par les politiques nouvelles et tous les articles de ce numéro thématique y ont accordé une place de choix. Les pays francophones ont, à cet égard, partagé un certain nombre de défis comparables et d'exigences nouvelles (Laveault, sous presse). En voici une brève description.

Dans tous les pays, deux choses semblent acquises. La première est la conviction largement répandue du rôle constructif que l'évaluation formative peut jouer en soutien à l'apprentissage. La seconde est l'intérêt pour des objets d'apprentissage qui favorisent l'intégration et la mobilisation des connaissances, des capacités et des attitudes des élèves dans des familles de situations complexes et authentiques. Il s'agit ici des compétences comme objectifs et résultats d'apprentissage.

L'évaluation formative et l'approche par compétences relèvent toutes deux de finalités communes : élever les standards d'apprentissage et accroître la réussite du plus grand nombre d'élèves. Si, dans leur formulation la plus globale, ces finalités remportent partout l'adhésion, les moyens d'y parvenir font l'objet de nombreuses discussions sur le plan scientifique et de controverses tant sur le plan politique qu'en ce qui a trait à la pratique. Le dicton anglais

« The devil is in the details » s'applique bien ici. L'opérationnalisation de ces bonnes intentions ne va pas d'elle-même et les modalités d'implantation ne font pas l'unanimité.

Dans le domaine de l'évaluation formative, Allal et Mottier Lopez (2005) font état des avancées des pays francophones sur le plan théorique, mais déplorent en même temps le peu de recherches empiriques sur le sujet. Nous pouvons probablement affirmer la même chose de l'évaluation de compétences, un domaine de recherche en plein essor. Conjuguées simultanément, l'évaluation formative et l'évaluation de compétences posent quant à elles de nombreux défis.

Les réformes des programmes d'étude ont donc été l'occasion non seulement de proposer de nouvelles pratiques, mais aussi d'interroger à nouveau le savoir disciplinaire dans le domaine de l'évaluation scolaire. La prudence et la méfiance du corps enseignant face aux objectifs poursuivis par les réformes ne sont donc pas toutes à ranger du côté des facteurs de résistance aux

changements. Les difficultés de mise en œuvre de ces réformes ont contribué à remettre en question certaines définitions et notions et à les actualiser dans de nouveaux contextes. Enfin, ces réformes ont permis de prendre conscience qu'en élevant la complexité des situations d'évaluation, le caractère indéterminé de l'évaluation progressait en proportion, nécessitant l'exercice plus approfondi du «jugement professionnel», chaque fois qu'une solution générale ne peut être envisagée (Lafortune & Allal, 2008; Laveault, 2008).

Fullan (2009) distingue plusieurs périodes dans l'établissement de réformes en éducation. La première période, entamée dans les années 1960, tenait pour acquis que les bonnes idées en éducation se répandent d'elles-mêmes et que leur adoption ne nécessite pas d'intervention précise. Or, ce ne fut pas le cas. La période de réformes qui suivit (1970-1980) part du constat d'échec précédent et entreprend la diffusion active d'innovations pédagogiques. Le terme «implantation» apparaît alors. La période suivante (1980-1996) se caractérise, selon Fullan (2009), par des pressions accrues de réformes qui se sont traduites par des innovations de surface. Les changements introduits ne se sont pas avérés durables dans la majorité des cas. La période 1997-2002 se caractérise par des changements à grande échelle impliquant simultanément de larges pans des systèmes éducatifs : programme d'études, évaluation des élèves à l'interne mais aussi à l'externe, changements sur le plan des organisations scolaires, etc. La période la plus récente (2002 à nos jours) se caractérise par une «professionnalisation de la réforme des systèmes», une expression que Fullan emprunte à Barber (2008, cité par Fullan, 2009). Cette période coïncide avec l'étude des facteurs qui favorisent une amélioration durable des effets escomptés des réformes.

Dans l'esprit de ce mouvement de professionnalisation de la réforme des systèmes, le but de cet article d'introduction n'est pas de résumer ou de présenter chaque contribution des pays francophones. Le caractère particulier de chaque juridiction rend risquée toute généralisation des solutions et des initiatives prises dans chaque pays. Les résultats des recherches les plus récentes nous mettent en garde : si l'on arrive à déterminer des facteurs importants d'amélioration, la façon de les opérationnaliser prend des formes très variées d'un pays à l'autre (Allal, 2007a; Atkin, 2007; Fullan, 2009). Le but de cet article est plutôt de fournir un cadre de référence pour l'étude de l'implantation réussie des politiques et de leurs effets attendus en matière d'évaluation. Peut-on dégager un enseignement et tirer des leçons des expériences diverses sur la manière d'implanter une politique, d'en susciter l'adhésion et l'engagement des différents intervenants? Plusieurs chercheurs

sont de cet avis. L'étude approfondie de la mise en place de réformes éducatives dans les pays francophones et anglo-saxons a permis d'établir un certain nombre de constantes et de principes de qualité en matière de réformes en éducation. Quels principes gouvernent le passage réussi des politiques aux pratiques en matière d'évaluation et comment y parvenir ? Voilà la question que nous tâcherons d'aborder dans cet article.

Améliorer l'évaluation ou les évaluateurs ?

Parmi les facteurs d'efficacité reconnus par les plus récentes études sur les réformes scolaires figure l'importance de pouvoir compter sur une profession enseignante de haute qualité (Hargreaves, 2009, cité par Fullan, 2009), sur un nouveau professionnalisme (Barber, 2008, 2009, cité par Fullan, 2009) et sur un enseignement livré par des professionnels formés adéquatement (Neumans, 2009, cité par Fullan, 2009). Il semble que les réformes dans le domaine de l'évaluation n'échappent pas à la question suivante : les politiques peuvent-elles vraiment améliorer l'évaluation en salle de classe sans améliorer les compétences des évaluateurs ?

Si l'ensemble des politiques recensées dans ce numéro spécial supportent les pratiques d'évaluation formative, celles-ci ne sont pas toujours suivies des effets voulus, même si des efforts considérables ont été investis dans la formation des enseignants. Une étude de généralisabilité de Heritage, Kim, Vendlinski et Herman (2009) démontre que le jugement des enseignants est particulièrement fiable lorsqu'il s'agit de relever les principes clés à l'œuvre dans une tâche, de même que lorsqu'il s'agit du degré de compréhension des enseignants envers la tâche. Par contre, le jugement des enseignants est moins fiable ou assuré lorsqu'il s'agit de planifier l'étape suivante de l'enseignement ou de l'apprentissage à la suite de la découverte d'une difficulté chez l'élève.

La composante de l'évaluation formative la plus difficile pour l'enseignant ne semble donc pas résider dans sa visée d'information, telle que décrite par Allal (1979) mais dans sa visée d'adaptation. Heritage et al. (2009) affirment qu'il est de première importance de développer la capacité de l'enseignant à savoir ce qu'il doit enseigner par la suite et comment il doit adapter son enseignement en fonction des résultats obtenus. Le passage de l'interprétation des résultats de l'élève à l'application de stratégies efficaces de régulation ne va pas de soi, ni ne se réalise spontanément (Allal, 2007b ; Laveault, 2007). Tant que ce passage de l'interprétation à l'action ne se fera pas mieux,

l'évaluation formative ne pourra donner sa pleine mesure. Ceci interpelle la formation des enseignants à l'évaluation formative. À quoi bon former des enseignants aux techniques d'évaluation formative, si une fois adéquatement informés des difficultés des élèves, ils se trouvent démunis quant à la connaissance de la progression de ces apprentissages et sur la façon dont ils doivent poursuivre leur action ?

Selon Shepard (2009) et Perie, Marion et Gong (2009), les responsables politiques devraient investir dans le développement professionnel des enseignants. Celui-ci ne devrait pas porter uniquement sur l'apprentissage de techniques d'évaluation formative, mais inclure l'apprentissage des méthodes les plus productives de recours au *feedback* et à l'autoévaluation (Shepard, 2009). James (2006) fait ressortir l'importance d'aligner l'évaluation non seulement sur le curriculum, mais aussi sur les théories d'apprentissage :

«

implementing assessment for learning/formative assessment may require a teacher to rethink what effective learning is, and his or her role in bringing it about. Similarly, a change in their view of learning is likely to require assessment practice to be modified» (p. 48).

Une forme de consensus semble se développer autour du principe selon lequel l'information apportée par l'évaluation formative, tout aussi valide qu'elle puisse être, est de peu d'utilité pratique si les enseignants ignorent ce qu'ils doivent faire lorsque les étudiants sont incapables d'appréhender un concept important (Shepard, 2009). Dans les pays francophones, cette problématique a été abordée non seulement à travers les théories d'apprentissage générales, mais également à travers les apports de la didactique des disciplines pour concevoir les programmes et analyser les difficultés des élèves. Pour Bain (1992, p. 7), les docimologues et les didacticiens «se trouvent à traiter d'un même concept, central et essentiel à la fois pour l'évaluation formative et pour l'enseignement (et donc pour les apprentissages visés) : celui de régulation». Pour Bodin (1992, p. 15), «c'est bien la marque du succès de l'évaluation [...] que de voir certaines pratiques issues de la problématique évaluative se résorber dans la didactique». Bain (1992, p. 10) déplore déjà que sans efforts d'intégration entre évaluation, didactique et théories d'apprentissage «l'évaluation formative reste pour longtemps au niveau d'un bricolage éclairé, plus ou moins lié à ses objectifs didactiques».

Des évaluations conséquentes qui soutiennent l'apprentissage

Le cas particulier de l'application de l'évaluation formative décrit précédemment est assez caractéristique des difficultés qui accompagnent le passage des politiques aux pratiques. En effet, l'évaluation formative décrit d'abord une finalité : utiliser l'information recueillie sur l'apprentissage de l'élève à des fins de régulation de l'apprentissage ou de l'enseignement. Appuyer davantage sur cette finalité, la réaffirmer sur le plan des politiques d'évaluation n'est pas de nature à elle seule à entraîner des changements quant aux résultats d'apprentissage.

Dans le cas de l'évaluation formative, le passage réussi des politiques aux pratiques se ramène à une question de « validité de conséquence » (Messick, 1988). L'évaluation formative, si sa finalité est bien le soutien à l'apprentissage, a-t-elle les conséquences attendues sur les résultats d'apprentissage de l'élève ? Mais en quoi alors l'évaluation formative serait-elle différente, en termes de ses conséquences, de l'évaluation sommative ?

Tourneur (1985) évoquait « le déséquilibre extraordinaire en faveur de l'évaluation formative » (p. 7) et soulignait le tort injustifié que la réaction négative à l'encontre des examens traditionnels faisait subir à l'évaluation sommative. Il déplore la confusion conceptuelle qui en résulte :

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 » (p. 7).

Selon Mottier Lopez et Laveault (2008), le développement de la recherche sur l'évaluation formative et l'évolution des pratiques devra sortir l'évaluation formative de l'aporie qui l'oppose à l'évaluation sommative. Cette opposition confond les finalités de l'évaluation avec la démarche employée (De Ketele, 2006 ; Figari, 2006 ; Harlen, 2005). L'évaluation formative décrit une finalité (supporter l'apprentissage), alors que l'évaluation sommative décrit une démarche (comment établir un bilan).

Dans le contexte récent, il ne s'agit plus de déterminer si l'évaluation sommative a ou pourrait avoir des effets négatifs sur l'apprentissage, mais de concevoir des moyens qui font en sorte que les méthodes qui consistent à établir des bilans des apprentissages puissent aussi contribuer à l'apprentissage

des élèves. Il s'agirait alors, selon les termes de Kennedy, Sang Chan, Fok et Yu (2008), de rendre le sommatif formatif. Comment pourrait-il en être autrement? Peut-on demeurer sans rien faire pour réduire les effets négatifs que peut avoir parfois l'évaluation sommative sur l'apprentissage des élèves, surtout lorsque de tels effets risqueraient d'anéantir les effets positifs de l'évaluation formative? Selon 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 sur le plan cognitif ou motivationnel, tout en reconnaissant évidemment sa fonction première d'attestation des apprentissages réalisés. Les entretiens maître-élève, avec ou sans la présence de parents, peuvent être conçus et élaborés pour apporter un soutien et non seulement pour dresser un bilan (p. 18).

De l'avis de Stiggins, Arter, Chappuis et Chappuis (2004), des bilans représentatifs des apprentissages des élèves peuvent s'inscrire également, même si moins fréquemment et de manière plus globale, dans un plan de soutien à l'apprentissage de l'élève. De plus, selon Atkin (2007, p. 55) le rôle de l'évaluation sommative ne se limite pas qu'à informer l'enseignant sur les résultats de l'élève: «while teachers inevitably play a central role in formative assessments, their summative judgments about the quality of student work are also significant. Furthermore they are policy-relevant».

Le rôle des évaluations externes

Les grandes enquêtes sur les niveaux de rendement des élèves, qu'elles soient internationales, nationales ou entreprises par des juridictions locales, remettent le rôle de bilan des apprentissages à l'avant-scène, non seulement pour tout ce qui concerne la régulation directe des apprentissages des élèves, mais aussi pour la régulation des systèmes éducatifs avec leurs conséquences indirectes sur le rendement des élèves. En effet, comme le font remarquer Black et William (1998, p. 148): «Teachers should play a far greater role in contributing to summative assessments for accountability. One strong reason for giving teachers a greater role is that they have access to the performance of their pupils in a variety of contexts and over extended periods of time.»

Avant même d'exercer leur rôle de régulateurs des systèmes éducatifs, les épreuves externes contribuent à modifier l'objet même qu'elles prétendent étudier. En effet, leur contenu et leur format ne sont pas sans influencer les choix des enseignants. Dans le meilleur des scénarios, ces épreuves constituent une référence concrète sur les objectifs des programmes. Elles permettent

d'améliorer l'alignement de l'enseignement et des évaluations en classe sur les objectifs des programmes d'études et d'accroître ainsi leur validité de contenu. Dans les pires scénarios cependant, ces épreuves sont considérées au premier niveau et l'enseignement se concentre sur la réussite du test et non sur celle des attentes du programme («*teach to the test*»). Ce peut être le cas lorsque « les enseignants réutilisent cette épreuve comme exercice formatif en cours d'année. Ainsi, l'examen externe aux établissements scolaires est devenu un des moteurs de la rénovation » (Delory, cité par Paquay, 2007, p. 19).

Avant même de pouvoir utiliser les résultats des épreuves externes à des fins de régulation des systèmes éducatifs, il convient de minimiser les impacts négatifs de leur utilisation et d'en maximiser les impacts positifs. La validité de conséquence de ces épreuves externes n'en sera qu'améliorée. Dans la mesure où un alignement sur les programmes d'étude permet aux enseignants de mieux établir la progression des élèves dans le programme de formation, d'éviter les chevauchements et de préparer adéquatement les élèves aux niveaux suivants, les épreuves externes peuvent contribuer indirectement à améliorer les résultats d'apprentissage. Considérant que la maîtrise des pré-alables est l'un des meilleurs prédicteurs de la réussite scolaire (Vienneau, 2005), toute amélioration de l'alignement des pratiques sur le programme d'études comme conséquence des épreuves externes est hautement souhaitable.

Jusqu'à présent, le passage des politiques aux pratiques a reposé sur un postulat fort : celui que les politiques sont valides et que leur implantation est souhaitable et désirable pour le meilleur bénéfice de tous. L'un des effets inattendus des enquêtes internationales a été de mettre en relief le caractère parfois arbitraire des politiques en matière d'éducation et le caractère imparfait de leur mise en pratique. Celles-ci sont fortement inscrites dans une tradition, dans une histoire ainsi que dans un contexte culturel et sociopolitique particulier. Elles ne reposent pas toujours sur une interprétation valable des modèles théoriques de l'apprentissage, de l'enseignement ou de l'évaluation.

Les tests internationaux (PISA, TIMMS, PIRLS) ont permis de comparer les résultats d'apprentissage produits par différents systèmes d'éducation et de faire ressortir les caractéristiques des systèmes qui produisent les meilleurs rendements. Même en prenant en considération les limites de ces tests internationaux, force nous est d'admettre qu'ils ont contribué à remettre en question non seulement la qualité des interventions éducatives effectuées sur le plan micro (salle de classe), mais aussi sur les plans méso (l'établissement d'enseignement) et macro (gouvernance locale et nationale). Sans tomber dans

une certaine forme de manichéisme selon lequel il y aurait de « bonnes » et de « mauvaises » politiques, les études internationales ont permis de mettre en évidence le fait que la réussite de l'élève comporte une responsabilité collective, engageant l'élève et l'enseignant au premier titre, mais aussi les intervenants à tous les niveaux d'un système d'éducation.

Cette prise de conscience de la responsabilité collective des systèmes éducatifs sur la réussite des élèves a pour principale conséquence d'encourager une régulation plus fréquente à différents niveaux du système éducatif. Si la réussite de l'élève résulte à la fois des interactions qui se produisent en salle de classe, mais aussi des décisions prises à différents niveaux du système éducatif, l'évaluation formative devrait également s'y retrouver de manière plus continue. Selon Perie et al. (2009), des évaluations bilans réalisées fréquemment, entre le début et la fin d'une année scolaire, peuvent jouer un rôle d'évaluation intérimaire (« *interim assessment* ») définie de la façon suivante : « Assessments administered during instruction to evaluate students' knowledge and skills relative to a specific set of academic goals in order to inform policymaker or educator decisions at the classroom, school or districts » (p. 6).

L'évaluation intérimaire se veut formative en permettant une régulation plus fréquente des systèmes éducatifs à tous les niveaux et en intégrant ce type de régulations à des régulations plus fines se produisant en salle de classe. Pour permettre de profiter de ce type de régulations plus fréquentes, il faut disposer d'épreuves qui fournissent des résultats valides à tous les niveaux du système éducatif. Perie et al. (2009) formulent donc un ensemble de critères de qualité dans le développement et l'utilisation de ces épreuves à différentes fins de régulation des systèmes éducatifs. Dans un tel contexte, le rôle des décideurs et des enseignants est davantage celui de consommateurs ou concepteurs éclairés de ces épreuves et d'utilisateurs prudents et compétents des résultats de ces épreuves.

L'introduction d'un niveau intérimaire d'évaluation est-il vraiment de nature à améliorer les résultats d'apprentissage des élèves ? N'y a-t-il pas risque d'une surcharge d'information qui peut être plus paralysante que motivante ? Pour optimiser l'impact de ce genre d'évaluation, il semble important de différencier le degré d'intervention et de soutien. Celui-ci devrait être

inversement proportionnel au rendement des différentes constituantes d'un système éducatif : les écoles et les secteurs les moins performants devraient recevoir plus de soutien (Barber, 2008, cité par Fullan, 2009, p. 104).

Tout comme l'évaluation soutien d'apprentissage s'attarde sur les élèves en difficulté, un bon système d'indicateurs doit permettre de déterminer les écoles et les catégories d'élèves qui ont vraiment besoin d'aide (Laveault, 2009). C'est le cas de la distinction d'établissements scolaires qui présentent des « cas extrêmes », ou en termes statistiques des « *multivariate outliers* », au moyen des indicateurs fournis par ces épreuves externes. Un bon système d'indicateurs doit conduire à plus d'équité et, à cet égard, la comparaison des progrès réalisés par différents groupes d'élèves aux niveaux micro (salle de classe), méso (établissement) et macro (district, conseil scolaire ou autre juridiction locale) est primordiale. La distinction des groupes ou catégories d'élèves qui ne présentent pas de progression revêt toute son importance pour le repérage précoce des populations scolaires à risque. Enfin, les indicateurs de résultats doivent pouvoir être interprétés avec des indicateurs d'input et de processus qui sont à même de donner un sens aux résultats locaux. C'est à de telles conditions que, selon Laveault (2009), l'interprétation des résultats d'épreuves externes permet aux systèmes éducatifs de réagir rapidement et adéquatement aux changements et d'adopter des plans d'intervention plus proactifs et potentiellement plus efficaces en tenant compte des complexités qui devraient être prises en compte dans leur mise en œuvre.

Une évaluation scolaire « équilibrée »

L'évaluation intérimaire joue un rôle complémentaire à l'évaluation-soutien d'apprentissage en salle de classe dans la mesure où elle parvient à combiner des renseignements utiles sur l'apprentissage des élèves et peut apporter des changements à plusieurs niveaux du système éducatif – district, établissement, salle de classe. Évaluations internes et externes peuvent se compléter. Dans un tel cas, Stiggins (2003) parle d'évaluation équilibrée (« *balanced assessment* »).

Une évaluation équilibrée cumule les avantages de l'évaluation en salle de classe pour réguler l'apprentissage et l'enseignement avec ceux de l'évaluation externe à des fins de reddition de compte et d'évaluation de l'efficacité des programmes d'étude (McCloskey & McMunn, 2003). Elle repose sur les principes suivants :

- Aucune forme d'évaluation ne peut satisfaire à tous les besoins d'information de chaque intervenant : une diversité est nécessaire pour évaluer l'élève de façon à assurer sa réussite ;
- Alors que l'évaluation en classe intervient pour supporter l'apprentissage, le rôle de l'évaluation externe est d'estimer le rendement à tous les niveaux du système éducatif par agrégation des résultats ;
- La communication des résultats doit s'effectuer en tenant compte du destinataire. L'évaluation en classe informe les élèves à leur propre sujet et les aide à se fixer des cibles individuelles d'amélioration. L'évaluation externe informe les différents intervenants quant à l'impact et à l'efficacité de leurs choix et de leurs décisions sur la réussite des élèves à tous les niveaux du système éducatif ;
- L'enseignant doit être formé à utiliser une variété de stratégies d'évaluation et à choisir celles qui conviennent le mieux selon les situations. L'enseignant est alors partie prenante et non seulement un exécutant des épreuves externes.

Quelles politiques et quelles pratiques ?

Pour améliorer les systèmes éducatifs, il faut être en mesure d'aller au-delà de la simple description de leurs rendements. Il faut pouvoir expliquer ceux-ci pour éventuellement, non pas en reproduire les conditions, mais les adapter au contexte particulier de chaque système. C'est l'approche « professionnalissante » des réformes éducatives (Fullan, 2009).

Les études de cas multiples des résultats des pays participants aux enquêtes internationales permettent déjà de tirer quelques enseignements quant aux facteurs de succès récurrents de certains pays dont les résultats sont non seulement plus élevés, mais aussi se maintiennent et continuent à s'améliorer. Selon Darling-Hammond et McCloskey (2008), il y a peu d'intérêt à tenter d'expliquer les résultats des pays qui se classent le mieux en ne considérant que les contenus des programmes d'études et les niveaux de rendement attendus dans les standards nationaux. Il importe davantage de savoir comment ces pays s'y prennent pour atteindre des niveaux élevés de rendement :

These analyses reveal that higher-achieving countries teach fewer topics more deeply each year; focus more on reasoning skills and applications of knowledge, rather than mere coverage; and have a more thoughtful sequence of expectations

based on developmental learning progressions within and across domains (p. 264).

Le cas de la Finlande a été particulièrement étudié. 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, prenant en compte le haut niveau de formation du personnel enseignant, insistent moins sur les détails des programmes de formation que sur l'approfondissement de quelques sujets essentiels : «Finland's national core curriculum is a much leaner document, reduced from hundred of pages of highly specific prescriptions to descriptions of a small number of skills and core concepts each year» (Darling-Hammond & McCloskey, 2008, p. 266).

L'état australien du Queensland est souvent mentionné pour illustrer les avantages d'un engagement accru des enseignants dans les systèmes d'évaluation nationaux (Atkin, 2007; Darling-Hammon & McCloskey, 2008). Ce système a été élaboré et raffiné au cours d'une période de 30 ans et a nécessité des investissements importants dans la formation des enseignants à ses débuts. Il ne tient pas pour acquis que le jugement professionnel des enseignants puisse reposer uniquement sur la communication de standards nationaux, ni sur les efforts du système d'éducation à en faire la promotion. Dans ce système, la notion de modération joue un rôle capital : «Moderation as judgment practice [...] involves opportunities for teachers when they use their own judgments of assessment data to integrate these judgments with those of other teachers and in so doing share interpretations of criteria and standards» (Maxwell, 2007, cité par Wyatt-Smith, Klenowski & Gunn, 2010, p. 61).

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 d'acquérir une représentation commune des standards et des indicateurs nécessaires à la reconnaissance des performances qui démontrent l'atteinte de ces standards (Maxwell, 2001). 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 travers une communauté professionnelle d'utilisateurs (Wyatt-Smith, Klenowski & Gunn, 2010). Selon ces auteurs, la modération

sociale des standards exerce également un effet en retour sur la qualité de la pratique : « Such sharing enables teachers to review their teaching and assessment practice, to centre on issue of quality, and where appropriate, modify practice » (p. 72).

Comment opérer les changements ?

Les exemples finlandais et australiens se caractérisent par leur très grande flexibilité et par le caractère professionnel de l'organisation du travail. Huberman (1983), dans son livre intitulé *Comment s'opèrent les changements en éducation : contribution à l'étude de l'innovation*, considère que l'organisation « professionnelle » du travail est plus apte à favoriser l'innovation et le changement qu'une structure « bureaucratique » qui caractérise certains systèmes éducatifs. Il relève plusieurs facteurs de limitation qui rendraient plus difficiles les changements et l'introduction d'innovations en éducation. Ces facteurs sont les suivants :

- *Cloisonnement du personnel et des services.* Les ordres professionnels stimulent l'interdépendance entre membres. Huberman (1983) cite en exemple les agriculteurs et les médecins qui « discutent des idées nouvelles et s'imitent les uns les autres, alors qu'il semble y avoir entre les enseignants peu de communications de nature à susciter les innovations » (p. 35) ;
- *Hierarchie et différences de statut.* Selon Huberman (1983), « la plupart des organisations professionnelles ont un taux d'innovation supérieur à celui des organisations bureaucratiques du fait que l'accent y est davantage mis sur la compétence que sur le rang, du fait aussi de la plus grande souplesse des membres de la profession, de la plus grande clarté des buts et des critères de production... » (p. 35). Dans ce cas-ci, la recherche établit un lien clair entre le caractère professionnel d'une organisation et son potentiel de changement. Dans les systèmes bureaucratiques, «
n'importe qui peut recevoir l'ordre d'adopter quelque chose, mais personne ne peut recevoir l'ordre de créer quelque chose de nouveau » (p. 35). Or, si les politiques comptent favoriser l'apparition de solutions locales et novatrices, elles doivent compter non seulement sur la capacité des enseignants à adopter de nouvelles stratégies, mais à en créer de nouvelles ;

- *Absence de procédures et de formation en vue du changement.* Sans politique claire de développement professionnel, les enseignants n'ont pas accès aux pratiques nouvelles et aux innovations mises en place par leurs pairs. Dans une structure hiérarchique de type bureaucratique, les formations sont données «*top down*», suivant la relation employeur-employés. Dans une profession, les formations sont conçues par et pour les membres en fonction des exigences d'exercice de la profession.

Les systèmes finlandais et australiens fournissent des exemples de l'efficacité de structures professionnelles d'organisation. Elles confirment, chacune à leur manière, l'analyse des systèmes d'éducation d'Huberman (1983) et ne sont pas sans rapport avec le stade «professionnel» du développement des réformes décrit par Fullan (2009).

Conclusion : joindre politiques et pratiques

Les résultats des recherches sur l'évaluation scolaire ont permis de soulever les difficultés des politiques et des pratiques actuelles en plus de relever de nouvelles problématiques. Le suivi de la trajectoire qui conduit des politiques aux pratiques soulève de nombreuses questions. En voici une courte énumération :

- a) La trajectoire est influencée par la nécessaire intégration de la théorie et de la pratique. Les deux catégories de savoir possèdent leurs propres ambiguïtés et leurs terminologies respectives sont le reflet autant que l'une des causes du maintien de ces ambiguïtés ;
- b) La communication de l'impulsion politique se heurte aux limites des canaux de communication et de formation. L'alignement des pratiques d'évaluation sur les programmes d'étude demeure un objectif souhaité, mais dans les faits, difficilement communicable dans sa totalité ;
- c) Le passage des politiques aux pratiques se heurte également aux difficultés que présentent les changements de pratique alimentées par les facteurs de résistance à l'innovation et au changement. Ces facteurs de résistance ne résultent pas uniquement de l'opposition au changement mais aussi de la difficulté d'opérer ce changement sans perte d'efficacité. Une grande partie du savoir pédagogique est un savoir d'expérience difficile à coder et long à acquérir. L'apprentissage de nouvelles pratiques, même lorsqu'il est amplement justifié, peut donner lieu à des réajustements qui peuvent

se traduire, temporairement, par une perte d'efficacité. Un temps d'acquisition et de mûrissement est nécessaire pour que ces nouvelles pratiques donnent leur pleine mesure ;

- d) Certaines politiques – tout comme certaines pratiques enseignantes – ne reposent pas sur un savoir théorique ou d'expérience validé. Certaines vont même dans le sens contraire de ce que nous estimons connaître de l'évaluation et de l'apprentissage. Elles existent du fait de l'habitude ou de la simple volonté des acteurs. Ceci explique la persistance voire l'introduction de politiques ou de pratiques dysfonctionnelles qui peuvent faire obstacle à un changement constructif. La confrontation des volontés se substitue alors à la confrontation des idées et des savoirs. La résolution de ces confrontations est souvent politique et n'a pas ou peu de fondements scientifiques.

La tentation est naturellement forte de copier les politiques des systèmes performants. Or, l'imitation ne donne pas forcément de bons résultats hors du contexte où ces politiques ont été établies (Allal, 2007a; Atkin, 2007). Des éléments de solution ont été apportés ici et là. D'une certaine manière, il y a eu, au cours de la dernière période des réformes scolaires, une accumulation importante de nouvelles données et de nouvelles connaissances. La difficulté se pose maintenant de mobiliser ces différents savoirs théoriques et pratiques pour assurer non pas le passage réussi des politiques aux pratiques, qui relève d'une logique de l'implantation, mais la coordination efficace des politiques et des pratiques dans le but d'améliorer les résultats d'apprentissage.

La « professionnalisation des réformes » suppose que les systèmes éducatifs s'éloignent progressivement d'un paradigme d'implantation pour se rapprocher d'un paradigme de jonction ou, si l'on préfère, de construction conjointe. La jonction réussie des politiques aux pratiques soulève deux questions d'importance : sur quoi devraient porter les politiques en matière d'évaluation et comment celles-ci peuvent avoir un impact réel sur l'amélioration des pratiques ? Le « quoi » pose la question de la validité de ces politiques, tant du point de vue d'un savoir théorique, ce qui implique une mobilisation des connaissances, que du point de vue des pratiques, ce qui ne peut se faire sans tenir compte du savoir d'expérience et de l'expertise des enseignants sur le terrain. Tout comme Paquay (2007, p. 16) l'a fait pour l'implantation des programmes d'études, le « comment » nécessite de « dépasser la question habituelle en évaluation qui cherche à déterminer dans quelle mesure les enseignants appliquent-ils les politiques ».

L'évaluation des élèves, lorsqu'elle s'accompagne de pratiques réflexives sur l'interprétation des résultats, constitue un instrument formidable de régulation des pratiques d'enseignement à tous les niveaux. En favorisant le développement de compétences individuelles et collectives en matière d'évaluation, les politiques contribuent indirectement au développement professionnel continu des enseignants et à la création d'un cercle vertueux entre enseignement et évaluation. Comment en effet pourrait-on améliorer les compétences des enseignants en matière d'évaluation et d'enseignement sans les impliquer davantage dans le processus? Les résultats des enquêtes internationales semblent indiquer qu'il y a tout intérêt à impliquer davantage les enseignants dans tous les aspects de l'évaluation scolaire et à mieux comprendre leur démarche d'appropriation, leurs besoins de ressources et de formation.

La jonction réussie des politiques et des pratiques suppose un certain nombre de changements :

1. L'évaluation scolaire, qu'elle soit formative, sommative, intérimaire, doit posséder une validité de conséquence élevée quant à l'amélioration des apprentissages des élèves : les bénéfices apportés par l'évaluation scolaire devront dépasser les inconvénients lorsque ceux-ci existent ;
2. Les pratiques d'évaluation devront s'inscrire dans un système d'évaluation équilibrée et continue qui favorise des régulations à tous les niveaux du système éducatif ;
3. Le suivi des résultats d'évaluation devra être assuré individuellement et collectivement à travers des réseaux de communauté de pratique. Il devra se fonder sur un double alignement : à la fois sur le programme d'étude et sur les théories d'apprentissage.

La mobilisation des savoirs théoriques et d'expérience est au cœur de la jonction réussie entre politiques et pratiques. Plutôt que de s'interroger sur l'efficacité des modalités hiérarchiques de transmission entre politiques et pratiques, il semble plus profitable de se demander comment joindre politiques et pratiques dans un processus d'information mutuelle. Il est sans doute plus prometteur de se demander comment les enseignants peuvent, individuellement et collectivement, s'approprier et développer de nouvelles compétences pour améliorer les résultats du système d'éducation à tous les niveaux.

À cet égard, les systèmes d'éducation des pays francophones constituent des laboratoires privilégiés d'étude des différents facteurs influençant l'efficacité des modalités de mise en œuvre de nouvelles politiques et pratiques. Centrées principalement autour de la reformulation des objectifs d'apprentissage en termes de compétences et de la régulation de l'enseignement et de l'apprentissage au moyen de l'évaluation formative, ces expériences sont à même de développer un nouveau corpus de connaissances et de compétences, tant en matière d'évaluation que de politiques éducationnelles. Les articles du présent numéro thématique ouvrent la voie à un champ de recherche prometteur dans lequel les pays francophones auront tout avantage à s'investir pour assurer l'efficacité et la qualité globales de leurs systèmes d'éducation.

RÉFÉRENCES

- Allal, L. (1979). Stratégies d'évaluation formative : conceptions psychopédagogiques et modalités d'application. In L. Allal, J. Cardinet & P. Perrenoud (éds), *L'évaluation formative dans un enseignement différencié* (pp. 130-145). Berne : Peter Lang.
- Allal, L. (2007a). Réflexions transatlantiques sur les réformes scolaires. In L. Lafortune, M. Ettayebi & P. Jonnaert (éds), *Observer les réformes scolaires* (pp. 199-206). Québec : Presses de l'Université du Québec.
- Allal, L. (2007b). Régulation des apprentissages : orientations conceptuelles pour la recherche et la pratique en éducation. In L. Allal & L. Mottier Lopez (éds), *Régulation des apprentissages en situation scolaire et en formation* (pp. 7-23). Bruxelles : De Boeck.
- Allal, L., & 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), 99-106.
- Allal, L., & Mottier Lopez, L. (2005). Formative assessment of learning: A review of publications in French. In *Formative Assessment - Improving Learning in Secondary Classrooms* (pp. 241-264). Paris: OECD-CERI Publication (What works in innovation in education).
- Atkin, J.M. (2007). Swimming upstream: relying on teachers' summative assessments. *Measurement*, 5(1), 54-67.
- Black, P., & William, D. (1998). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.

- Bain, A. (1992). À propos des conditions d'une régulation pertinente et efficace. In D. Laveault (éd.), *Les pratiques d'évaluation en éducation*. Actes de colloque de la 15^e session d'études de l'ADMEE-Canada. Montréal : Éditions de l'ADMEE.
- Bodin, A. (1992). Évaluation et didactiques : distinguer ou intégrer? In D. Laveault (éd.), *Les pratiques d'évaluation en éducation*. Actes de colloque de la 15^e session d'études de l'ADMEE-Canada. Montréal : Éditions de l'ADMEE.
- Darling-Hammond, L., & 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.
- Figari, G. (2006). L'activité évaluative entre cognition et réponse sociale : nouveaux défis pour les évaluateurs. *Mesure et évaluation en éducation*, 29(1), 5-18.
- Fullan, M (2009). Large-scale reform comes of age. *Journal of Educational Change*, 10, 101-113.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *The Curriculum Journal*, 16(2), 207-223.
- Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: a seamless process in formative assessment? *Educational Measurement: Issues and Practice*, 28(3), 24-31.
- Huberman, A.M. (1983). *Comment s'opèrent les changements en éducation ; contribution à l'étude de l'innovation*. Paris : UNESCO.
- James, M. (2006) Assessment, teaching and theories of learning. In J. Gardner (éd.), *Assessment and Learning* (pp. 47-60). London: Sage.
- Kennedy, K. Sang Chan, J.K., Fok, P.K., & Yu, W.M. (2008). Forms of assessment and their potential for enhancing learning: conceptual and cultural issues. *Education Research Policy and Practice*, 7, 197-207.
- Lafortune, L., & Allal, L. (éds) (2008). *Jugement professionnel en évaluation : pratiques enseignantes au Québec et à Genève*. Québec : Presses de l'Université du Québec.
- Laveault, D. (2007). De la régulation au réglage : étude des dispositifs d'évaluation favorisant l'autorégulation des apprentissages. In L. Allal & L. Mottier Lopez (éds), *Régulation des apprentissages en situation scolaire et en formation* (pp. 207-234). Bruxelles : De Boeck.
- Laveault, D. (2008). Le jugement professionnel d'évaluation scolaire : enjeux, tensions et synergies nouvelles. *Revue suisse des sciences de l'éducation*, 30(3), 483-500.
- Laveault, D. (2009). L'amélioration de l'efficacité du système éducatif : sur quels indicateurs s'appuyer? In X. Dumay & V. Dupriez (éds), *L'efficacité dans l'enseignement : promesses et zones d'ombre* (pp. 177-194). Bruxelles : De Boeck.
- Laveault, D. (sous presse). Classroom Assessment in a Policy Context (French Sources), in *The International Encyclopedia of Education*, 3rd Edition, edited by Barry McGaw, Penelope Peterson and Eva Baker, Elsevier (in press).
- Maxwell, G. (2001). *Moderation of assessments in vocational education and training*. Brisbane, Queensland: Department of Employment and Training.
- McCloskey, W., & McMunn, N. (2003). Better Classroom Assessment in Support of Student Learning. *NCME Newsletter*, 11(3), 2-3.

- Messick, S. (1988). The once and the future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer & H.I. Braun, *Test Validity* (pp. 33-45). Hillsdale, NJ: Lawrence Erlbaum.
- Mottier Lopez, L., & Laveault D. (2008). L'évaluation des apprentissages en contexte scolaire : 30 ans de travaux. Développements, enjeux et controverses. *Mesure et évaluation en éducation*, 31(3), 5-34.
- Nichols, P.D., Meyers, J.L., & Burling, K.S. (2009a). A framework for evaluating and planning assessments intended to improve student achievement. *Educational Measurement: Issues and Practice*, 28(3), 14-23.
- Nichols, P.D., Meyers, J.L., & Burling, K.S. (2009b). A response to Shepard. *Educational Measurement: Issues and Practice*, 28(3), 38.
- Paquay, L. (2007). À quoi bon un curriculum de qualité s'il ne change pas les pratiques enseignantes! In M. Behrens (éd.), *La qualité en éducation* (pp. 57-98). Québec : PUQ.
- Perie, M., Marion, S., & Gong, B. (2009). Moving toward a comprehensive assessment system: a framework for considering interim assessment. *Educational Measurement: Issues and Practice*, 28(3), 5-13.
- Perrenoud, P. (2004). Obligation de compétence et analyse du travail : rendre compte dans le métier d'enseignant. In C. Lessard & P. Meirieu (éds), *L'obligation de résultats en éducation* (pp. 207-233). Sainte-Foy : Presses de l'Université Laval.
- Shepard, L. (2009). Commentary: evaluating the validity of formative and interim assessment. *Educational Measurement: Issues and Practice*, 28(3), 32-37.
- Stiggins, R.J. (2003). *Balanced Assessment: The Key to Accountability and Improved Student Learning*. National Education Association.
- Stiggins, R.J., Arter, J.A., Chappuis, J., & Chappuis, S. (2004). *Classroom assessment for student learning. Doing it right – Using it well*. Portland, OR: Assessment Training Institute.
- Tourneur, Y. (1985). La certification des compétences. *Mesure et évaluation en éducation*, 7(4), 5-20.
- Vienneau, R. (2005). *Apprentissage et enseignement, Théories et pratiques*. Montréal : Gaëtan Morin Éditeur.
- Wyatt-Smith, C., Klenowski, V. & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: a study of standards in moderation. *Assessment in Education: Principles, Policy & Practice*, 17(1), 59-75.